

**RECONHECIMENTO DE PADRÕES NO DIAGNÓSTICO DE  
DISTÚRBIOS, VOCAIS DE DOCENTES**

DIVANETE MARIA BITDINGER DE OLIVEIRA

**RECONHECIMENTO DE PADRÕES NO DIAGNÓSTICO DE DISTÚRBIOS  
VOCAIS DE DOCENTES**

Dissertação apresentada como requisito parcial à obtenção do grau de mestre em Ciências, do Programa de Pós-Graduação em Métodos Numéricos em Engenharia, na Área de Concentração em Programação Matemática, dos setores de Ciências Exatas e de Tecnologia da Universidade Federal do Paraná.

Orientadora

Prof<sup>a</sup> Dr<sup>a</sup> Maria Teresinha Arns Steiner

CURITIBA

2010

## **TERMO DE APROVAÇÃO**

DIVANETE MARIA BITDINGER DE OLIVEIRA

### **RECONHECIMENTO DE PADRÕES NO DIAGNÓSTICO DE DISTÚRBIOS VOCAIS DE DOCENTES**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciências, do Programa de Pós-Graduação em Métodos Numéricos em Engenharia, na Área de Concentração em Programação Matemática, dos setores de Ciências Exatas e de Tecnologia da Universidade Federal do Paraná.

Orientadora: \_\_\_\_\_

Profa. Maria Teresinha Arns Steiner, D. Eng.  
Departamento de Engenharia de Produção e PPGMNE, UFPR

\_\_\_\_\_  
Profa. Deise Maria Bertholdi Costa, D.Eng.  
Departamento de Expressão Gráfica e PPGMNE, UFPR

\_\_\_\_\_  
Prof. Jair Mendes Marques, D.Sc.  
Universidade Tuiuti do Paraná e PPGMNE, UFPR

\_\_\_\_\_  
Prof. Júlio Nievola, D.Eng.  
Programa de Pós-Graduação em Informática, PUCPR

**Curitiba, 20 de agosto de 2010.**

Aos meus pais, meus grandes exemplos, e  
ao meu esposo; por todo o amor, por quem  
sou e por tudo que alcancei.

## **AGRADECIMENTOS**

À Deus...

Se vencemos, alguém esteve conosco, se nada conseguimos, Ele continua junto de nós. Se persistirmos, veremos realmente que quem nos fez continuar, sorrirá para nós, mesmo que Dele, na felicidade, nos tenhamos esquecido.

Em Vossas mãos deposito meu futuro...

Aos pais e ao esposo...

Que compartilharam os meus ideais e os alimentaram, mantendo-se sempre ao meu lado, lutando comigo, dedico mais essa conquista e a mais profunda admiração e carinho.

Aos professores e colegas...

Vivemos tantas lutas juntos e delas carregamos a marca da experiência. Que tudo que aprendemos seja luz para nossos caminhos.

Á minha orientadora...

Agradeço-vos imensamente, por sua dedicação, paciência, incentivo e confiança ao longo dessa jornada; a sua sabedoria me impulsiona a ser cada vez melhor. Obrigada!

Á fonoaudióloga Sônia Maria Arns Guimarães...

Incansável na busca da excelência em seu trabalho, pela sua disponibilidade e prontidão. Pelo tempo que destinou coletando os dados e pelas suas orientações.

Ao Colégio Bom Jesus, pelo fornecimento do problema e dos dados, em especial, à fonoaudióloga do colégio, Sônia Maria Arns Guimarães.

Á minha colega e amiga Tereza Rachel Mafioleti, pela grande ajuda com o programa para realizar os testes das RNAs.

E a todos aqueles que de alguma forma contribuíram para esta dissertação tornar-se realidade, o meu MUITO OBRIGADO.

A pessoa verdadeiramente  
sábia é sempre humilde.

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>13</b>
1.1 OBJETIVO DO TRABALHO.....	14
1.2 ESTRUTURA DO TRABALHO.....	14
<b>2 ASPECTOS GERAIS SOBRE A FONOAUDIOLOGIA E OS ATRIBUTOS CONSIDERADOS NESTE TRABALHO .....</b>	<b>15</b>
2.1 INTRODUÇÃO.....	15
2.2 DADOS HISTÓRICOS.....	15
2.3 COMO SE PRODUZ A VOZ.....	16
2.4 FATORES DE RISCO PARA A VOZ.....	17
<b>3 KDD E DATA MINING.....</b>	<b>21</b>
3.1 INTRODUÇÃO.....	21
3.2 O PROCESSO KDD .....	23
3.3 TÉCNICAS DE DATA MINING.....	25
3.3.1 <i>Redes Neurais Artificiais (RNAs)</i> .....	26
3.3.1.1 Tipos de Aprendizado .....	32
3.3.1.2 Modelos de RNAs .....	32
3.3.2 <i>Função Discriminante Linear de Fisher</i> .....	43
3.3.2.1 Método de Fisher para duas populações .....	43
3.3.2.2 Método de Fisher para discriminar entre várias populações .	46
3.4 REVISÃO DA LITERATURA – TRABALHOS CORRELATOS .....	49
3.4.1 <i>Trabalhos que utilizaram técnicas semelhantes</i> .....	49
3.4.2 <i>Trabalhos relacionados a problemas de voz de docentes</i> .....	51
<b>4 MATERIAL E MÉTODOS.....</b>	<b>53</b>
4.1 ATRIBUTOS CONSIDERADOS NO TRABALHO.....	53
4.2 TRATAMENTO DOS DADOS .....	55
<b>5 IMPLEMENTAÇÃO DAS TÉCNICAS PARA RESOLUÇÃO DO PROBLEMA E ANÁLISE DOS RESULTADOS .....</b>	<b>57</b>
5.1 INTRODUÇÃO .....	57
5.2 REDES NEURAIS ARTIFICIAIS (RNAs) .....	58
5.2.1 <i>Teste I</i> .....	59
5.2.2 <i>Teste II</i> .....	60
5.2.3 <i>Teste III</i> .....	62
5.2.4 <i>Teste IV</i> .....	66
5.3 FUNÇÃO DISCRIMINANTE LINEAR DE FISHER (FDLF) .....	70

5.3.1	Teste V.....	70
5.3.2	Teste VI.....	70
5.4	ANÁLISE DOS RESULTADOS.....	71
<b>6</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>74</b>
6.1	CONCLUSÕES.....	74
6.2	SUGESTÕES PARA TRABALHOS FUTUROS.....	75
<b>7</b>	<b>REFERÊNCIAS.....</b>	<b>77</b>
<b>8</b>	<b>APÊNDICES.....</b>	<b>83</b>



## LISTA DE FIGURAS

FIGURA 3.1 – ETAPAS DO PROCESSO <i>KDD</i> .....	24
FIGURA 3.2 – NEURÔNIO ARTIFICIAL PROJETADO POR MCCULLOCH E PITTS.....	26
FIGURA 3.3 – FUNÇÃO IDENTIDADE.....	27
FIGURA 3.4 – FUNÇÃO LINEAR.....	28
FIGURA 3.5 - GRÁFICO DA FUNÇÃO LIMIAR.....	28
FIGURA 3.6 - FUNÇÃO SIGMÓIDE.....	29
FIGURA 3.7 – FUNÇÃO SIGMÓIDE, COM $a$ TENDENDO AO INFINITO.....	29
FIGURA 3.8 - FUNÇÃO SIGNUM.....	30
FIGURA 3.9 – FUNÇÃO TANGENTE HIPERBÓLICA.....	30
FIGURA 3.10 – PERCEPTRON.....	33
FIGURA 3.11 - MODELO DE NEURÔNIO.....	35
FIGURA 3.12 – REDE DE MÚLTIPLAS CAMADAS.....	35
FIGURA 3.13 - REDE DE MÚLTIPLAS CAMADAS.....	37
FIGURA 3.14 – REDE NEURAL DE BASE RADIAL.....	41
FIGURA 4.1 – ESPAÇO DISCRIMINANTE.....	73

## LISTA DE QUADROS

QUADRO 2.1 – ATRIBUTOS UTILIZADOS.....	54
QUADRO 2.2 – PONTUAÇÃO UTILIZADA PELA ESCOLA.....	55
QUADRO 4.1 – MATRIZ DE CONFUSÃO.....	58
QUADRO 4.2 – MATRIZ DE CONFUSÃO TESTE I.....	60
QUADRO 4.3 – MATRIZ DE CONFUSÃO TESTE II.....	61
QUADRO 4.4 – MATRIZ DE CONFUSÃO TESTE III(1 e 2,3,4).....	64
QUADRO 4.5 – MATRIZ DE CONFUSÃO TESTE III(2 e 1,3,4).....	65
QUADRO 4.6 – MATRIZ DE CONFUSÃO TESTE III(3 e 1,2,4).....	65
QUADRO 4.7 – MATRIZ DE CONFUSÃO TESTE III(3 e 1,2,4).....	65
QUADRO 4.8 – MATRIZ DE CONFUSÃO TESTE III(4 e 1,2,3).....	66
QUADRO 4.9 – MATRIZ DE CONFUSÃO TESTE IV(1 e 2,3,4).....	68
QUADRO 4.10 – MATRIZ DE CONFUSÃO TESTE IV(2 e 1,3,4).....	69
QUADRO 4.11 – MATRIZ DE CONFUSÃO TESTE IV(3 e 1,2,4).....	69
QUADRO 4.12 – MATRIZ DE CONFUSÃO TESTE IV(4 e 1,2,3).....	69
QUADRO 4.13 – MATRIZ DE CONFUSÃO TESTE IV(e).....	69
QUADRO 4.14 – MATRIZ DE CONFUSÃO TESTE V.....	70
QUADRO 4.15 – MATRIZ DE CONFUSÃO TESTE VI.....	71

## LISTA DE TABELAS

TABELA 4.1 – TESTE I.....	59
TABELA 4.2 – TESTE II.....	61
TABELA 4.3 – TESTE III.....	62
TABELA 4.4 – TAXA APARENTE DE ERRO TESTE III.....	66
TABELA 4.5 – TESTE IV.....	66
TABELA 4.6 – TAXA APARENTE DE ERRO TESTE IV.....	70
TABELA 4.7 – TESTE VI.....	71
TABELA 4.8 – MELHORES RESULTADOS.....	72

## RESUMO

A voz é uma das faculdades essenciais à comunicação humana; é adquirida e vai se formando através de nosso crescimento físico e emocional. Os profissionais da voz, em especial o professor, devido ao uso intensivo e por estar, em geral, despreparado para a grande demanda vocal diária, utilizam-na de forma inadequada em seu ofício, podendo produzir os chamados distúrbios vocais. Desta forma, é importante que esses profissionais sejam orientados a desenvolverem medidas preventivas visando a manter a qualidade da própria voz. Este trabalho tem como objetivo fazer a classificação quanto a gravidade de distúrbios vocais de docentes de uma escola localizada no município de Curitiba, PR. O interesse da escola com relação a esta classificação está no fato de que o docente com distúrbio classificado como leve poderá aumentar a sua carga horária, se assim o desejar; já o docente com distúrbio classificado como grave, terá que reduzir a sua carga horária, visando a sua saúde e bem-estar. Para tanto, foram utilizados dados de 100 docentes desta escola, com 10

informações (atributos) de cada um deles e, também, as suas respectivas classificações, fornecidos pela especialista da área, a fonoaudióloga da escola. Para alcançar o objetivo proposto, foi aqui utilizado o processo *KDD (Knowledge Discovery in Databases* ou Descoberta de Conhecimento em Bases de Dados), usando na etapa de *Data Mining* (ou Mineração de Dados), duas técnicas de Reconhecimento de Padrões, Redes Neurais Artificiais (RNAs) e a Função Discriminante Linear de Fisher (FDLF) para realizar a classificação dos padrões (docentes). As referidas técnicas (RNAs e FDLF) foram treinadas, comparativamente, para que forneçam com a máxima acurácia possível, a classificação para cada um dos padrões. Assim, dado um novo padrão (novo docente) com seus atributos, pode-se utilizar a técnica treinada que tenha apresentado o melhor desempenho (RNAs ou FDLF) para classificar este novo padrão, sendo que, desta forma, a especialista terá um respaldo adicional para o seu diagnóstico. As técnicas utilizadas mostraram-se de grande eficiência, visto que apresentaram uma margem de erro aceitável e resultados consistentes. Nesta comparação, as RNAs mostraram um maior desempenho, apresentando um resultado de 100% de acerto, no teste IV, contra 84,61% de acerto, no teste VI, da FDLF.

## ABSTRACT

The voice is one of the essential powers to human communication, it is acquired and it is being built through our physical and emotional growth. The voice's professionals, especially the teacher, due to intensive use of the voice and generally for being unprepared for the daily vocal demand, use it inappropriately in their work, and it can cause the so-called vocal disorders. Thus, it is important these professionals to be advised to develop preventive measures that objective to maintain the quality of their own voice. This work aims to classify the severity of vocal disorders in teachers that work in a school located in the city of Curitiba, PR. The interest of the school in this classification is related to the fact that teachers with a disorder classified as mild may increase their workload, if they wish; however, teachers with a disorder classified as grave will have to reduce their workload, aiming their health and wellbeing. For this, data from 100 teachers of that school were used, with 10 information (attributes) from each of them and also their respective classifications, provided by the area specialist, the school's speech therapist. To achieve the proposed objective, in this study were used the process called KDD (Knowledge Discovery in Databases or Knowledge Discovery in Databases), used in the stage of Data Mining, two techniques of Patterns Recognition, Artificial Neural Networks (RNAs) and the Fisher Linear Discriminant Function (FDLF) to perform patterns' classification (teachers). These techniques (RNAs and FDLF) were trained, comparatively, to provide with the utmost possible accuracy the rating for each of the standards. Thus, with a new standard (new teacher) with its attributes it is possible to use the trained technique that has shown the best performance (RNAs or FDLF) to classify this new standard, in a way the expert will have an additional support for its diagnosis. The techniques used have been very efficient once they showed an acceptable margin of error and consistent results. In this comparison, the RNAs showed a higher performance, which represents an result of 100% accuracy in test IV, compared with 84,61% accuracy, in test VI of FDLF.

## CAPÍTULO I

### 1. INTRODUÇÃO

Com o aumento na quantidade de dados a serem armazenados para análise, que superam as habilidades humanas de interpretá-los, surgiu a necessidade de técnicas que analisem esses dados de forma inteligente.

O processo *KDD* (*Knowledge Discovery in Data Bases*) tem esse objetivo, transformar esses dados armazenados em conhecimento. Esse processo é dividido em cinco etapas, onde a principal chama-se *Data Mining* ou Mineração de Dados, uma vez que é nesta que são aplicados os algoritmos para extrair modelos dos dados.

No problema aqui apontado, têm-se 10 informações ou dados de cada um dos 100 docentes, donde nessas 10 informações (critérios), têm-se ainda 33 variáveis. Diante dessa quantidade de elementos, é interessante ter um subsídio para a profissional, na classificação dos docentes quanto a sua voz.

A voz é um importante instrumento de comunicação e, para os docentes, ocupa lugar de destaque, já que influencia a relação entre o professor e os alunos. Porém, há uma grande predominância de distúrbios vocais nessa classe de trabalhadores, devido as condições precárias de trabalho e, conseqüentemente, ao mau uso da voz.

Dessa forma, o objetivo da fonoaudiologia é o retorno a uma função vocal normal, e para isto todos os fatores ou causas devem ser analisados.

Na etapa de *Data Mining* foram utilizadas as Redes Neurais Artificiais (RNAs) e a Função Discriminante Linear de Fisher (FDLF), com a finalidade de classificar padrões.

As RNAs apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência, e geralmente são empregadas no reconhecimento de padrões; e a FDLF, tem como

alvo encontrar “discriminantes” cujos valores numéricos sejam tais que as populações sejam separadas tanto quanto possível.

### 1.1 OBJETIVO DO TRABALHO

O principal objetivo desse trabalho é encontrar uma técnica que consiga classificar o grau de distúrbio vocal de um docente com a maior precisão possível. Deste modo, ter-se-á um respaldo adicional ao diagnóstico do especialista na área, no caso, a fonoaudióloga, em cumprir para com o interesse da escola de prevenir e detectar precocemente possíveis disfonias nos docentes.

Buscando alcançar esse objetivo, foram utilizadas duas técnicas, as RNAs e a FDLF, com a intenção de comparar seus resultados, verificando qual dentre elas oferece o menor percentual de erros.

### 1.2 ESTRUTURA DO TRABALHO

O presente trabalho está dividido em mais quatro capítulos, além deste capítulo introdutório.

O capítulo II descreve os aspectos gerais sobre a fonoaudiologia e os distúrbios de voz.

No capítulo III são apresentados, em detalhes, os conceitos de *KDD* e *Data Mining*, de RNAs e da FDLF, técnicas utilizadas neste trabalho e os trabalhos correlatos.

O capítulo IV expõe os atributos considerados neste trabalho e o tratamento dos dados.

No capítulo V é apresentada a aplicação das técnicas estudadas e implementadas ao problema. Também é apresentada a análise e discussão dos resultados obtidos.

Finalmente, no capítulo VI, são apresentadas as conclusões do trabalho, baseadas na análise dos resultados, além de sugestões para trabalhos futuros.

## CAPÍTULO II

### **2 ASPECTOS GERAIS SOBRE A FONOAUDIOLOGIA E OS ATRIBUTOS CONSIDERADOS NESTE TRABALHO**

#### 2.1 INTRODUÇÃO

Com o intuito de melhor posicionar o problema aqui abordado (Critérios para Classificação de Gravidade de Distúrbios Vocais) no contexto da área de Fonoaudiologia, são apresentados neste capítulo, alguns aspectos gerais sobre esta área, assim como os atributos selecionados por especialistas, para o desenvolvimento do trabalho.

#### 2.2 DADOS HISTÓRICOS

Segundo PINHO (1998) tem-se que “A especialidade de voz estava restrita aos professores de canto e às professoras de oratória, sem estar ainda voltada especificamente para as patologias laríngeas. Os terapeutas que se dedicavam ao trabalho de voz o faziam de uma forma mais empírica, tentando suprir a falta de informação mais específica e, em consequência, os tratamentos eram mais demorados e os resultados nem sempre satisfatórios”.

O grande impulso desse campo da voz deu-se graças ao trabalho conjunto e integrado entre os profissionais da área da Medicina e da Fonoaudiologia. Segundo BEHLAU e PONTES (1995), somente na segunda metade do séc. XIX e no início do séc. XX, aparecem artigos científicos descrevendo casos clínicos tratados por fonoterapia. Esses avanços têm auxiliado em proporcionar terapias mais eficazes.

Sendo assim, o objetivo da terapia de voz é o retorno a uma função vocal normal, não se podendo atuar sobre parâmetros vocais independentemente. Segundo BEHLAU e PONTES (1995), um tratamento bem-sucedido depende de um



trabalho em equipe, onde o paciente é o primeiro elemento dessa equipe, ou seja, é de grande importância a colaboração do paciente em seu próprio tratamento.

Assim, a avaliação da voz é dividida em três partes: a anamnese, a avaliação do comportamento vocal e o exame otorrinolaringológico. A anamnese é a identificação pessoal, onde a queixa é o motivo da consulta. Ela revela o grau de conscientização do paciente sobre sua alteração, sendo que a impressão do paciente sobre sua própria voz reflete não somente o grau de consciência sobre a emissão, mas também a auto-imagem e as fantasias sobre seus recursos vocais, obrigando o paciente à mudança de hábitos de fala se não gostar de sua voz.

A avaliação do comportamento vocal é uma investigação básica sobre os principais parâmetros empregados pelo indivíduo em sua comunicação habitual e, o exame otorrinolaringológico é quando o paciente procura um otorrinolaringologista.

### 2.3 COMO SE PRODUZ A VOZ

A voz é uma das faculdades essenciais à comunicação humana, afinal, é por meio dela que o homem transmite seus conhecimentos, vontades, sentimentos, emoções e pensamentos.

Segundo BEHLAU e PONTES (1995), a voz se produz na laringe que é um tubo que contém as pregas vocais. Ao inspirar, o ar entra nos pulmões e as pregas vocais se afastam. Ao expirar, o ar sai dos pulmões e, passando pela laringe, põe em vibração as pregas vocais. Nesse momento, é produzido um som de pequena intensidade que vai ser amplificado nas cavidades de ressonância (laringe, faringe, boca e nariz), se transformando nos diversos sons da fala através dos movimentos dos órgãos miofuncionais (boca, língua, lábios, bochechas, mandíbula, dentes e palato). Assim sendo, para que o som seja produzido com boa qualidade, há necessidade da integridade do sistema respiratório, sistema digestivo e de toda musculatura envolvente neste processo.

Portanto, a voz é adquirida e vai se formando através de nosso crescimento físico e emocional. Segundo PINHO (1997), principalmente os profissionais da voz (professores, locutores, cantores, entre outros) utilizam atividades verbais

inadequadas em sua profissão, podendo danificar os tecidos da laringe e produzir um distúrbio vocal resultante do abuso ou mau uso da voz.

Segundo GUIMARÃES (2001), dentre todos os profissionais da voz, o professor é o mais suscetível às alterações vocais, devido ao uso intensivo de sua voz e por estar despreparado para a grande demanda vocal diária. Por isso, orientações sobre o mecanismo de produção da voz, noções de higiene e técnica vocal são indispensáveis para auxiliar esses profissionais na manutenção da qualidade da própria voz.

Dessa forma, é importante determinar o comportamento vocal do professor, as situações de abuso vocal e os possíveis agentes prejudiciais à voz, assim como a influência de fatores psicológicos e de estresse, pois todos esses dados serão importantes para a orientação e o tratamento das alterações vocais.

## 2.4 FATORES DE RISCO PARA A VOZ

Os fatores de risco para distúrbios vocais, segundo PINHO (1997 e 1998), basicamente são os seguintes:

- não hidratar o organismo e falar em ambientes secos e empoeirados: o profissional da voz deve tomar muita água, pois pregas vocais hidratadas tem melhor flexibilidade e vibração. Segundo BEHLAU e PONTES (1995), a hidratoterapia (ingestão de água) é o melhor coadjuvante no tratamento das disfonias funcionais e orgânico-funcionais, por reduzir a viscosidade do muco da laringe e oferecer uma maior flexibilidade de vibração da mucosa;
- gritar sem suporte vibratório: é uma das atitudes mais agressoras para as pregas vocais. No grito ocorre uma verdadeira “tombada” entre as pregas vocais, momento em que o ar pode chegar a passar a 80 km/h na região glótica;
- golpe de glote (falar com ataques vocais bruscos): uma prega vocal é golpeada contra a outra;

- tossir ou pigarrear excessivamente: ao pigarrear uma prega vocal é “raspada” na outra, criando uma região de atrito na glote. As pregas vocais reagem a esse atrito constante produzindo muco para se protegerem contra o impacto, e isso se torna um círculo vicioso, pois o muco atrapalha a emissão vocal, forçando o indivíduo a pigarrear novamente. Quanto ao tossir, é um abuso enorme para as pregas vocais, que ficam batendo com força uma contra a outra;
- falar em ambientes ruidosos ou abertos (competição vocal): em ambientes ruidosos fala-se mais alto, pois perde-se o retorno auditivo da própria voz, e em ambientes abertos, porque a onda sonora se dissipa com mais facilidade. Isso rende um desgaste vocal muito maior;
- utilizar tom grave ou agudo demais: cada indivíduo possui uma extensão vocal diferente que deve ser respeitada, que é o número de notas da mais grave à mais aguda que o indivíduo pode produzir, não importando a qualidade (varia de 13 a 44 semi-tons);
- falar excessivamente durante quadros gripais ou crises alérgicas: pode ser extremamente prejudicial, causando danos irreversíveis à mucosa que reveste as pregas vocais;
- praticar exercícios físicos falando: pois durante o exercício físico ocorre um aumento no fechamento das cordas vocais e ao falar, pode gerar uma sobrecarga;
- fumar ou falar muito em ambientes de fumantes: o cigarro é altamente irritante às mucosas do trato vocal e, a fumaça, resseca as pregas vocais, dificultando sua vibração, além de agredir todo o sistema respiratório e trato vocal. Pode levar à irritação do trato vocal, edema em pregas vocais e proporcionar o aparecimento de pigarro e de tosse em decorrência do aumento da secreção;
- utilizar álcool em excesso: o álcool, além de irritante, tem um efeito anestésico, que mascara a dor de garganta, permitindo que vários abusos vocais sejam cometidos sem serem percebidos. Sprays e pastilhas também podem apresentar efeitos semelhantes;

- falar ou cantar abusivamente em período pré-menstrual: pois nesse período ocorre um aumento seguido de uma redução do nível de estrogênio no corpo da mulher, o que causa inchaço em várias regiões do corpo, assim como nas pregas vocais, o que torna mais prejudicial abusar da voz;
- rir alto: visto que geralmente o riso é apoiado na musculatura respiratória;
- falar muito após ingerir grandes quantidades de aspirinas, calmantes ou diuréticos: a aspirina resulta no aumento da circulação sanguínea na periferia das pregas vocais, conseqüentemente, com muita fala há o extravazamento de sangue para fora dos tecidos; os diuréticos e calmantes ressecam as mucosas e causam imprecisão articulatória;
- discutir com freqüência: pois a pessoa perde o controle por estar sob domínio das emoções e comete abusos vocais;
- cantar inadequada ou abusivamente e fazer parte de corais sem preparo vocal: o canto sem preparo e técnica pode ser extremamente prejudicial, causando sérios distúrbios orgânicos secundários;
- alimentação: evitar alimentos que causam azia, má digestão e refluxo de substâncias gástricas, que podem banhar as pregas vocais causando irritação das mesmas; não se alimentar muito próximo às performances, pois impede a livre movimentação do diafragma dificultando a fonação, da mesma forma que ingerir alimentos muito pesados e condimentados. Além disso, evitar os derivados do leite, que aumentam e “engrossam” a secreção do trato vocal, e o café e chá preto, que ressecam a mucosa das pregas vocais.

Por outro lado, comer uma maçã antes da atividade profissional da fala limpa a boca e a faringe, além de exercitar a mastigação relaxando a musculatura da boca. (BEHLAU e PONTES 1995).

- ar condicionado: resseca toda a mucosa do trato vocal, pois diminui a umidade do ar;
- vestuário: evitar roupas muito justas, que apertem especialmente a região do pescoço e do abdome, pois causam dificuldades no movimento respiratório e na liberdade da movimentação muscular.

Além disso, as atividades mais indicadas para quem utiliza a voz profissionalmente são natação, caminhadas, ginástica sem impacto, exercícios de alongamento e ioga, onde não há nenhuma movimentação violenta que causaria tensão muscular nas regiões do pescoço, ombro, tórax e costas.

## CAPÍTULO III

### 3 KDD E DATA MINING

#### 3.1 INTRODUÇÃO

Nessas últimas décadas, grandes empresas, hospitais, instituições de ensino e outros, vêm armazenando imensas quantidades de dados, para análise e interpretação, objetivando melhorar a eficiência e eficácia de seus processos. Desta forma, surge a necessidade de se gerar técnicas e ferramentas que transformem esses dados em conhecimento, visto que essas quantidades superam a capacidade humana de entendimento. Segundo FAYYAD *et al.* (1996), esse é o objetivo do processo denominado *Knowledge Discovery in Databases* (*KDD* ou Descoberta de Conhecimento em Bases de Dados), criado em 1989.

Muitos trabalhos têm sido desenvolvidos com a finalidade de analisar, discriminar e classificar dados, identificando as potencialidades e limitações dos diversos métodos. Em geral, estes métodos recebem uma amostra de dados (pontos, observações, padrões, exemplos, instâncias) definidos por um conjunto de variáveis escolhidas e geram uma função discriminante que serve como classificador.

O processo *KDD* é um conjunto de atividades contínuas que compartilham o conhecimento descoberto a partir de bases de dados. Segundo FAYYAD *et al.* (1996), esse conjunto é composto de cinco etapas: seleção dos dados; pré-processamento e limpeza dos dados; transformação dos dados; Mineração de Dados (*Data Mining* ou reconhecimento de padrões - principal etapa, núcleo deste trabalho); interpretação e avaliação dos resultados.

Até 1995, muitos autores consideravam os termos *KDD* e *Data Mining* como sinônimos, porém o processo *KDD* refere-se a todo processo de descoberta de conhecimento útil nos dados, enquanto *Data Mining*, que é uma etapa do processo *KDD*, refere-se à aplicação de algoritmos para extrair modelos dos dados. Segundo

FREITAS (2000), o método de descoberta do conhecimento deve apresentar as seguintes características: ser eficiente, genérico (aplicável a vários tipos de dados) e flexível (facilmente modificável).

*Data Mining* é uma metodologia para melhorar a qualidade e eficiência das decisões, e as empresas, de um modo geral, buscam formas de acrescentar algum diferencial mercadológico, proporcionando uma maior rentabilidade, eficiência e segurança nas suas relações comerciais ou prestação de serviços.

As técnicas de *Data Mining* têm a finalidade de classificar ou agrupar os dados ou, ainda, descobrir regras de associação entre os dados. Dentre os métodos de *Data Mining*, pode-se citar os algoritmos que geram as árvores de decisão, as máquinas de suporte de vetores (*Support Vector Machines, SVM*), os métodos estatísticos, e as metaheurísticas (redes neurais, algoritmos genéticos; busca tabu e outras).

No presente trabalho será utilizado o processo *KDD*, onde na etapa de *Data Mining* será empregado um Método Estatístico, Função Discriminante Linear de Fisher (FDLF) e, também, um Método da área de Inteligência Artificial, técnica das Redes Neurais Artificiais (RNAs).

O método estatístico abordado nesse trabalho, a FDLF, pode ser apresentado da seguinte forma: dadas duas populações de observações multivariadas com uma certa dimensão  $n$ , a idéia de Fisher é transformar estas observações multivariadas em observações univariadas tal que estejam separadas tanto quanto possível. Deve-se determinar variáveis que melhor discriminem esses grupos, utilizando essas variáveis para criar funções discriminantes que serão utilizadas para alocar novos indivíduos, objetos ou observações no grupo mais adequado (a função discriminante otimiza a alocação de novos padrões). Outro fator importante é que esse método é de fácil cálculo, pois utiliza a combinação linear das observações multivariadas para criar as observações univariadas.

Já as RNAs são técnicas computacionais que adquirem conhecimento através da experiência. Têm uma grande capacidade de processamento e de armazenamento de informação, simulando o cérebro humano e o comportamento humano, ou seja, aprendendo, errando e fazendo descobertas.

A propriedade mais importante das RNAs é a habilidade de aprender a partir de seu ambiente e com isso melhorar o seu desempenho. Modelos de RNAs podem lidar com dados imprecisos e situações não totalmente definidas. Uma rede treinada tem a habilidade de generalizar quando é apresentada à entradas que não estavam presentes nos dados já conhecidos por ela, isso é feito através de um processo iterativo de ajustes aplicado a seus pesos, o treinamento. O aprendizado ocorre quando a rede RNA atinge uma solução generalizada para certa classe de problemas.

Tem-se, resumidamente, que para aplicar os dois métodos (RNAs e FDLF), as seguintes etapas devem ser tratadas: coleta de dados relativos ao problema; separação destes dados em conjuntos de treinamento (dados que serão usados para treinar os métodos) e conjuntos de testes (dados que serão usados para testar os referidos métodos, verificando o aprendizado das técnicas, ou seja, a capacidade de generalização das mesmas). Com os métodos treinados e avaliados, eles poderão ser integrados em um sistema do ambiente operacional da aplicação, ou seja, pode-se apresentar aos mesmos novos padrões até então desconhecidos para obter a sua classificação.

### 3.2 O PROCESSO *KDD*

*KDD* é um processo que extrai informações das base de dados, que cria relações importantes normalmente não observadas pelo especialista no assunto, além de auxiliar na validação do conhecimento extraído (WIKIPÉDIA, 2009). Enquanto os métodos tradicionais são capazes de tratar apenas as informações explícitas, a extração de conhecimento é capaz de detectar informações implícitas armazenadas nos bancos de dados.

O processo de *KDD* é composto de cinco etapas (LEMOS, 2003):

- Seleção dos dados;
- Pré-processamento e limpeza dos dados;
- Transformação dos dados;



- *Data Mining*;
- Interpretação e avaliação dos resultados.

Essas etapas podem ser observadas na figura 3.1, onde as três primeiras podem ser interpretadas como a análise exploratória dos dados.

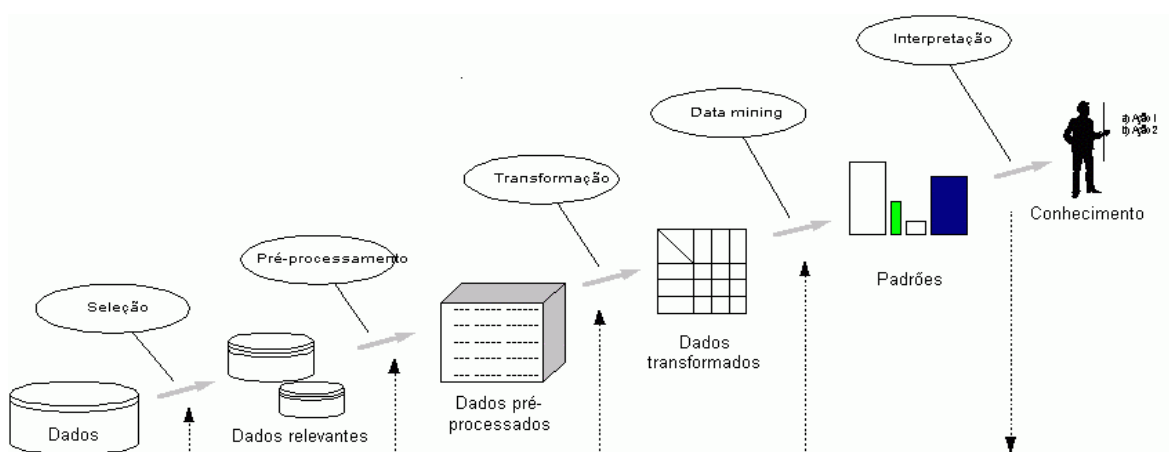


FIGURA 3.1 – ETAPAS DO PROCESSO *KDD*. FONTE: FAYYAD *et al.* (1996)

Assim, segundo LEMOS (2003), o processo de *KDD* começa com o entendimento do domínio da aplicação e dos objetivos finais a serem atingidos. A partir dessa etapa, os dados são selecionados de acordo com os critérios definidos; na etapa de pré-processamento ou limpeza dos dados, são removidas as informações julgadas desnecessárias.

Os dados pré-processados devem ainda passar por uma transformação que os armazena adequadamente, facilitando o uso das técnicas de *Data Mining*. Após essa etapa chega-se à fase de *Data Mining*, que começa com a escolha das ferramentas (algoritmos) a serem utilizadas, escolha essa que depende fundamentalmente do objetivo do processo de *KDD*: classificação, agrupamento, associação ou outras.

Essas ferramentas escolhidas procuram os padrões de comportamento dos dados. Segundo LEMOS (2003), essa busca pode ser efetuada automaticamente pelo sistema, vasculhando o banco de dados de forma livre (*roams*), ou interativamente com um analista, responsável pela geração de hipóteses, (*directed analysis* - análise direcionada - ou *supervised learning* - aprendizado

supervisionado), onde o analista diria, por exemplo, quando uma premissa foi ou não correta.

Muitas ferramentas podem ser aplicadas ao problema nessa fase, tanto isoladamente quanto em combinação. Geralmente, o processamento de busca é interativo, ou seja, os analistas revêem o resultado, formam um novo conjunto de questões com o objetivo de refinar a busca em um determinado aspecto das descobertas e realimentam o sistema com novos parâmetros.

Ao final do processo, o sistema de *Data Mining* gera um relatório das descobertas, que será interpretado pelos analistas. Após essa interpretação é que se encontra conhecimento. Pelo fato das técnicas de *Data Mining* não contarem com a experiência e intuição humana para reconhecer a diferença entre uma correlação relevante ou irrelevante, analistas serão sempre necessários para avaliar modelos e aprová-los.

Existem técnicas tradicionais que verificam os conhecimentos armazenados nas bases de dados, porém, segundo LEMOS (2003), aproximadamente 5% de todas as relações podem ser encontradas por esses métodos, enquanto *Data Mining* pode descobrir as outras 95%. *Data Mining* explora as bases de dados através de centenas de pontos de vista diferentes, assim toda informação escondida será enfatizada.

Ainda segundo LEMOS (2003), o uso de *Data Mining* para construção de um modelo traz as seguintes vantagens: modelos de fácil compreensão, grandes bases de dados podem ser analisadas, descobre informações não esperadas, as variáveis não necessitam de recodificação e os modelos são precisos.

### 3.3 TÉCNICAS DE *DATA MINING*

Como já citado anteriormente, a fase de *Data Mining* começa com a escolha das ferramentas (algoritmos) a serem utilizadas, escolha essa que depende do objetivo do processo de *KDD*: classificação, agrupamento, regras associativas ou outras.

Cada tipo de tecnologia tem suas vantagens e desvantagens, assim nenhuma ferramenta consegue atender a todas as necessidades em todas as aplicações. Nas seções seguintes, são descritas, particularmente, as ferramentas utilizadas nesse trabalho, que são as Redes Neurais Artificiais e Função Discriminante Linear de Fisher.

### 3.3.1 Redes Neurais Artificiais (RNAs)

Baseado nos estudos do cérebro, nas idéias de redes de neurônios e nos modelos dos neurônios, foram propostas as estruturas de redes neurais, ou redes de neurônios artificiais, no entanto, uma RNA pode ter centenas ou milhares de unidades de processamento, enquanto o cérebro de um mamífero possui bilhões de neurônios (OLIVEIRA, 2002).

Segundo TONSIG (2000), as primeiras informações sobre neurocomputação surgiram em 1943, em artigos do neurofisiologista Warren McCulloch, do Instituto Tecnológico de Massachusetts e do matemático Walter Pitts, da Universidade de Illinois. Ambos fizeram uma analogia entre células nervosas vivas e o processo eletrônico, simulando o comportamento do neurônio natural, no qual o neurônio possuía apenas uma saída, que era uma função da soma de valor de suas diversas entradas, conforme ilustrado na figura 3.2.

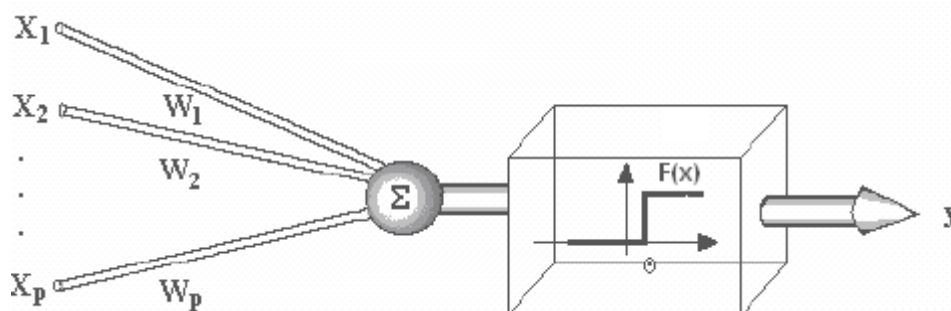


FIGURA 3.2 – NEURÔNIO ARTIFICIAL PROJETADO POR MCCULLOCH E PITTS.

FONTE: TONSIG (2000)

Nesta figura 3.2, o vetor  $x$  representa um conjunto de  $p$  entradas, que multiplicado por um vetor peso  $w$  nos dá o produto  $p = x w$ , aplicado aos canais de entrada do neurônio. A função de ativação,  $F(x)$ , que produz o sinal de saída  $y$  do neurônio (binária ou contínua) é dada por:

$$y = F\left(\sum_0^p x_i w_i\right), \text{ para } x_0 = 1 \text{ e } w_0 = \theta \quad (3.1)$$

Onde o parâmetro  $\theta$  é o bias ou vício, que aumenta o número de graus de liberdade disponíveis no modelo, permitindo que a RNA tenha maior capacidade de se ajustar ao conhecimento a ela fornecido. Segundo PAVANELLI (2007), *bias* é uma entrada adicional acrescentada ao neurônio artificial, cujo valor inicial é sempre igual a 1, sendo ajustado, assim como os demais pesos da rede.

Existem várias funções de ativação, entre elas: função identidade, função linear, função limiar, função sigmoidal, função signum, função tangente hiperbólica, entre outras, representadas a seguir.

- *Função Identidade*

Segundo WIKIPÉDIA (2009), essa função dá como imagem de cada elemento o próprio elemento. É da forma (3.2):

$$f(x) = x \quad (3.2)$$

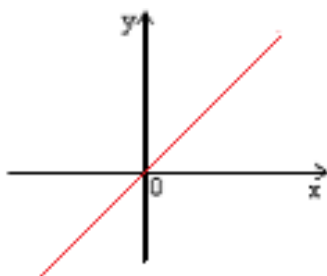


FIGURA 3.3 – FUNÇÃO IDENTIDADE

- *Função Linear*

Segundo STEINER (2008), essa função elimina a descontinuidade em  $x = \theta$ . É da forma (3.3):

$$\begin{cases} f(x) = 0, x < b \\ f(x) = ax, b \leq x \leq c \\ f(x) = 1, x > c \end{cases} \quad (3.3)$$

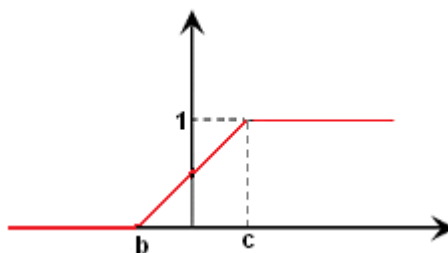


FIGURA 3.4 – FUNÇÃO LINEAR POR PARTES. FONTE: GUIMARÃES (2009)

- *Função Limiar (Degrau ou Passo)*

Segundo LNCC (2009), essa função foi utilizada no modelo de McCulloch e Pitts. Ela modela a característica "tudo-ou-nada" deste neurônio e é expressa da seguinte forma (3.4):

$$f(v) = \begin{cases} 1, & \text{se } v \geq 0; \\ 0, & \text{se } v < 0; \end{cases} \quad (3.4)$$

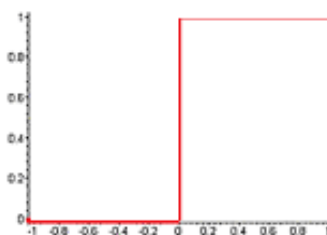


FIGURA 3.5 - GRÁFICO DA FUNÇÃO LIMIAR. FONTE: LNCC (2009)

- *Função Sigmoidal*

Essa função pode assumir valores no intervalo (0, 1) e, segundo LNCC (2009), a representação mais utilizada para essa função é a função logística, definida pela equação (3.5)

$$f(v) = \frac{1}{1 + e^{(-av)}} \quad (3.5)$$

onde  $a$  é o parâmetro de inclinação da função sigmóide e  $v$  é o valor de ativação do neurônio.

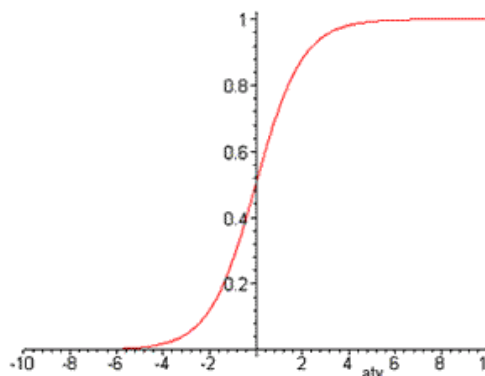


FIGURA 3.6 - FUNÇÃO SIGMÓIDE. FONTE: LNCC (2009)

Ainda segundo LNCC (2009), quando aumentamos o valor do parâmetro  $a$ , tendendo-o ao infinito, esta função comporta-se como uma função de limiar, conforme figura 3.7 a seguir.

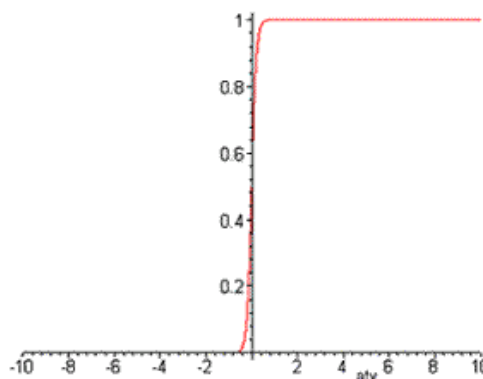


FIGURA 3.7 – FUNÇÃO SIGMÓIDE, COM  $a$  TENDENDO AO INFINITO. FONTE: LNCC (2009)

- *Função Signum*

Essa função está limitada ao intervalo  $(-1, 1)$  e apresenta as mesmas características da Função Limiar. É representada pela equação (3.6).

$$f(v) = b \frac{v}{|v|} \text{ para } v \neq 0 \quad (3.6)$$

onde  $b$  são os limites inferiores e superiores ( $b=|1|$ , no gráfico), e  $v$  é o valor de ativação da unidade.

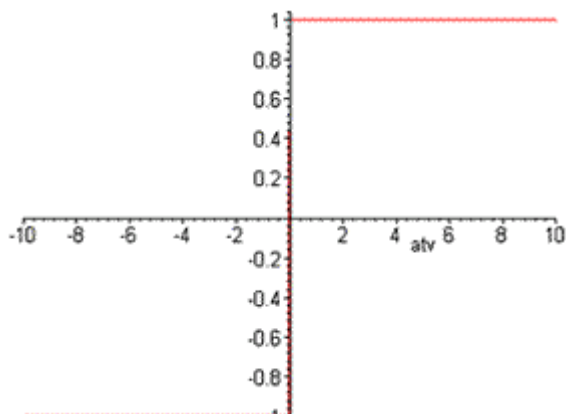


FIGURA 3.8 - FUNÇÃO SIGNUM. FONTE: LNCC (2009)

- *Função Tangente Hiperbólica*

Representada pela equação (3.7).

$$f(v) = a \frac{e^{(bv)} - e^{(-bv)}}{e^{(bv)} + e^{(bv)}} \quad (3.7)$$

onde  $a$  é o parâmetro de inclinação da curva;  $b$  são os limites inferiores e superiores ( $b=|1|$ , no gráfico) e  $v$  é o valor de ativação da unidade.

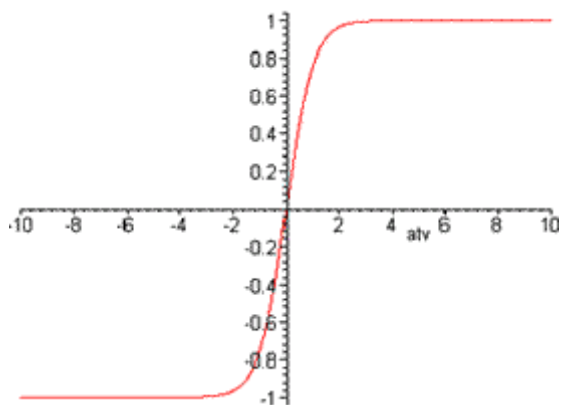


FIGURA 3.9 – FUNÇÃO TANGENTE HIPERBÓLICA. FONTE: LNCC (2009)

As RNAs apresentam desempenho satisfatório em diversas tarefas, dentre as quais pode-se citar: reconhecimento de padrões; classificação de padrões; aproximação de funções e previsão de séries temporais (WIKIPÉDIA, 2009).

A RNA se baseia nos padrões apresentados para extrair um modelo geral, portanto, a fase de aprendizado deve ser rigorosa a fim de se evitar modelos espúrios. “Os dados coletados devem ser significativos e cobrir amplamente o domínio do problema; não devem cobrir apenas as operações normais ou rotineiras,

mas também as exceções e as condições nos limites do domínio do problema” (TATIBANA, 2009).

Todo o conhecimento de uma RNA está armazenado nas sinapses, ou seja, nos pesos atribuídos às conexões entre os neurônios. Segundo TATIBANA (2009), 50 a 90% do total de padrões deve ser separado para o treinamento da RNA, escolhidos aleatoriamente para que a rede realmente "aprenda" as regras. O restante dos dados só é apresentado à RNA na fase de testes, a fim de averiguar a capacidade de generalização da rede, ou seja, o quanto a rede, de fato, "aprendeu".

Ainda segundo TATIBANA (2009), existem seis passos necessários para o desenvolvimento de aplicações utilizando RNAs.

Os dois primeiros passos do processo para o treinamento de RNAs são, conforme já comentado, a coleta de dados relativos ao problema e a sua separação em um conjunto de treinamento e um conjunto de testes.

O terceiro passo é a definição da configuração da rede, que pode ser dividida em três etapas: seleção do paradigma neural apropriado à aplicação, determinação da topologia da rede a ser utilizada - o número de camadas e o número de unidades em cada camada - e a determinação de parâmetros do algoritmo de treinamento e funções de ativação. Este passo tem um grande impacto no desempenho do sistema resultante.

O quarto passo é o treinamento da rede. Nesta fase serão ajustados os pesos das conexões, onde geralmente, seus valores iniciais são números aleatórios uniformemente distribuídos, em um intervalo definido. Uma escolha adequada dos valores iniciais dos pesos da rede pode diminuir o tempo necessário para o treinamento, e a escolha errada destes pesos pode levar a uma saturação prematura.

O treinamento deve ser interrompido quando a rede apresentar uma capacidade apropriada de generalização e quando a taxa de erro for suficientemente pequena, ou seja, menor que um erro admissível. Assim, deve-se encontrar um ponto ótimo de parada com erro mínimo e capacidade de generalização máxima.



O quinto passo é o teste da rede, onde seu desempenho é uma boa indicação de seu desempenho real. Durante esta fase, o conjunto de teste é utilizado para determinar o desempenho da rede com padrões que não foram utilizados durante o treinamento.

Finalmente, com a rede treinada e avaliada, pode-se integrá-la em um sistema do ambiente operacional da aplicação. Esse é o sexto passo, integração.

#### 3.3.1.1 Tipos de Aprendizado

Segundo ALECRIM (2004), os tipos de aprendizado das RNAs são basicamente três: supervisionado, não-supervisionado e híbrido.

O aprendizado diz-se supervisionado quando a rede neural recebe um conjunto de entradas (padrões) e suas correspondentes saídas, sendo que para cada entrada a rede produz uma saída, que deverá ser comparada com a saída desejada. Essa diferença é o erro, que deverá ser minimizado, ou seja, ocorrem ajustes nos pesos sinápticos até que esse erro seja mínimo.

Não-supervisionado é quando temos apenas o conjunto de entrada para classificar, não existe uma saída desejada. Já o híbrido, é o caso onde há uma “mistura” dos dois casos citados anteriormente: uma camada pode trabalhar com um tipo de aprendizado enquanto outra camada trabalha com o outro tipo.

#### 3.3.1.2 Modelos de RNAs

São muitos os modelos de RNAs, dentre os quais pode-se citar: o Perceptron, Redes Lineares e Redes de Múltiplas Camadas, que são as mais comuns, e outras como redes de base radial. Segundo TATAI (2000), ainda tem-se as redes recorrentes, redes de Hopfield, redes de Kohonen e redes construtivas.

Os modelos a serem utilizados neste trabalho serão as redes de múltiplas camadas e, por este motivo, a mesma será mais enfatizada.

##### a) Perceptron

Em 1957, Rosenblatt propôs uma rede neural simples, o Perceptron, e um algoritmo de aprendizagem para essa rede. A rede do tipo Perceptron foi a primeira rede neural que obteve sucesso prático. Segundo STEINER (2008), depois de um

número finito de tentativas, o Perceptron tem determinada habilidade para aprender a reconhecer os padrões de um conjunto. Na Figura 3.10 tem-se uma representação do Perceptron.

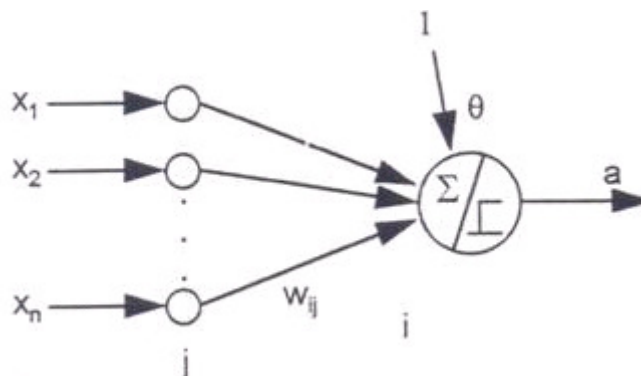


FIGURA 3.10 – PERCEPTRON. FONTE: STEINER (2008)

Nesta figura 3.10, tem-se que o Perceptron consiste de uma única camada de  $i$  neurônios conectados as  $n$  entradas através de um conjunto de pesos  $w(i, j)$ , onde  $i$  e  $j$ , indicam que  $w(i, j)$  é a "força" da conexão da  $j$ -ésima entrada ao  $i$ -ésimo neurônio.

Dados um vetor de entrada,  $\underline{x}$  e o erro do Perceptron,  $E$ , que é a diferença entre a resposta do neurônio,  $a$ , e da saída desejada ao padrão,  $t$  (ou  $d^p$ ), a regra de aprendizagem do Perceptron calcula as trocas nos pesos deste e o *bias*.

A saída desejada,  $t$ , deve conter os valores "0" e "1". A cada ajuste de pesos e *bias*, o Perceptron terá uma melhor chance de obter as saídas corretas,  $a = t$  (ou  $a = d^p$ ), dados os vetores de entrada  $\underline{x}$ .

Porém, como o Perceptron possui uma única camada de neurônios e a sua função de ativação é a função passo, seus cálculos são limitados, e só pode classificar problemas que são linearmente separáveis (TONSIG 2000). Assim, em 1969, Minsky e Papert constituíram um livro no qual mostravam as limitações teóricas do Perceptron, o que desencadeou em desinteresse por esse assunto (TODESCO, 1995).

Somente em 1982, com o trabalho de Hopfield e com o trabalho de Rumelhart, em 1986, os interesses pelas redes neurais foram realçados novamente. Esses trabalhos propuseram um novo algoritmo para aprendizagem, chamado de

retropropagação do erro (*back-propagation*), capaz de realizar tarefas bem mais complexas do que o Perceptron (OLIVEIRA 2002).

## b) Redes Lineares

As Redes Lineares, por sua vez, diferem do Perceptron por possuírem mais de uma camada de neurônios e, também, pela função de transferência que é linear, o que permite que as saídas tomem qualquer valor no intervalo (0, 1). Essas redes utilizam a regra de aprendizagem de Widrow-Hoff (Regra do Mínimos Quadrados) para ajustar os pesos e *bias* de acordo com a magnitude dos erros.

Segundo MACHADO *et al* (2005), essa regra de aprendizagem também é conhecida como “Regra Delta” – LMS (*Least Mean Square* ou minimização do erro médio quadrático) e foi introduzida em 1960-1962 para redes do tipo “feedforward perceptron”. Ela ajusta os pesos das conexões entre os neurônios da rede de acordo com o erro, ou seja, tem como objetivo encontrar um conjunto de pesos e polarizações que minimizem a função erro apresentada pela equação (3.8).

$$E = \frac{1}{2} \sum_{p=1}^R \sum_{i=1}^S (y_{p,i} - y'_{p,i})^2 \quad (3.8)$$

onde  $R$  é o número de padrões ou vetores de entrada;  $S$  é o número de neurônios de saída – dimensão do vetor de saída;  $y_{p,i}$  é a saída desejada no  $i$ -ésimo neurônio, quando o  $p$ -ésimo padrão é apresentado e  $y'_{p,i}$  é a saída obtida pela rede no  $i$ -ésimo neurônio, quando o  $p$ -ésimo padrão é apresentado.

A alteração dos pesos  $W_{i,j}$  da regra “Widrow-Hoff” é calculada da seguinte maneira através da equação (3.9).

$$\Delta W_{i,j} = -\eta \partial E / \partial W_{i,j} \quad (3.9)$$

onde  $\eta$  é a taxa de aprendizado e  $\partial E / \partial W_{i,j}$  é a derivada parcial do erro em relação ao peso da respectiva conexão – gradiente. A principal restrição na minimização do erro no sentido do gradiente decrescente é que a função de transferência do neurônio tem que ser monotônica e diferenciável em qualquer ponto.

### c) Redes de Múltiplas Camadas

A Rede de Múltiplas Camadas é um sistema artificial composto de células elementares (figura 3.11) – neurônios – organizadas em camadas sucessivas que são conectadas entre si, semelhante a um neurônio.

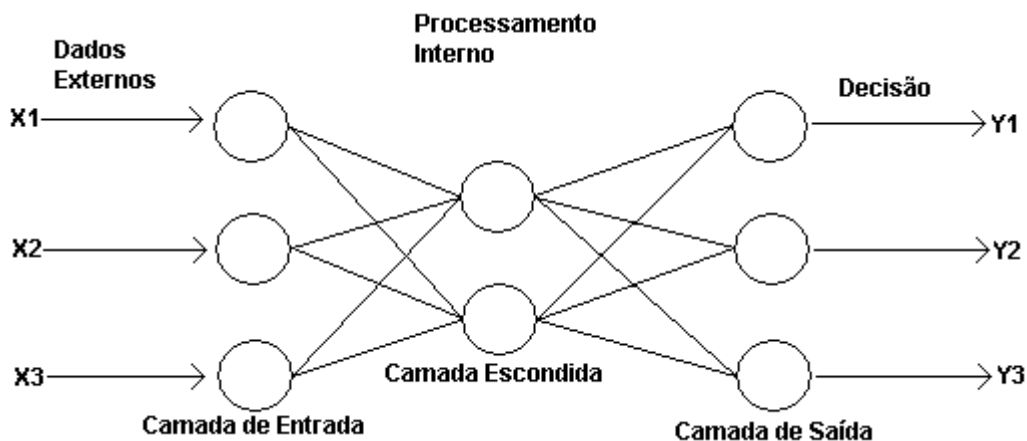


FIGURA 3.11 – REDE DE MÚLTIPLAS CAMADAS. FONTE: SAITO (2004)

Segundo TATIBANA (2009), o neurônio humano é constituído pelo corpo celular - que contém o núcleo do neurônio - os dendritos e os axônios. O núcleo do neurônio é responsável pelas transformações bioquímicas necessárias à síntese das enzimas e das outras moléculas que garantem a vida do neurônio; os dendritos são os receptores do neurônio, responsáveis por captar sinais que lhes são destinados e, finalmente, os axônios são uma fibra nervosa que serve de meio de transporte para os sinais emitidos pelo neurônio, conforme ilustrado na figura 3.12.

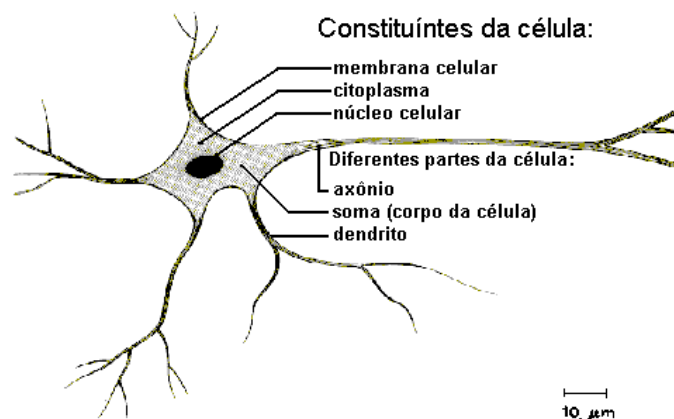


FIGURA 3.12 - MODELO DE NEURÔNIO. FONTE: TATIBANA (2009)

Assim, uma rede neural é um processador paralelo distribuído de forma massiva, com capacidade de armazenar conhecimento por experiência adquirida e torná-lo disponível para uso. Suas características se assemelham ao funcionamento do cérebro em dois aspectos: o conhecimento é adquirido pela rede através de um processo de aprendizagem e a força das conexões entre os neurônios ou pesos sinápticos é utilizada para armazenar o conhecimento (MARTINS, 2002).

As Redes de Múltiplas Camadas foram criadas generalizando a regra de aprendizagem de Windrow-Hoff para redes de múltiplas camadas e funções de transferência diferenciáveis não-lineares. Elas podem ser treinadas com o algoritmo de treinamento *back-propagation* e os vetores de entrada e saída são usados para treinar a rede até que ela possa aproximar uma função que classifique os vetores de entrada de maneira apropriada.

Essas redes frequentemente usam função de transferência sigmoideal, que gera saídas no intervalo (0, 1) para entradas variando no intervalo  $(-\infty, +\infty)$  e, em geral, possuem três camadas: a camada de entrada, onde os padrões são apresentados à rede; a camada intermediária ou oculta, onde é realizado o processamento, e a camada de saída, que apresenta os valores de saídas da rede. Segundo o Teorema de Kolmogorov, uma RN com apenas uma camada oculta pode calcular uma função arbitrária qualquer a partir dos dados fornecidos.

Também, segundo RANGEL (2010), quanto menos complexas essas redes forem, maior pode ser sua sincronia, ou seja, “um menor número de iterações pode melhorar a sincronia e o funcionamento da rede como um todo”.

Cada uma destas camadas é totalmente conectada à camada seguinte e à anterior. O sinal que chega à camada de entrada se propaga, camada à camada, até a saída. Os valores de saída para cada um dos padrões são comparados com os valores desejados para os padrões e o erro  $E$  é calculado. Com base no valor do erro, os pesos da rede são ajustados.

Dessa forma, os parâmetros que caracterizam uma rede multicamada são os seguintes: número de camadas, número de neurônios por camada, escolha dos valores iniciais para as conexões (pesos), tipo de funções de ativação dos neurônios (sigmoideal) e valores iniciais aos parâmetros  $\alpha$  (taxa de momento) e  $\gamma$  (taxa de

aprendizagem), descritos mais adiante, assim como a forma de ajuste destes dois parâmetros.

A seguir é feita a descrição dos passos do algoritmo *back-propagation*, dividido em duas fases: a propagação *forward* e *backward*, conforme a RN da figura 3.13.

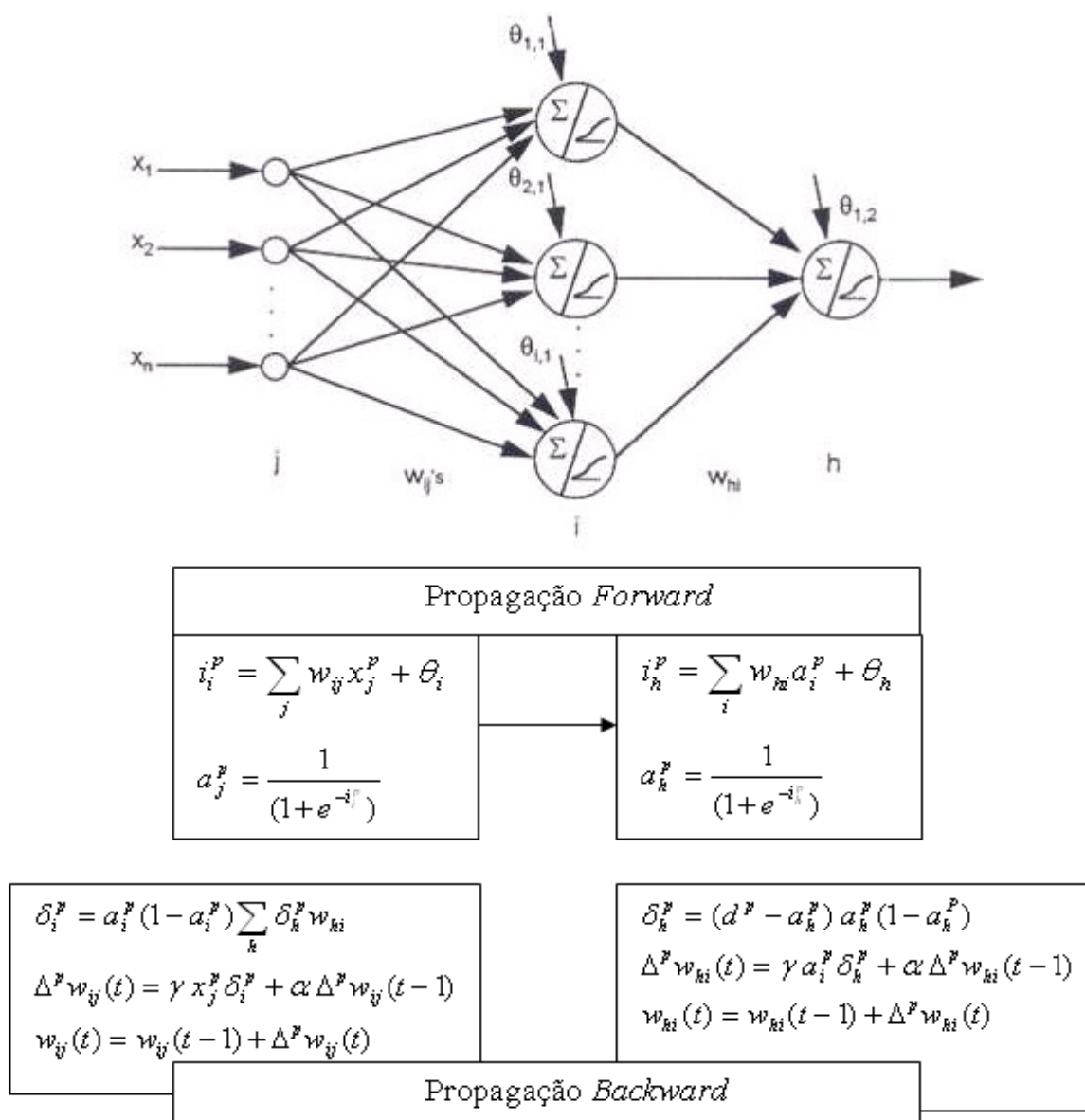


FIGURA 3.13 - REDE DE MÚLTIPLAS CAMADAS. FONTE: STEINER (2008)

O algoritmo *back-propagation* inicia com a propagação *forward*, onde todos os padrões (do treinamento), descritos por suas coordenadas  $x_j^p$ , são apresentados à rede. Cada unidade  $i$ , da camada escondida, recebe os sinais de entrada e os agrupa conforme a expressão (3.10).

$$i_i^p = \sum_j w_{ij} x_j^p + \theta_i \quad (3.10)$$

onde  $i_i^p$  é a entrada da unidade  $i$  para o padrão  $p$ ;  $w_{ij}$  é a conexão peso entre as unidades  $i$  e  $j$ ;  $x_j^p$  são as entradas do padrão  $p$  e  $\theta_i$  são as *bias* das unidades  $i$ .

Com base no resultado encontrado em (3.10), é gerado um sinal de saída,  $a_i^p$ , para o padrão  $p$ , utilizando a função sigmoïdal, conforme a função (3.11).

$$a_i^p = \frac{1}{(1 + e^{-i_i^p})} \quad (3.11)$$

Os sinais de saída são enviados para a única unidade da camada  $h$ , que os agrupa em  $i_h^p$ , como mostra a expressão 3.12.

$$i_h^p = \sum_i w_{hi} a_i^p + \theta_h, \quad h = 1, 2, 4 \text{ (para o caso aqui estudado)} \quad (3.12)$$

gerando a saída conforme (3.13).

$$a_h^p = \frac{1}{(1 + e^{-i_h^p})} \quad (3.13)$$

Essa saída obtida é comparada com o valor da saída desejada, calculando-se o erro quadrático  $E$ , onde  $d^p$  é a saída desejada para o padrão  $p$ , e  $a_h^p$ , a saída obtida, conforme (3.14).

$$E = \sum_{p=1}^{m+k} \frac{(d^p - a_h^p)^2}{2} \quad (3.14)$$

Segundo Steiner (2008), é o vetor de pesos  $W$  que constitui o que a rede neural “sabe”; o objetivo é, ajustando estes pesos, minimizar o erro  $E$ , de tal modo que todos os vetores de entrada sejam corretamente mapeados em suas correspondentes saídas.

Terminada essa fase (propagação *forward*), inicia-se a propagação *backward*, que executa o gradiente descendente do erro  $E$  em relação a  $W$ , fazendo o ajuste dos pesos.

Nessa fase, começa-se calculando delta, conforme expressão (3.15). Depois os pesos entre a camada de saída e a camada oculta são atualizados, conforme expressão (3.16) e (3.17), onde  $t$  = número da iteração ou do padrão,  $\gamma$  é a taxa de aprendizagem,  $0 < \gamma < 1$  e  $\alpha$  é a taxa momento.

$$\delta_h^p = (d^p - a_h^p) a_h^p (1 - a_h^p) \quad (3.15)$$

$$\Delta^p w_{hi}(t) = \gamma a_i^p \delta_h^p + \alpha \Delta^p w_{hi}(t-1) \quad (3.16)$$

$$w_{hi}(t) = w_{hi}(t-1) + \Delta^p w_{hi}(t) \quad (3.17)$$

O próximo passo, então, é atualizar os pesos entre a camada oculta e a camada de entrada, conforme expressão (3.18):

$$\begin{aligned} \delta_i^p &= a_i^p (1 - a_i^p) \sum_h \delta_h^p w_{hi} \\ \Delta^p w_{ij}(t) &= \gamma x_j^p \delta_i^p + \alpha \Delta^p w_{ij}(t-1) \\ w_{ij}(t) &= w_{ij}(t-1) + \Delta^p w_{ij}(t) \end{aligned} \quad (3.18)$$

Depois de todos os padrões terem passado pela rede, tem-se definida uma iteração. Em geral, são necessárias milhares de iterações para se obter a convergência da rede.

Com referência a rapidez na convergência, alguns autores recomendam que a taxa de aprendizagem da RN,  $\gamma$ , seja alta no início do treinamento, cerca de 0,8, declinando gradativamente a medida que ele evolui, e que o termo momento  $\alpha$  seja considerado, proporcionando assim, rapidez na convergência do treinamento, estabilidade e resistência ao aparecimento de mínimos locais (STEINER 2008).

Também que, se houver um coeficiente de aprendizado inicial muito alto; dados com magnitude incompatível com a função de ativação dos neurônios ou uma rede mal dimensionada pode não haver convergência já na fase de treinamento.

#### d) Rede Neural de Base Radial (RBF)

Segundo BARRETO (2008), as redes de funções de Base Radial (RBFs) são redes neurais multicamadas, ou seja, com neurônios ocultos não-lineares. Dessa forma, elas podem resolver problemas não-linearmente separáveis.



Segundo TODESCO (1995), existem muitas diferenças entre o perceptron multi-camadas e as redes de função de base radial. Por exemplo: aquele pode ter mais de uma camada intermediária, enquanto as RBF tem uma única; a RBF utiliza a distância Euclidiana entre o vetor de entrada e o centro de cada unidade na camada intermediária como função de ativação, já o perceptron multi-camadas faz o cálculo utilizando o produto interno do vetor de entrada e o vetor peso sináptico daquela unidade.

Segundo HAYKIN (2001), a forma mais básica de construção de uma RBF envolve três camadas, com papéis completamente diferentes. A camada de entrada constitui-se de unidades sensoriais, ou nós de fonte, que conectam a rede ao seu ambiente; a camada oculta (única) aplica uma transformação não-linear do espaço de entrada para o oculto, que na maioria das aplicações, é de alta dimensionalidade; a camada de saída, que é linear, fornece a resposta da rede ao sinal de ativação aplicado à camada de entrada.

Como se pode observar no parágrafo anterior, há uma transformação não-linear seguida de uma linear, que pode ser esclarecida com o *Teorema de Cover* sobre a separabilidade de padrões. Segundo HAYKIN (2001), esse teorema pode ser formulado da seguinte forma: um problema de classificação de padrões disposto não linearmente em um espaço de alta dimensão tem maior probabilidade de ser linearmente separável do que em um espaço de baixa dimensionalidade, isso justifica o fato do espaço oculto ter alta dimensão.

Os neurônios de saída são sempre lineares e os neurônios da camada intermediária têm apenas funções de base radial como função de ativação, ao invés de funções sigmoidais ou outras. Essa função de ativação calcula a distância euclidiana entre o vetor de entrada e o centro daquela unidade. Segundo HAYKIN (2001) as características lineares da camada de saída da RBF indicam que ela está mais proximamente relacionada com o perceptron de Rosenblatt do que com o perceptron de múltiplas camadas.

A estrutura da rede é apresentada na figura 3.14, para o caso de uma única saída.

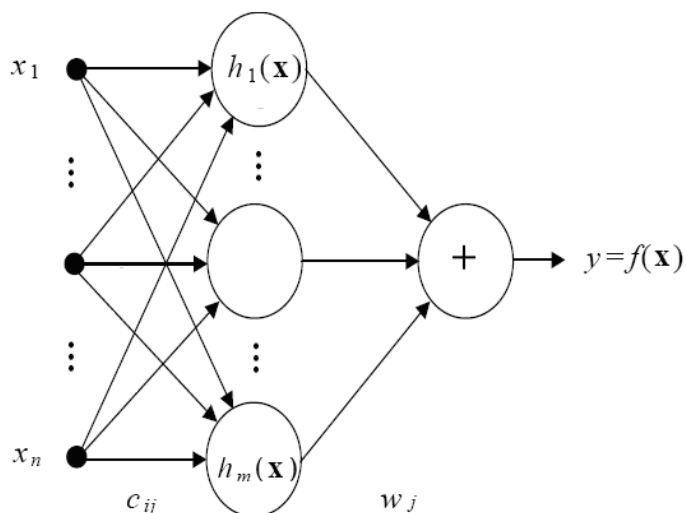


FIGURA 3.14 – REDE NEURAL DE BASE RADIAL. FONTE: ZUBEN e ATTUX (2009)

onde  $i$  vai de 1 até  $n$  e  $j$  vai de 1 até  $m$ .

#### e) Redes Recorrentes

São redes que possuem realimentação, ou seja, a saída da rede é utilizada como entrada para ela mesma. Essas redes modelam sistemas dinâmicos não-lineares e são adequadas para tarefas como mapeamento de dados de entrada-saída, memória associativa, previsão de séries temporais, entre outros.

Segundo HAYKIN (2001), a rede recorrente não sofre nem de escassez (*underfitting*) nem de excesso de ajuste (*overfitting*), ela incorpora um perceptron de múltiplas camadas estático ou partes dele e explora a capacidade de mapeamento não-linear do perceptron de múltiplas camadas. Haykin também confirma que há dois modos de treinar uma rede recorrente, treinamento por iterações e treinamento contínuo.

No treinamento por iterações a rede inicia a execução de algum estado inicial em uma dada iteração, até alcançar um novo estado, onde o treinamento é parado e a rede é reinicializada em um estado inicial para a próxima iteração. Esse estado inicial não precisa ser igual ao anterior, além de ser importante que este seja diferente do estado alcançado ao final da iteração anterior.

O treinamento contínuo, porém, nunca pára, assim a rede aprende enquanto realiza o processamento de sinal. Esse método é adequado a situações em que não

haja estado de reinício e/ou em que seja requerida aprendizagem em tempo de execução.

#### f) Redes de Hopfield

Foi desenvolvida em 1982 por J.J.Hopfield, e é utilizada, de uma forma geral, para reconhecimento de imagens.

Segundo TATAI (2000), é um caso característico de uma rede recorrente totalmente conectada, normalmente utilizada para o armazenamento de padrões e a recuperação dos mesmos a partir de dados ruidosos ou incompletos. O algoritmo de aprendizado utilizado é a regra de Hebb generalizada que define os pesos da rede para armazenamento do padrão dado. A recuperação dos padrões é dada através do ajuste sucessivo das saídas dos neurônios da rede até ser atingida a estabilidade.

A capacidade de armazenamento desta rede é limitada tanto pelo número de neurônios da rede quanto pelas instabilidades decorrentes da não-linearidade da rede.

#### g) Redes de Kohonen

Desenvolvida em 1987 pelo Prof. Teuvo Kohonen, é utilizada para classificação de padrões, otimização de problemas e simulações (OLIVEIRA, 2002).

A rede de Kohonen é constituída por neurônios dispostos em vizinhanças geralmente uni-dimensionais ou bi-dimensionais, capazes de se auto-organizar, através do algoritmo de aprendizado, de forma a detectar padrões recorrentes. Essa rede realiza um mapeamento entre um espaço  $n$ -dimensional (a dimensão do problema original) ao espaço de vizinhança da rede (uni ou bi-dimensional), com perda de informação.

Simplificando, para cada padrão fornecido o algoritmo de aprendizagem determina o neurônio mais próximo ao mesmo, aproximando este neurônio e seus vizinhos ao padrão. Após certo número de iterações, os neurônios tendem a agrupar-se em torno dos padrões fornecidos. Assim, a partir das relações de vizinhança entre os neurônios, podem ser determinados os padrões (*clusters*) recorrentes (TATAI 2000).

## h) Redes construtivas

Segundo TATAI (2000), as redes construtivas “não constituem propriamente uma arquitetura *per se*”, ou seja, geralmente são redes em camadas diferenciadas. Porém, elas se diferenciam das redes em camadas por que suas funções de ativação e topologia são determinadas através de um processo de aprendizagem, de forma a se adaptar ao problema e representá-lo melhor, geralmente resultando em economia de neurônios e recursos computacionais.

### 3.3.2 Função Discriminante Linear de Fisher

Segundo JOHNSON e WICHERN (1998), foi Ronald A. Fisher que introduziu a terminologia "discriminar" e "classificar" no primeiro tratamento moderno dos problemas de separação de conjuntos na década de 30.

Assim, o objetivo básico do método de Fisher é separar populações, além de também poder ser usado com o propósito de classificar. Esse método pode ser usado tanto para duas populações quanto para diversas populações.

#### 3.3.2.1 Método de Fisher para duas populações

Segundo LANDIM (2008), dadas duas populações  $\pi_1$  e  $\pi_2$  de observações multivariadas  $\underline{X}$ 's, de dimensão  $n$ , a idéia de Fisher foi transformar estas observações multivariadas em observações univariadas  $Y$ 's tal que os  $Y$ 's nas populações  $\pi_1$  e  $\pi_2$  sejam separados tanto quanto possível. Pelo fato da combinação linear ser de fácil obtenção matematicamente, Fisher propôs o uso desta nas  $n$  variáveis aleatórias componentes de  $\underline{X}$  para obtenção dos  $Y$ 's.

O método funciona da seguinte forma: seja  $\mu_{1Y}$  a média dos  $Y$ 's obtida dos  $\underline{X}$ 's pertencentes à população  $\pi_1$  e  $\mu_{2Y}$  a média dos  $Y$ 's obtida dos  $\underline{X}$ 's pertencentes à população  $\pi_2$ , seleciona-se a combinação linear que maximiza a distância quadrática entre  $\mu_{1Y}$  e  $\mu_{2Y}$  relativamente à variabilidade dos  $Y$ 's, ou seja, tendo:

$$\underline{\mu}_1 = E(\underline{X}/\pi_1): \text{ valor esperado de uma observação multivariada de } \pi_1$$

$$\underline{\mu}_2 = E(\underline{X}/\pi_2): \text{ valor esperado de uma observação multivariada de } \pi_2$$

e supondo a matriz de covariância dada em (3.19) (a mesma para ambas as populações):

$$\underline{\Sigma} = E[(\underline{X} - \underline{\mu}_i)(\underline{X} - \underline{\mu}_i)'] , \quad i = 1, 2 \quad (3.19)$$

E considerando a combinação linear (3.20):

$$Y = \underline{C}' \underline{X} \quad (3.20)$$

têm-se as médias dadas em (3.21) e (3.22):

$$\mu_{1Y} = E(Y/\pi_1) = E(\underline{C}' \underline{X}/\pi_1) = \underline{C}' E(\underline{X}/\pi_1) = \underline{C}' \underline{\mu}_1, \quad (3.21)$$

$$\mu_{2Y} = E(Y/\pi_2) = E(\underline{C}' \underline{X}/\pi_2) = \underline{C}' E(\underline{X}/\pi_2) = \underline{C}' \underline{\mu}_2 \quad (3.22)$$

e sua variância conforme (3.23) (a mesma para as duas populações)

$$\sigma_Y^2 = V(Y) = V(\underline{C}' \underline{X}) = \underline{C}' V(\underline{X}) \underline{C} = \underline{C}' \underline{\Sigma} \underline{C} \quad (3.23)$$

Segundo Fisher, a melhor combinação linear vem da razão entre o “quadrado da distância entre as médias” e a “variância de Y”, ou seja, conforme (3.24):

$$\frac{(\mu_{1Y} - \mu_{2Y})^2}{\sigma_Y^2} = \frac{(\underline{C}' \underline{\mu}_1 - \underline{C}' \underline{\mu}_2)^2}{\underline{C}' \underline{\Sigma} \underline{C}} = \frac{(\underline{C}' \underline{\delta})^2}{\underline{C}' \underline{\Sigma} \underline{C}} \quad (3.24)$$

onde:  $\underline{\delta} = \underline{\mu}_1 - \underline{\mu}_2$  e  $V(\underline{X}) = \underline{\Sigma}$ .

Segundo JOHNSON e WICHERN (1998), seja  $\underline{\delta} = \underline{\mu}_1 - \underline{\mu}_2$  e  $Y = \underline{C}' \underline{X}$ , então a razão  $\frac{(\underline{C}' \underline{\delta})^2}{\underline{C}' \underline{\Sigma} \underline{C}}$  é maximizada por  $\underline{C} = K \underline{\Sigma}^{-1} \underline{\delta} = K \underline{\Sigma}^{-1} (\underline{\mu}_1 - \underline{\mu}_2)$  para qualquer  $K \neq 0$ . Escolhendo-se  $K=1$  tem-se  $\underline{C} = \underline{\Sigma}^{-1} (\underline{\mu}_1 - \underline{\mu}_2)$  e então Y é dado por (3.25):

$$Y = \underline{C}' \underline{X} = (\underline{\mu}_1 - \underline{\mu}_2)' \underline{\Sigma}^{-1} \underline{X} \quad (3.25)$$

que é denominada Função Discriminante Linear de Fisher (FDLF).

Assim, definindo:

$$Y_0 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} \underline{X}_0 \quad (3.26)$$

como o valor da função para uma nova observação  $\underline{X}_0$  e considerando  $m$  o ponto médio entre as duas populações univariadas, tem-se conforme (3.27):

$$\begin{aligned} m &= \frac{1}{2}(\mu_{1Y} + \mu_{2Y}) = \frac{1}{2}(\underline{C}' \underline{\mu}_1 + \underline{C}' \underline{\mu}_2) = \frac{1}{2} \underline{C}' (\underline{\mu}_1 + \underline{\mu}_2) \\ &= \frac{1}{2} [(\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} (\underline{\mu}_1 + \underline{\mu}_2)] \end{aligned} \quad (3.27)$$

Subtraindo  $m$  dos valores esperados de  $Y_0$  condicionados a cada uma das duas populações obtém-se (3.28):

$$E(Y_0/\pi_1) - m \geq 0 \text{ e } E(Y_0/\pi_2) - m < 0 \quad (3.28)$$

ou seja, se  $\underline{X}_0 \in \pi_1$ , é esperado que  $Y_0 \geq m$  e se  $\underline{X}_0 \in \pi_2$ , é esperado que  $Y_0 < m$ . Em outras palavras, se  $Y_0$  provém da população  $\pi_1$ , espera-se que ele seja superior ao ponto médio  $m$ ; caso contrário, espera-se que ele seja inferior ao ponto médio  $m$  (JOHNSON e WICHERN, 1998).

Desse modo, a Regra de Classificação é: alocar  $\underline{X}_0$  em  $\pi_1$  se  $Y_0 \geq m$  ou alocar  $\underline{X}_0$  em  $\pi_2$  se  $Y_0 < m$ .

Contudo, como  $\pi_1$ ,  $\pi_2$  e  $\Sigma$  geralmente são desconhecidos, é necessário estimá-los a partir das observações amostrais.

Suponha então que se tenham amostras de tamanho  $n_1$  e  $n_2$  de  $\pi_1$  e  $\pi_2$ , respectivamente, e que  $\underline{X}_1$  e  $\underline{X}_2$  sejam as matrizes de dados, onde

$$\begin{aligned} \bar{\underline{X}}_1 &= \frac{1}{n_1} \sum_{i=1}^{n_1} X_{i1} \quad \text{estima} \quad \underline{\mu}_1, & \bar{\underline{X}}_2 &= \frac{1}{n_2} \sum_{i=1}^{n_2} X_{i2} \quad \text{estima} \quad \underline{\mu}_2 \quad \text{e} \\ S_p &= \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \quad \text{estima} \quad \Sigma, \text{ com:} \end{aligned}$$

$$S_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (\underline{X}_{i1} - \bar{\underline{X}}_1)(\underline{X}_{i1} - \bar{\underline{X}}_1)' \quad (3.29)$$

$$S_2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (\underline{X}_{i2} - \bar{\underline{X}}_2)(\underline{X}_{i2} - \bar{\underline{X}}_2)' \quad (3.30)$$

A FDL de Fisher estimada será dada por:

$$\hat{Y} = \hat{\underline{C}}' \underline{X} = (\underline{X}_1 - \underline{X}_2)' S_p^{-1} \underline{X} \quad (3.31)$$

sendo o ponto médio  $m$  estimado por:

$$\hat{m} = \frac{1}{2} (\underline{X}_1 - \underline{X}_2)' S_p^{-1} (\underline{X}_1 + \underline{X}_2) = \frac{\bar{Y}_1 + \bar{Y}_2}{2} \quad (3.32)$$

Assim, a regra de alocação baseada na F.D.L. de Fisher estimada fica:

$$\text{- Se } \hat{Y}_0 - \hat{m} \geq 0 \text{ ou } \hat{Y}_0 \geq \hat{m} \text{ alocar } \underline{X}_0 \text{ em } \pi_1 \quad (3.32.a)$$

$$\text{- Se } \hat{Y}_0 - \hat{m} < 0 \text{ ou } \hat{Y}_0 < \hat{m} \text{ alocar } \underline{X}_0 \text{ em } \pi_2. \quad (3.32.b)$$

### 3.3.2.2 Método de Fisher para discriminar entre várias populações

Segundo JOHNSON e WICHERN (1998), este método não necessita da suposição de que as diversas populações sejam normais, porém assume-se que as matrizes de covariâncias populacionais  $\Sigma$ 's são iguais, isto é,  $\Sigma_1 = \Sigma_2 = \dots \Sigma_g = \Sigma$ .

Admitindo  $\bar{\underline{\mu}}$  como o vetor médio das  $g$  populações e  $B_0$  a matriz "de variação entre grupos populacionais" tal que

$$\bar{\underline{\mu}} = \frac{1}{g} \sum_{i=1}^g \underline{\mu}_i \quad (3.33)$$

$$B_0 = \sum_{i=1}^g (\underline{\mu}_i - \bar{\underline{\mu}})(\underline{\mu}_i - \bar{\underline{\mu}})' \quad (3.34)$$

Considerando a combinação linear  $Y = \underline{C}' \underline{X}$  com esperança dada por

$$E(Y) = \underline{\mathbf{C}}' E(\underline{\mathbf{X}} / \pi_i) = \underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_i \quad (3.35)$$

para  $i = 1, 2, \dots, g$ , e variância dada por

$$\sigma_Y^2 = V(Y) = V(\underline{\mathbf{C}}' \underline{\mathbf{X}}) = \underline{\mathbf{C}}' V(\underline{\mathbf{X}}) \underline{\mathbf{C}} = \underline{\mathbf{C}}' \underline{\boldsymbol{\Sigma}} \underline{\mathbf{C}} \quad (3.36)$$

para todas as populações. Por conseguinte o valor esperado univariado

$$E(Y) = \mu_{iY} = \underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_i \quad (3.37)$$

se altera conforme muda a população da qual  $\underline{\mathbf{X}}$  provém. Temos então uma média global dada por

$$\bar{\mu}_Y = \frac{1}{g} \sum_{i=1}^g \mu_{iY} = \frac{1}{g} \sum_{i=1}^g \underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_i = \underline{\mathbf{C}}' \bar{\boldsymbol{\mu}} \quad (3.38)$$

e portanto a razão entre a “soma dos quadrados das distâncias das populações para a média global de Y” e a “variância de Y” é

$$\begin{aligned} \frac{\sum_{i=1}^g (\mu_{iY} - \bar{\mu}_Y)^2}{\sigma_Y^2} &= \frac{\sum_{i=1}^g (\underline{\mathbf{C}}' \underline{\boldsymbol{\mu}}_i - \underline{\mathbf{C}}' \bar{\boldsymbol{\mu}})^2}{\underline{\mathbf{C}}' \underline{\boldsymbol{\Sigma}} \underline{\mathbf{C}}} = \frac{\sum_{i=1}^g [\underline{\mathbf{C}}' (\underline{\boldsymbol{\mu}}_i - \bar{\boldsymbol{\mu}}) (\underline{\boldsymbol{\mu}}_i - \bar{\boldsymbol{\mu}})' \underline{\mathbf{C}}]}{\underline{\mathbf{C}}' \underline{\boldsymbol{\Sigma}} \underline{\mathbf{C}}} \\ &= \frac{\underline{\mathbf{C}}' [\sum_{i=1}^g (\underline{\boldsymbol{\mu}}_i - \bar{\boldsymbol{\mu}}) (\underline{\boldsymbol{\mu}}_i - \bar{\boldsymbol{\mu}})' ] \underline{\mathbf{C}}}{\underline{\mathbf{C}}' \underline{\boldsymbol{\Sigma}} \underline{\mathbf{C}}} = \frac{\underline{\mathbf{C}}' \underline{\mathbf{B}}^0 \underline{\mathbf{C}}}{\underline{\mathbf{C}}' \underline{\boldsymbol{\Sigma}} \underline{\mathbf{C}}} \end{aligned} \quad (3.39)$$

que é uma generalização multi-grupal do caso de duas populações.

Segundo JOHNSON e WICHERN (1998), esta razão mede a variabilidade entre os grupos de valores Y relativamente à variabilidade comum entre os grupos. Equivalentemente ao problema de duas populações podemos selecionar  $\underline{\mathbf{C}}$  de modo a maximizar esta razão. É conveniente escalonar  $\underline{\mathbf{C}}$  tal que  $\underline{\mathbf{C}}' \underline{\boldsymbol{\Sigma}} \underline{\mathbf{C}} = 1$ .

Assim temos o seguinte resultado (JOHNSON e WICHERN 1998): sejam  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_s > 0$  os  $s \leq \text{mín}(g - 1, p)$  autovalores não-nulos ( $g = n^\circ$  de populações,  $p = n^\circ$  de variáveis) de  $\underline{\boldsymbol{\Sigma}}^{-1} \underline{\mathbf{B}}_0$  e  $\underline{\mathbf{e}}_1, \underline{\mathbf{e}}_2, \dots, \underline{\mathbf{e}}_s$  os correspondentes autovetores



escalonados tal que  $\underline{e}' \sum \underline{e} = 1$ . Então o vetor de coeficiente  $\underline{C}$  que maximiza a razão  $\frac{\underline{C}' B_0 \underline{C}}{\underline{C}' \sum \underline{C}}$  é dado por  $\underline{C}_1 = \underline{e}_1$ .

A combinação linear  $\underline{C}'_1 \underline{X} = \underline{e}'_1 \underline{X}$  é chamada primeiro discriminante. O vetor  $\underline{C}_2 = \underline{e}_2$  maximiza aquela razão sujeito à condição  $Cov(\underline{C}'_1 \underline{X}, \underline{C}'_2 \underline{X}) = 0$ .

A combinação linear  $\underline{C}'_2 \underline{X} = \underline{e}'_2 \underline{X}$  é chamada segundo discriminante, e assim sucessivamente, obtemos os vários discriminantes. Ainda, temos que  $V(\underline{C}'_i \underline{X}) = 1$ .

Ordinariamente, a matriz de covariância  $\Sigma$  e os vetores de média da população  $\underline{\mu}_i$  são desconhecidos, entretanto podemos estimar esses parâmetros desconhecidos a partir de amostras aleatórias de tamanhos  $n_i$  das populações  $\pi_i$  ( $i = 1, 2, \dots, g$ ), assim obtemos os discriminantes lineares amostrais de Fisher:

$$\bar{\underline{X}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \underline{X}_{ij}, \text{ estima } \underline{\mu}_i \quad (3.40)$$

$$\bar{\underline{X}} = \frac{\sum_{i=1}^g n_i \bar{\underline{X}}_i}{\sum_{i=1}^g n_i} = \frac{\sum_{i=1}^g \sum_{j=1}^{n_i} \underline{X}_{ij}}{\sum_{i=1}^g n_i}, \text{ estima } \bar{\underline{\mu}} \text{ (média global)} \quad (3.41)$$

e a matriz soma de produtos cruzados “variação entre grupos amostrais” é dada por

$$\hat{B}_0 = \sum_{i=1}^g (\bar{\underline{X}}_i - \bar{\underline{X}})(\bar{\underline{X}}_i - \bar{\underline{X}})' \quad (3.42)$$

Um estimador da matriz  $\Sigma$  pode ser obtido com base na matriz  $W$ , “variação dentro dos grupos amostrais”

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (\underline{X}_{ij} - \bar{\underline{X}}_i)(\underline{X}_{ij} - \bar{\underline{X}}_i)' \quad (3.43)$$

Conseqüentemente

$$\frac{W}{n_1 + n_2 + \dots + n_g - g} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2 + \dots + (n_g - 1)S_g}{n_1 + n_2 + \dots + n_g - g} = S_p \quad (3.44)$$

é um estimador para  $\Sigma$ .

Demonstra-se que para maximizar a razão  $\frac{\hat{\underline{C}}' \hat{\underline{B}}_0 \hat{\underline{C}}}{\hat{\underline{C}}' \underline{S}_p \hat{\underline{C}}}$  ou  $\frac{\hat{\underline{C}}' \hat{\underline{B}}_0 \hat{\underline{C}}}{\hat{\underline{C}}' \underline{W} \hat{\underline{C}}}$ , o vetor  $\hat{\underline{C}}$

que a maximiza é dado por  $\hat{\underline{C}}_1 = \hat{\underline{e}}_1$ , e a combinação linear  $\hat{\underline{C}}_1' \underline{X}$  é chamada 1º discriminante amostral. Continuando, da mesma forma que, no caso populacional, temos no passo  $K$ ,  $\hat{\underline{C}}_K = \hat{\underline{e}}_K$  ( $K \leq s$ ) e  $\hat{\underline{C}}_K' \underline{X}$  o  $K$ -ésimo discriminante amostral, onde:  $\hat{\underline{C}}_1 = \hat{\underline{e}}_1$ ,  $\hat{\underline{C}}_2 = \hat{\underline{e}}_2$ , ... ,  $\hat{\underline{C}}_K = \hat{\underline{e}}_K$  são os autovetores correspondentes dos autovalores obtidos da matriz  $\underline{W}^{-1} \hat{\underline{B}}_0$ .

### 3.4 REVISÃO DA LITERATURA – TRABALHOS CORRELATOS

Nesta seção encontram-se resumos de alguns trabalhos que também utilizaram a ideia do processo *KDD* e *Data Mining* para as mais diversas aplicações, além de trabalhos relacionados aos problemas vocais dos docentes.

#### 3.4.1 Trabalhos que utilizaram técnicas semelhantes

BALDIN (2002) utiliza as técnicas de Redes Neurais como ferramenta auxiliar na Orientação Profissional dos jovens. Foram testadas três topologias de RNAs: com um único neurônio dicotômico na camada de saída; com um neurônio tricotômico (um único neurônio na camada de saída dividido em três intervalos) e, finalmente, com três neurônios na camada de saída; as três redes foram treinadas com o algoritmo do *back-propagation*. Além disso, foram utilizados cinco tipos de conjuntos de pesos iniciais: pesos iguais; soma de pesos = 1; pesos randômicos no intervalo [0, 1]; pesos randômicos no intervalo [-1, +1] e pesos randômicos no intervalo [-0,1; +0,1]. Em todas as topologias, com os diferentes conjuntos de pesos iniciais, o número de neurônios na camada intermediária variou de 1 até 10. Os resultados encontrados foram bastante satisfatórios, com exceção do tipo pesos iguais.

Algumas das conclusões obtidas deste trabalho foram: colocar um neurônio tricotômico na saída, com valores divididos em intervalos é equivalente a colocar um número de neurônios igual ao número de faixas, cada uma delas representada por um neurônio dicotômico. Assim, a segunda e a terceira topologias são equivalentes. Um neurônio tricotômico “consome” menos tempo de processamento para o treinamento do que a saída com três neurônios. Além disso, consome um menor número de iterações.

LEMOS (2003) utilizou as seguintes técnicas de *Data Mining*: RNAs e árvores de decisão, através do uso dos softwares *MATLAB Neural Networks Toolbox* e *WEKA (Waikato Environment for Knowledge Analysis)*, respectivamente, na análise de crédito bancário para auxiliar na tomada de decisão sobre conceder ou não crédito a um novo cliente. Os resultados foram bastante satisfatórios, mostrando que, para este problema específico, as Redes Neurais apresentaram um percentual menor de erros.

BRAULIO (2005) trabalhou com Técnicas Estatísticas Multivariadas a fim de construir um modelo estatístico que avaliasse o valor de imóveis urbanos (casas, apartamentos e terrenos) da cidade de Campo Mourão, PR, com a máxima precisão. Utilizou o software *Statgraphics Plus 5.0* para o processamento dos dados e o *Software Excel*®. Os resultados obtidos com os modelos de Regressão Linear Múltipla foram altamente precisos, assim, a metodologia aplicada mostrou-se viável e altamente apropriada para avaliação dos imóveis.

BAPTISTELLA (2005) fez uma comparação das técnicas de RNAs e de Regressão Linear Múltipla, na estimação dos valores venais de imóveis urbanos da cidade de Guarapuava, PR. As duas técnicas se mostraram eficientes, porém as RNAs apresentaram um desempenho superior aos resultados obtidos pela outra técnica. Foi utilizado o *Software Matlab*.

ALVES (2005) em seu trabalho com uma amostra de 119 imóveis, tem o objetivo principal de desenvolver um programa computacional em *MATLAB* que avalie o valor de imóveis urbanos utilizando técnicas estatísticas da Análise Multivariada, considerando os atributos de cada imóvel como um vetor de características. Seus resultados foram satisfatórios, tanto que o programa criado nesse trabalho pode ser utilizado por qualquer cidade, bastando alterar os dados.

PAVANELLI (2007) implementou uma RNA para fornecer uma previsão do tempo de duração dos processos trabalhistas. Foram treinadas várias RNAs, com as mais variadas topologias, utilizando a Análise de Componentes Principais (ACP) para refinar ainda mais os resultados obtidos. As ferramentas utilizadas mostraram-se de grande eficiência e os resultados mostraram que a ACP é uma ferramenta que melhora o desempenho da RNA.

MENDES (2008) aplicou as RNAs como ferramenta de reconhecimento de padrões nos diagnósticos médicos dos tipos mais freqüentes de Cefaléia, utilizando dados de 2.177 pacientes. Serviu-se do *software MATLAB 7.0* e do componente *Neural Network Toolbox* para a sua implementação e seu treinamento. Na primeira etapa foram consideradas uma única RNA com cinco saídas e, na segunda etapa, foram consideradas cinco RNAs, cada uma delas com uma saída, sendo cada uma das RNAs responsável em fazer o treinamento de cada um dos diagnósticos. Os resultados obtidos foram satisfatórios e, o modelo que apresentou o melhor desempenho foi o que utilizou 5 redes neurais com uma saída.

#### 3.4.2 Trabalhos relacionados a problemas de voz de docentes

JARDIM (2006), com o objetivo de investigar as relações entre qualidade de vida e voz, dos docentes, realizou uma revisão bibliográfica sistemática sobre disfonias em professores por meio de consulta a diversas bases de dados, considerando os artigos publicados a partir de 1990. Concluiu que há uma elevada prevalência de distúrbios de voz nessa classe de trabalhadores, especialmente pelas condições de trabalho muitas vezes precárias, e que o professor tende a sentir-se culpado quando está “sem voz” e se vê obrigado a se afastar da sala de aula.

AMORIM (2006) fez um estudo exploratório quantitativo e qualitativo sobre a ocorrência de problemas vocais e estresse em professores(as) do ensino fundamental de Goiânia, GO, estabelecendo uma relação entre a precariedade do trabalho dos docentes, os distúrbios vocais e estresse. Mostrou um maior impacto sobre as mulheres e identificou uma alta incidência desses distúrbios. Ressalta que seu interesse por esse assunto deu-se pela sua experiência enquanto fonoaudióloga docente da Universidade Estadual de Goiás, trabalhando a saúde vocal de professores, quando detectou uma alta incidência de distúrbios vocais nesta categoria. Assim selecionou uma amostra aleatória de 64 professores de oito

escolas distintas, onde os mesmos responderam alguns questionários e foram entrevistados, além de se levar em consideração questões como o espaço físico das escolas, relacionamento professor/aluno, entre outros. Como conclusão do seu trabalho, observou que há a necessidade urgente de estabelecer programas que levem em consideração a influência dos fatores do trabalho sobre a saúde profissional, visto que há uma grande relação entre estes.

MEDEIROS (2006) realizou um trabalho com o objetivo de conhecer o perfil dos professores em readaptação funcional definitiva por disfonia, determinar a prevalência e fatores associados à disfonia em professores e associar com fatores relacionados ao trabalho docente. Para tanto, fez um estudo descritivo de dados secundários referente ao perfil de professores em readaptação por disfonia, além de um estudo com amostra aleatória de 2.103 professores de 83 escolas. Concluiu que os problemas vocais que afastaram professores das salas de aula atingiu predominantemente as mulheres, 98% dos readaptados por disfonias foram professores do sexo feminino; que a disfonia mostrou-se associada a precárias condições de trabalho e que há a necessidade de implementar novas investigações para compreender melhor os fatores associados e subsidiar melhorias nas condições de trabalho.

## CAPÍTULO IV

### 4 MATERIAL E MÉTODOS

#### 4.1 ATRIBUTOS CONSIDERADOS NO TRABALHO

A voz é alterada todo tempo dependendo, além dos aspectos já citados anteriormente, também da carga afetiva. Assim, cada pessoa possui uma qualidade vocal individual e, portanto, algumas pessoas são mais suscetíveis às disfonias do que outras.

A preocupação é com um profissional diferenciado, aquele que ao produzir sua voz profissionalmente, como os profissionais da educação, tenham nela um instrumento “aliado” ao seu trabalho.

No trabalho docente a voz ocupa lugar de destaque, influenciando a relação professor-aluno. É um importante recurso para ganhar respeito, atenção e tornar o trabalho em sala de aula mais interessante, já que a qualidade vocal e a forma de expressão do professor influenciam a receptividade dos alunos a determinados ensinamentos.

Segundo JARDIM (2006), várias pesquisas têm mostrado prevalência de disfonia em professores. Esses resultados reforçam a necessidade de melhorias nas condições ambientais e organizacionais das escolas.

Ao se fazer referência aos profissionais da educação, a adequação do ambiente de trabalho, acústica, hidratação, tempo de uso da voz, quantidades de hora de trabalho, número de alunos por sala de aula, condições do ar no ambiente de trabalho (uso de giz anti-alérgico, outros poluentes em suspensão no ar, condições técnicas de condicionamento do ar), orientação preventiva sob supervisão médica e fonoaudiológica, se tornam fatores imprescindíveis.

Portanto, como já citado anteriormente, o objetivo desse trabalho é classificar os docentes quanto aos distúrbios vocais, valendo-se de alguns critérios já citados neste capítulo como fatores de risco para a voz, que se encontram no

quadro 2.1, a seguir. Nesse quadro, encontram-se os atributos utilizados pela escola para a classificação da gravidade de distúrbios vocais de docentes, onde a pontuação varia de zero (menor impacto sobre a voz) até 7. Na sequência, o quadro 2.2 apresenta as pontuações utilizadas pela escola para a classificação da gravidade de distúrbios vocais dos docentes.

Os dados utilizados no presente trabalho foram obtidos junto a um colégio da cidade de Curitiba, PR, mais especificamente, junto a fonoaudióloga do colégio, que forneceu os dados de 100 docentes, já avaliados e, também, já classificados em quatro classes: 21 como tendo distúrbios do tipo “leve”; 17 do tipo “leve a moderado”; 49 do tipo “moderado” e 13 do tipo “grave”. Esses dados encontram-se no apêndice I.

<b>Critério</b>	<b>Variáveis</b>	<b>Pontuação</b>
1- Fatores de risco	Ausentes	0
	RGE / ou suspeita de RGE	3
	Atopia / ou suspeita de atopia	3
	Tabagismo	3
2- Medidas de Prevenção: Aquecimento vocal e ingestão de água	Cumpre os 2	0
	Cumpre só 1	1
	Cumpre nenhum	2
3- Fono terapia	Não necessita / Alta fonoterápica	0
	Realiza regularmente	1
	Interrupção / Realiza irregularmente	3
	Não realiza, apesar da indicação	5
4- Sintomas vocais	Ausentes	0
	Não associados a patologia	3
	Associados a patologia	5
5- Patologia	Ausente	0
	Adquirida	3
	Congênita	5
6- Cirurgia	Não necessita	0
	Indicada, mas não realizada	4
	Realizada há – de 1 ano	3
	Realizada há + de 1 ano	2
7- Carga horária <u>total</u> atual (semanal)	até 25 horas/aula	1
	26 a 35 h/a	2
	36 a 45 h/a	4
	46 a 55 h/a	5
	Acima de 55 h/a	7
8- Tempo de docência	Até 5 anos ou acima de 20	2
	6 a 19 anos	1
9- Outra profissão da voz – não docente (exemplo: cantor)	Não	0
	Uso profissional da voz esporádico	1
	Uso profissional da voz freqüente	4
10- Uso de microfone	Sim	0
	Não	1

QUADRO 2.1 - ATRIBUTOS UTILIZADOS

<b>Pontuação</b>	<b>Classificação</b>	<b>Sugestão de carga horária</b>
0 a 10	Caso leve	Aumentar até 10 h/a
11 a 15	Caso leve a moderado	Aumentar até 5 h/a
16 a 25	Caso moderado	Manter carga horária
26 ou mais	Caso grave	Reduzir carga horária

QUADRO 2.2 - PONTUAÇÃO UTILIZADA PELA ESCOLA

Onde RGE ou DRGE é o Refluxo Gastro-Esofágico e as Atopias são alergias – no caso dos professores, a rinite alérgica é a mais comum.

## 4.2 TRATAMENTO DOS DADOS

De cada docente (padrão) foram obtidos 10 critérios e, a partir destes, 33 variáveis, conforme quadro 2.1 (capítulo II). Dentre os 100 docentes, nenhum apresentou o critério 9 (“outra profissão da voz – não docente”) e, da mesma forma, a variável referente ao critério 7 (“acima de 55 h/a”), de tal modo que os mesmos foram retirados. Desta forma, têm-se nove critérios (variáveis) originais e 29 critérios (variáveis) codificados de forma binária.

A fim de se obter um desempenho adequado das técnicas utilizadas neste trabalho, os atributos contidos no quadro 2.1 foram tratados de duas formas.

Na primeira, os 9 critérios definem as entradas variando de 0 a 7, conforme a pontuação que consta no quadro 2.1 anterior; na segunda, as variáveis são, primeiramente, transformadas em coordenadas binárias. Por exemplo, se o docente faz uso de microfone (critério 10), ele terá uma pontuação igual a “0” como entrada para o primeiro caso; para o segundo caso, ele terá uma entrada igual a (1 0), ou seja, o valor “1” por apresentar a primeira variável desse critério (sim) e “0” por não apresentar a segunda variável (não), conforme apêndices II e III.

Para se trabalhar com as saídas, mencionadas no quadro 2.2 anterior, para ambas as técnicas utilizadas (RNAs e FDLF), trabalhou-se de diversas formas, que se encontram no apêndice IV e explicadas a seguir:

1. com uma saída pertencente ao intervalo (0; 1), onde, para o grupo 1 (“caso leve”), a saída deverá estar no intervalo (0; 0,25); para o grupo 2 (“caso leve a moderado”),



a saída deverá estar entre  $[0,25; 0,5)$ ; para o grupo 3 (“caso moderado”), entre  $[0,5; 0,75)$  e para o grupo 4 (“caso grave”), entre  $[0,75; 1)$ ;

2. com apenas uma saída pertencente ao intervalo  $(0; 1)$ , onde de  $(0; 0,5]$  se tem um dos quatro casos e de  $(0,5; 1)$  se tem os demais três casos;

3. com duas saídas, transformadas em coordenadas binárias, onde de  $(1 0)$  se tem um dos casos e de  $(0 1)$  se tem os demais três casos;

4. com quatro saídas, transformadas em coordenadas binárias, onde para o grupo 1, a saída deverá ser  $(1 0 0 0)$ ; para o grupo 2,  $(0 1 0 0)$ ; para o grupo 3,  $(0 0 1 0)$  e para o grupo 4,  $(0 0 0 1)$ .

## CAPÍTULO V

### 5 IMPLEMENTAÇÃO DAS TÉCNICAS PARA RESOLUÇÃO DO PROBLEMA E ANÁLISE DOS RESULTADOS

#### 5.1 INTRODUÇÃO

Neste capítulo é descrita a forma como foram implementadas as duas técnicas de *Data Mining*: Redes Neurais Artificiais e Função Discriminante Linear de Fisher, bem como a comparação entre elas no que diz respeito à sua eficiência.

As técnicas abordadas (RNAs e FDLF) foram programadas no *MATLAB* e, em ambas foi utilizado o procedimento de avaliação *holdout*, que separa dois terços dos dados para o treinamento e um terço para testes. Além disso, a amostragem foi estratificada, assegurando que cada grupo seria representado de forma proporcional nos dois conjuntos de dados (treinamento e teste).

Ainda, com a finalidade de avaliar a classificação destas ferramentas de Reconhecimento de Padrões, calculou-se a Taxa Aparente de Erro (APER) para cada um dos resultados apresentados. Segundo MENDES (2008), essa taxa é definida como sendo a fração das observações no treinamento amostral referente ao reconhecimento errôneo pela função obtida, e deve ser interpretada como a proporção de observações classificadas incorretamente.

Essa taxa é calculada pela Matriz de Confusão que apresenta a situação real das observações nos grupos comparando-a com o reconhecimento apresentado pelo modelo encontrado.

O quadro 4.1 retrata a forma que tem a matriz de confusão, onde têm-se  $n_1$  observações de  $\pi_1$  e  $n_2$  observações de  $\pi_2$ .

		Classificação prevista	
		$\pi_1$	$\pi_2$
Classificação real	$\pi_1$	$n_{11}$	$n_{12}$
	$\pi_2$	$n_{21}$	$n_{22}$

QUADRO 4.1 – MATRIZ DE CONFUSÃO

Nesta matriz,  $n_{ij}$  são as observações pertencentes à população de origem  $i$  classificadas pela função discriminante como pertencentes à população  $j$ , ou seja, pertencentes a  $\pi_i$  classificadas corretamente se  $i=j$ , ou incorretamente se  $i \neq j$ . Assim sendo,  $n_{11}$  é o número de observações de  $\pi_1$  classificados corretamente como  $\pi_1$ ;  $n_{21}$  é o número de observações de  $\pi_2$  classificados incorretamente como  $\pi_1$ ; e assim por diante.

Com base nesses dados, a Taxa Aparente de Erro é calculada por (4.1).

$$APER = \frac{n_{12} + n_{21}}{n_1 + n_2} \quad (4.1)$$

onde  $n_1$  representa o tamanho da amostra 1 e  $n_2$ , o tamanho da amostra 2.

## 5.2 REDES NEURAIAS ARTIFICIAIS (RNAs)

Para realizar o treinamento das redes neurais foi utilizado o software *Matlab* 7.0.4, utilizando o algoritmo *back-propagation*, apresentado no item *c* da seção 3.3.1.2, do capítulo III.

As redes são constituídas pela camada de entrada; em algumas delas, pela camada escondida, e pela camada de saída. Foram utilizados dois tipos de entradas: nove entradas (apêndice II) que se referem aos critérios originais ou 29 entradas (apêndice III), que se referem aos critérios codificados. As redes ficaram com uma saída, duas saídas e quatro saídas (apêndice IV) e, a camada oculta, variou de zero a 20 neurônios. Os quatro testes estão descritos mais adiante.

Além disso, foram utilizados cinco conjuntos de pesos iniciais, sendo a escolha dos pesos aleatória no intervalo (0, 1), visto que uma escolha adequada para estes pesos pode diminuir o tempo de treinamento e aumentar a acurácia da rede. As taxas momento e aprendizagem foram ambas iguais a 0,8 em todos os testes. Cada uma das diferentes topologias das RNAs foi treinada por cinco vezes e o critério de parada utilizado foi a quantidade de padrões classificados corretamente.

### 5.2.1 Teste I

Neste teste, as redes neurais foram treinadas utilizando nove e 29 entradas; uma variação de zero a 20 neurônios na camada escondida e um neurônio na saída. Este neurônio representará a saída para os quatro grupos, sendo que para o grupo 1 (“caso leve”), a saída deverá estar no intervalo (0; 0,25); para o grupo 2 (“caso leve a moderado”), a saída deverá estar entre [0,25; 0,5); para o grupo 3 (“caso moderado”), entre [0,5; 0,75) e para o grupo 4 (“caso grave”), entre [0,75; 1).

Foram separados 75% dos padrões para treinamento e os 25% restantes para teste e, conforme já mencionado, de forma estratificada, ou seja, dos padrões referente ao caso 1, 75% foram para treinamento e 25% para teste, e assim sucessivamente, sendo que por serem obtidos números decimais nessas porcentagens, os mesmos foram arredondados. Assim, para este teste I e para o teste II, tem-se 72 padrões separados para treinamento e 28 para teste.

Os melhores resultados das cinco simulações encontram-se na tabela 4.1, sendo que os melhores resultados estão em negrito.

TABELA 4.1 – TESTE I

Camada Oculta	Entrada	Saída	% acerto	Camada Oculta	Entrada	Saída	% acerto
0	9	1	25	7	9	1	57,14
0	29	1	<b>32,14</b>	7	29	1	<b>60,71</b>
1	9	1	<b>53,57</b>	8	9	1	<b>57,14</b>
1	29	1	50	8	29	1	<b>57,14</b>
2	9	1	57,14	9	9	1	60,71
2	29	1	<b>60,71</b>	9	29	1	<b>71,43</b>
3	9	1	53,57	10	9	1	<b>71,43</b>
3	29	1	<b>60,71</b>	10	29	1	64,29
4	9	1	<b>57,14</b>	11	9	1	<b>71,43</b>
4	29	1	53,57	11	29	1	<b>46,43</b>
5	9	1	<b>57,14</b>	12	9	1	<b>60,71</b>
5	29	1	<b>57,14</b>	12	29	1	<b>60,71</b>
6	9	1	53,57	13	9	1	53,57
6	29	1	<b>75</b>	13	29	1	<b>60,71</b>

Camada Oculta	Entrada	Saída	% acerto
14	9	1	<b>57,14</b>
14	29	1	53,57
15	9	1	<b>64,29</b>
15	29	1	42,86
16	9	1	<b>57,14</b>
16	29	1	<b>57,14</b>
17	9	1	<b>53,57</b>
17	29	1	<b>53,57</b>

Camada Oculta	Entrada	Saída	% acerto
18	9	1	<b>53,57</b>
18	29	1	<b>53,57</b>
19	9	1	50
19	29	1	<b>53,57</b>
20	9	1	<b>57,14</b>
20	29	1	53,57

A partir da tabela 4.1, verifica-se que o melhor resultado é o que apresenta seis neurônios na camada oculta e 29 entradas, com 75% de casos classificados corretamente. O quadro 4.2 representa a matriz de confusão para esse resultado, que foi construída utilizando apenas o conjunto de teste (28 padrões), visto que esse conjunto é considerado o mais importante por mostrar a capacidade de generalização da rede.

		Classificação Prevista			
		CL	CLM	CM	CG
Classificação Real	CL	5	1	0	0
	CLM	0	5	0	0
	CM	0	2	11	0
	CG	0	1	3	0

QUADRO 4.2 – MATRIZ DE CONFUSÃO TESTE I

onde CL é o “caso leve”; CLM o é o “caso leve a moderado”; CM, o “caso moderado” e CG, “caso grave”.

A Taxa Aparente de Erro para esse teste é igual a  $(1+2+1+3)/28=0,25$ , ou seja, 25% dos casos foram classificados incorretamente.

### 5.2.2 Teste II

Neste teste as redes neurais foram treinadas utilizando também nove e 29 entradas; a camada escondida variando de zero a 20 neurônios e quatro neurônios na camada de saída. A saída ficou definida da seguinte forma: para o grupo 1, a saída deverá ser (1 0 0 0); para o grupo 2, (0 1 0 0); para o grupo 3, (0 0 1 0) e para o grupo 4, (0 0 0 1).

Os melhores resultados encontram-se na tabela 4.2, destacados em negrito.

TABELA 4.2 – TESTE II

Camada Oculta	Entrada	Saída	% acerto	Camada Oculta	Entrada	Saída	% acerto
0	9	4	<b>39,29</b>	11	9	4	<b>53,57</b>
0	29	4	<b>39,29</b>	11	29	4	50
1	9	4	<b>57,14</b>	12	9	4	<b>50</b>
1	29	4	42,85	12	29	4	<b>50</b>
2	9	4	<b>57,14</b>	13	9	4	<b>57,14</b>
2	29	4	46,43	13	29	4	53,57
3	9	4	<b>50</b>	14	9	4	<b>57,14</b>
3	29	4	<b>50</b>	14	29	4	<b>57,14</b>
4	9	4	50	15	9	4	<b>50</b>
4	29	4	<b>57,14</b>	15	29	4	<b>50</b>
5	9	4	<b>57,14</b>	16	9	4	<b>60,71</b>
5	29	4	<b>57,14</b>	16	29	4	53,57
6	9	4	<b>57,14</b>	17	9	4	<b>57,14</b>
6	29	4	50	17	29	4	50
7	9	4	<b>57,14</b>	18	9	4	<b>57,14</b>
7	29	4	<b>57,14</b>	18	29	4	50
8	9	4	53,57	19	9	4	<b>50</b>
8	29	4	<b>57,14</b>	19	29	4	<b>50</b>
9	9	4	50	20	9	4	<b>57,14</b>
9	29	4	<b>53,27</b>	20	29	4	50
10	9	4	<b>57,14</b>				
10	29	4	50				

Nesse caso, verifica-se que o melhor resultado, apresenta nove entradas com 16 neurônios na camada oculta e, com 60,71% de casos classificados corretamente. Pode-se visualizar a matriz de confusão, no quadro 4.3, a seguir, onde foram utilizados os 28 padrões separados para teste.

		Classificação Prevista			
		CL	CLM	CM	CG
Classificação Real	CL	5	0	1	0
	CLM	2	0	3	0
	CM	1	0	12	0
	CG	1	0	3	0

QUADRO 4.3 – MATRIZ DE CONFUSÃO TESTE II

A Taxa Aparente de Erro para esse teste é igual a 0,3928, ou seja, 39,28% dos casos foram classificados incorretamente.

Comparando as Taxas Aparentes de Erro desses dois testes, onde foram utilizados os quatro grupos, tem-se que o Teste I teve um resultado bem melhor, porém, mesmo assim, insatisfatório.

Com base nesses resultados, que foram pouco satisfatórios, foram feitos os Testes III e IV, a seguir.

### 5.2.3 Teste III

O Teste III apresenta o treinamento de redes neurais com nove e 29 entradas; uma variação de zero a 20 neurônios na camada escondida e um neurônio na saída, porém, diferentemente do Teste I, foram utilizados dois conjuntos a serem discriminados e quatro RNAs, uma para cada classe.

Na primeira RNA, um conjunto é formado pelos padrões da 1ª classe e o outro, pelo restante; na segunda RNA, um conjunto é formado pelos padrões da 2ª classe e o outro, pelo restante, e assim para a terceira e quarta RNAs.

No grupo 1, a saída deve estar no intervalo (0; 0,5] e no grupo 2, entre (0,5; 1). Cada uma dessas topologias foi treinada por cinco vezes e o critério de parada utilizado foi a quantidade de padrões classificados corretamente. Neste teste e no próximo (teste IV) têm-se 74 padrões separados para treinamento e 26 no conjunto de teste. Os melhores resultados encontram-se na tabela 4.3, destacados em negrito.

TABELA 4.3 – TESTE III

Camada Oculta	Entrada	Grupos	% Acerto	Camada Oculta	Entrada	Grupos	% Acerto
0	9	1 e 2,3,4	88,46	2	29	4 e 1,2,3	80,77
0	29	1 e 2,3,4	76,92	3	9	1 e 2,3,4	<b>88,46</b>
0	9	2 e 1,3,4	<b>92,31</b>	3	29	1 e 2,3,4	84,62
0	29	2 e 1,3,4	80,77	3	9	2 e 1,3,4	80,77
0	9	3 e 1,2,4	53,85	3	29	2 e 1,3,4	84,62
0	29	3 e 1,2,4	50	3	9	3 e 1,2,4	50
0	9	4 e 1,2,3	<b>92,31</b>	3	29	3 e 1,2,4	53,85
0	29	4 e 1,2,3	<b>92,31</b>	3	9	4 e 1,2,3	84,62
1	9	1 e 2,3,4	<b>92,31</b>	3	29	4 e 1,2,3	84,62
1	29	1 e 2,3,4	84,62	4	9	1 e 2,3,4	<b>88,46</b>
1	9	2 e 1,3,4	80,77	4	29	1 e 2,3,4	84,62
1	29	2 e 1,3,4	80,77	4	9	2 e 1,3,4	80,77
1	9	3 e 1,2,4	50	4	29	2 e 1,3,4	80,77
1	29	3 e 1,2,4	50	4	9	3 e 1,2,4	50
1	9	4 e 1,2,3	84,62	4	29	3 e 1,2,4	50
1	29	4 e 1,2,3	88,46	4	9	4 e 1,2,3	73,08
2	9	1 e 2,3,4	<b>92,31</b>	4	29	4 e 1,2,3	84,62
2	29	1 e 2,3,4	88,46	5	9	1 e 2,3,4	<b>88,46</b>
2	9	2 e 1,3,4	80,77	5	29	1 e 2,3,4	80,77
2	29	2 e 1,3,4	80,77	5	9	2 e 1,3,4	80,77
2	9	3 e 1,2,4	50	5	29	2 e 1,3,4	80,77
2	29	3 e 1,2,4	50	5	9	3 e 1,2,4	50
2	9	4 e 1,2,3	88,46	5	29	3 e 1,2,4	50

Camada Oculta	Entrada	Grupos	% Acerto
5	9	4 e 1,2,3	73,08
5	29	4 e 1,2,3	76,92
6	9	1 e 2,3,4	<b>92,31</b>
6	29	1 e 2,3,4	84,62
6	9	2 e 1,3,4	80,77
6	29	2 e 1,3,4	84,62
6	9	3 e 1,2,4	53,85
6	29	3 e 1,2,4	50
6	9	4 e 1,2,3	80,77
6	29	4 e 1,2,3	80,77
7	9	1 e 2,3,4	<b>92,31</b>
7	29	1 e 2,3,4	88,46
7	9	2 e 1,3,4	80,77
7	29	2 e 1,3,4	80,77
7	9	3 e 1,2,4	53,85
7	29	3 e 1,2,4	57,69
7	9	4 e 1,2,3	84,62
7	29	4 e 1,2,3	80,77
8	9	1 e 2,3,4	<b>92,31</b>
8	29	1 e 2,3,4	84,62
8	9	2 e 1,3,4	84,62
8	29	2 e 1,3,4	84,62
8	9	3 e 1,2,4	53,85
8	29	3 e 1,2,4	50
8	9	4 e 1,2,3	80,77
8	29	4 e 1,2,3	88,46
9	9	1 e 2,3,4	<b>92,31</b>
9	29	1 e 2,3,4	84,62
9	9	2 e 1,3,4	80,77
9	29	2 e 1,3,4	84,62
9	9	3 e 1,2,4	50
9	29	3 e 1,2,4	57,69
9	9	4 e 1,2,3	80,77
9	29	4 e 1,2,3	80,77
10	9	1 e 2,3,4	<b>92,31</b>
10	29	1 e 2,3,4	84,62
10	9	2 e 1,3,4	80,77
10	29	2 e 1,3,4	80,77
10	9	3 e 1,2,4	53,85
10	29	3 e 1,2,4	50
10	9	4 e 1,2,3	80,77
10	29	4 e 1,2,3	73,08
11	9	1 e 2,3,4	<b>88,46</b>
11	29	1 e 2,3,4	84,62
11	9	2 e 1,3,4	80,77
11	29	2 e 1,3,4	80,77
11	9	3 e 1,2,4	50
11	29	3 e 1,2,4	50
11	9	4 e 1,2,3	73,08
11	29	4 e 1,2,3	80,77
12	9	1 e 2,3,4	<b>88,46</b>
12	29	1 e 2,3,4	84,62
12	9	2 e 1,3,4	84,62
12	29	2 e 1,3,4	<b>88,46</b>
12	9	3 e 1,2,4	50
12	29	3 e 1,2,4	50
12	9	4 e 1,2,3	76,92
12	29	4 e 1,2,3	76,92

Camada Oculta	Entrada	Grupos	% Acerto
13	9	1 e 2,3,4	<b>88,46</b>
13	29	1 e 2,3,4	80,77
13	9	2 e 1,3,4	80,77
13	29	2 e 1,3,4	80,77
13	9	3 e 1,2,4	50
13	29	3 e 1,2,4	50
13	9	4 e 1,2,3	76,92
13	29	4 e 1,2,3	84,62
14	9	1 e 2,3,4	<b>92,31</b>
14	29	1 e 2,3,4	84,62
14	9	2 e 1,3,4	80,77
14	29	2 e 1,3,4	80,77
14	9	3 e 1,2,4	50
14	29	3 e 1,2,4	53,85
14	9	4 e 1,2,3	69,23
14	29	4 e 1,2,3	76,92
15	9	1 e 2,3,4	<b>92,31</b>
15	29	1 e 2,3,4	84,62
15	9	2 e 1,3,4	80,77
15	29	2 e 1,3,4	80,77
15	9	3 e 1,2,4	50
15	29	3 e 1,2,4	57,69
15	9	4 e 1,2,3	76,92
15	29	4 e 1,2,3	84,61
16	9	1 e 2,3,4	<b>96,15</b>
16	29	1 e 2,3,4	84,62
16	9	2 e 1,3,4	80,77
16	29	2 e 1,3,4	84,62
16	9	3 e 1,2,4	50
16	29	3 e 1,2,4	53,85
16	9	4 e 1,2,3	80,77
16	29	4 e 1,2,3	65,38
17	9	1 e 2,3,4	<b>92,31</b>
17	29	1 e 2,3,4	84,62
17	9	2 e 1,3,4	80,77
17	29	2 e 1,3,4	80,77
17	9	3 e 1,2,4	50
17	29	3 e 1,2,4	50
17	9	4 e 1,2,3	76,92
17	29	4 e 1,2,3	80,77
18	9	1 e 2,3,4	<b>92,31</b>
18	29	1 e 2,3,4	84,62
18	9	2 e 1,3,4	80,77
18	29	2 e 1,3,4	80,77
18	9	3 e 1,2,4	50
18	29	3 e 1,2,4	50
18	9	4 e 1,2,3	76,92
18	29	4 e 1,2,3	88,46
19	9	1 e 2,3,4	<b>92,31</b>
19	29	1 e 2,3,4	80,77
19	9	2 e 1,3,4	80,77
19	29	2 e 1,3,4	84,62
19	9	3 e 1,2,4	50
19	29	3 e 1,2,4	50
19	9	4 e 1,2,3	80,77
19	29	4 e 1,2,3	65,38
20	9	1 e 2,3,4	<b>92,31</b>
20	29	1 e 2,3,4	84,62



Camada Oculta	Entrada	Grupos	% Acerto
20	9	2 e 1,3,4	80,77
20	29	2 e 1,3,4	80,77
20	9	3 e 1,2,4	50

Camada Oculta	Entrada	Grupos	% Acerto
20	29	3 e 1,2,4	50
20	9	4 e 1,2,3	80,77
20	29	4 e 1,2,3	69,23

Nesse teste III, têm-se quatro situações diferentes e, portanto, quatro resultados. A partir da tabela 4.3, verifica-se que o melhor resultado apresentado pelo conjunto formado pelos padrões da 1ª classe e pelos padrões das demais é o possui 16 neurônios na camada oculta e nove entradas, com 96,31% de casos classificados corretamente.

Da mesma forma, o melhor resultado apresentado pelo conjunto formado pelos padrões da 2ª classe e pelos padrões das demais é o que possui 12 neurônios na camada oculta, com 29 entradas e com 88,46% de casos classificados corretamente. No conjunto formado pelos padrões da 3ª classe e pelos padrões das demais, tem-se um empate entre os resultados com sete, nove e 15 neurônios na camada oculta, todos com 29 entradas e com 57,69% de casos classificados corretamente. E, finalmente, para o conjunto formado pelos padrões da 4ª classe e pelos padrões das demais, o melhor resultado é a rede que possui zero neurônio na camada oculta, com nove ou 29 entradas e 92,31% de casos classificados corretamente.

Assim, o melhor resultado desse teste é o apresentado pelo conjunto formado pelos padrões da 1ª classe e pelos padrões das demais (96,31%).

Nesse teste e no seguinte, têm-se as matrizes de confusão referentes as quatro RNAs, cada uma com dois grupos a serem discriminados. O quadro 4.4, a seguir, indica a matriz de confusão referente a rede onde um conjunto é formado pelos padrões da 1ª classe e o outro pelo restante (2ª, 3ª e 4ª classes), para os 26 padrões (apenas para o conjunto de teste).

		Classificação prevista	
		G1	G2
Classificação real	G1	5	1
	G2	0	20

QUADRO 4.4 – MATRIZ DE CONFUSÃO TESTE III(1 e 2,3,4)

O quadro 4.5 representa o conjunto formado pelos padrões da 2ª classe e o outro pelo restante (1ª, 3ª e 4ª classes).

		Classificação prevista	
		G1	G2
Classificação real	G1	2	3
	G2	0	21

QUADRO 4.5 – MATRIZ DE CONFUSÃO TESTE III(2 e 1,3,4)

A terceira matriz apresenta a rede onde um conjunto é formado pelos padrões da 3ª classe e o outro pelo restante (1ª, 2ª e 4ª classes), porém com 7 e 9 neurônios na camada oculta.

		Classificação prevista	
		G1	G2
Classificação real	G1	4	9
	G2	2	11

QUADRO 4.6 – MATRIZ DE CONFUSÃO TESTE III (3 e 1,2,4)

Para o caso de 15 neurônios na camada oculta, tem-se a 4ª matriz, com a mesma porcentagem de casos classificados incorretamente, porém com uma classificação diferente.

		Classificação prevista	
		G1	G2
Classificação real	G1	2	11
	G2	0	13

QUADRO 4.7 – MATRIZ DE CONFUSÃO TESTE III(3 e 1,2,4)

Nesse caso, pode-se notar que os valores tendiam a “1”, sendo que os 13 últimos normalmente estariam corretos, pois estão no intervalo (0,5; 1). Dessa forma, a maior ocorrência era 50% de acertos.

Por fim, a quinta matriz, pela rede onde um conjunto é formado pelos padrões da 4ª classe e o outro pelo restante (1ª, 2ª e 3ª classes).

		Classificação prevista	
		G1	G2
Classificação real	G1	22	0
	G2	2	2

QUADRO 4.8 – MATRIZ DE CONFUSÃO TESTE III (4 e 1,2,3)

A tabela 4.7, a seguir, mostra o valor da Taxa Aparente de Erro para cada rede treinada no Teste III.

TABELA 4.4 – TAXA APARENTE DE ERRO TESTE III

QUADRO	APER
4.4	0,0384
4.5	0,1153
4.6	0,4230
4.7	0,4230
4.8	0,0769

Assim, a menor Taxa Aparente de Erro para o Teste III é a que se refere a quarta rede, com 3,84% de casos classificados incorretamente.

#### 5.2.4 Teste IV

O Teste IV apresenta o treinamento de redes neurais com nove e 29 entradas; uma variação de zero a 20 neurônios na camada escondida e dois neurônios na saída; e, como no teste anterior, foram utilizados dois conjuntos a serem discriminados e quatro RNAs, uma para cada classe. As saídas foram transformadas em coordenadas binárias, onde de (1 0) se tem um dos casos e de (0 1) se tem os demais três casos. Os melhores resultados encontram-se na tabela 4.5, destacados em negrito.

TABELA 4.5 – TESTE IV

Camada Oculta	Entrada	Grupos	% Acerto	Camada Oculta	Entrada	Grupos	% Acerto
0	9	1 e 2,3,4	96,15	1	29	2 e 1,3,4	<b>92,30</b>
0	29	1 e 2,3,4	<b>100</b>	1	9	3 e 1,2,4	76,92
0	9	2 e 1,3,4	80,76	1	29	3 e 1,2,4	61,53
0	29	2 e 1,3,4	80,76	1	9	4 e 1,2,3	84,61
0	9	3 e 1,2,4	57,69	1	29	4 e 1,2,3	84,61
0	29	3 e 1,2,4	53,84	2	9	1 e 2,3,4	<b>92,30</b>
0	9	4 e 1,2,3	84,61	2	29	1 e 2,3,4	84,61
0	29	4 e 1,2,3	84,61	2	9	2 e 1,3,4	80,76
1	9	1 e 2,3,4	<b>92,30</b>	2	29	2 e 1,3,4	84,61
1	29	1 e 2,3,4	<b>92,30</b>	2	9	3 e 1,2,4	57,69
1	9	2 e 1,3,4	88,46	2	29	3 e 1,2,4	57,69

Camada Oculta	Entrada	Grupos	% Acerto
2	9	4 e 1,2,3	84,61
2	29	4 e 1,2,3	88,46
3	9	1 e 2,3,4	80,76
3	29	1 e 2,3,4	80,76
3	9	2 e 1,3,4	<b>84,61</b>
3	29	2 e 1,3,4	<b>84,61</b>
3	9	3 e 1,2,4	57,69
3	29	3 e 1,2,4	73,07
3	9	4 e 1,2,3	<b>84,61</b>
3	29	4 e 1,2,3	<b>84,61</b>
4	9	1 e 2,3,4	80,76
4	29	1 e 2,3,4	76,92
4	9	2 e 1,3,4	88,46
4	29	2 e 1,3,4	<b>92,30</b>
4	9	3 e 1,2,4	73,07
4	29	3 e 1,2,4	57,69
4	9	4 e 1,2,3	88,46
4	29	4 e 1,2,3	84,61
5	9	1 e 2,3,4	84,61
5	29	1 e 2,3,4	76,92
5	9	2 e 1,3,4	<b>96,15</b>
5	29	2 e 1,3,4	80,76
5	9	3 e 1,2,4	57,69
5	29	3 e 1,2,4	57,69
5	9	4 e 1,2,3	84,61
5	29	4 e 1,2,3	88,46
6	9	1 e 2,3,4	76,92
6	29	1 e 2,3,4	80,76
6	9	2 e 1,3,4	<b>88,46</b>
6	29	2 e 1,3,4	80,76
6	9	3 e 1,2,4	57,69
6	29	3 e 1,2,4	65,38
6	9	4 e 1,2,3	84,61
6	29	4 e 1,2,3	<b>88,46</b>
7	9	1 e 2,3,4	76,92
7	29	1 e 2,3,4	80,76
7	9	2 e 1,3,4	<b>88,46</b>
7	29	2 e 1,3,4	84,61
7	9	3 e 1,2,4	53,84
7	29	3 e 1,2,4	61,53
7	9	4 e 1,2,3	84,61
7	29	4 e 1,2,3	<b>88,46</b>
8	9	1 e 2,3,4	80,76
8	29	1 e 2,3,4	80,76
8	9	2 e 1,3,4	84,61
8	29	2 e 1,3,4	84,61
8	9	3 e 1,2,4	61,53
8	29	3 e 1,2,4	57,69
8	9	4 e 1,2,3	84,61
8	29	4 e 1,2,3	<b>88,46</b>
9	9	1 e 2,3,4	<b>88,46</b>
9	29	1 e 2,3,4	80,76
9	9	2 e 1,3,4	84,61
9	29	2 e 1,3,4	80,76
9	9	3 e 1,2,4	61,53
9	29	3 e 1,2,4	61,53
9	9	4 e 1,2,3	84,61
9	29	4 e 1,2,3	<b>88,46</b>

Camada Oculta	Entrada	Grupos	% Acerto
10	9	1 e 2,3,4	84,61
10	29	1 e 2,3,4	80,76
10	9	2 e 1,3,4	84,61
10	29	2 e 1,3,4	<b>88,46</b>
10	9	3 e 1,2,4	57,69
10	29	3 e 1,2,4	57,69
10	9	4 e 1,2,3	<b>88,46</b>
10	29	4 e 1,2,3	<b>88,46</b>
11	9	1 e 2,3,4	80,76
11	29	1 e 2,3,4	<b>92,30</b>
11	9	2 e 1,3,4	84,61
11	29	2 e 1,3,4	80,76
11	9	3 e 1,2,4	57,69
11	29	3 e 1,2,4	61,53
11	9	4 e 1,2,3	88,46
11	29	4 e 1,2,3	84,61
12	9	1 e 2,3,4	84,61
12	29	1 e 2,3,4	80,76
12	9	2 e 1,3,4	<b>92,30</b>
12	29	2 e 1,3,4	88,46
12	9	3 e 1,2,4	80,76
12	29	3 e 1,2,4	73,07
12	9	4 e 1,2,3	84,61
12	29	4 e 1,2,3	88,46
13	9	1 e 2,3,4	84,61
13	29	1 e 2,3,4	84,61
13	9	2 e 1,3,4	<b>88,46</b>
13	29	2 e 1,3,4	84,61
13	9	3 e 1,2,4	61,53
13	29	3 e 1,2,4	65,38
13	9	4 e 1,2,3	<b>88,46</b>
13	29	4 e 1,2,3	<b>88,46</b>
14	9	1 e 2,3,4	80,76
14	29	1 e 2,3,4	80,76
14	9	2 e 1,3,4	<b>84,61</b>
14	29	2 e 1,3,4	80,76
14	9	3 e 1,2,4	57,69
14	29	3 e 1,2,4	50
14	9	4 e 1,2,3	<b>84,61</b>
14	29	4 e 1,2,3	<b>84,61</b>
15	9	1 e 2,3,4	88,46
15	29	1 e 2,3,4	<b>96,15</b>
15	9	2 e 1,3,4	80,76
15	29	2 e 1,3,4	80,76
15	9	3 e 1,2,4	65,38
15	29	3 e 1,2,4	65,38
15	9	4 e 1,2,3	84,61
15	29	4 e 1,2,3	84,61
16	9	1 e 2,3,4	<b>96,15</b>
16	29	1 e 2,3,4	84,61
16	9	2 e 1,3,4	88,46
16	29	2 e 1,3,4	80,76
16	9	3 e 1,2,4	65,38
16	29	3 e 1,2,4	65,38
16	9	4 e 1,2,3	84,61
16	29	4 e 1,2,3	84,61
17	9	1 e 2,3,4	<b>92,30</b>
17	29	1 e 2,3,4	88,46

Camada Oculta	Entrada	Grupos	% Acerto	Camada Oculta	Entrada	Grupos	% Acerto
17	9	2 e 1,3,4	84,61	19	29	1 e 2,3,4	84,61
17	29	2 e 1,3,4	80,76	19	9	2 e 1,3,4	80,76
17	9	3 e 1,2,4	69,23	19	29	2 e 1,3,4	80,76
17	29	3 e 1,2,4	73,07	19	9	3 e 1,2,4	65,38
17	9	4 e 1,2,3	84,61	19	29	3 e 1,2,4	76,92
17	29	4 e 1,2,3	88,46	19	9	4 e 1,2,3	88,46
18	9	1 e 2,3,4	<b>96,15</b>	19	29	4 e 1,2,3	84,61
18	29	1 e 2,3,4	76,92	20	9	1 e 2,3,4	84,61
18	9	2 e 1,3,4	92,30	20	29	1 e 2,3,4	84,61
18	29	2 e 1,3,4	80,76	20	9	2 e 1,3,4	<b>88,46</b>
18	9	3 e 1,2,4	69,23	20	29	2 e 1,3,4	<b>88,46</b>
18	29	3 e 1,2,4	69,23	20	9	3 e 1,2,4	65,38
18	9	4 e 1,2,3	88,46	20	29	3 e 1,2,4	65,38
18	29	4 e 1,2,3	88,46	20	9	4 e 1,2,3	<b>88,46</b>
19	9	1 e 2,3,4	<b>96,15</b>	20	29	4 e 1,2,3	84,61

Conforme resultados contidos na tabela 4.4, conclui-se que o melhor resultado corresponde a rede que classificou 100% dos casos corretamente, sem camada oculta, com 29 entradas e apresentado pelo conjunto formado pelos padrões da 1ª classe e pelos padrões das demais, ou seja, esses dados são linearmente separáveis.

Assim, analisando-se os melhores resultados para este teste IV envolvendo RNAs, tem-se que a maior porcentagem de dados classificados corretamente para o conjunto formado pelos padrões da 2ª. Classe e pelos padrões das demais é de 96,15%, com nove entradas e cinco neurônios na camada oculta. Para o conjunto formado pelos padrões da 3ª. classe e pelos padrões das demais, tem-se 80,77%, com nove entradas e 12 neurônios na camada oculta. Para o conjunto formado pelos padrões da 4ª. classe e pelos padrões das demais, tem-se 88,46%, com 4, 10, 11, 13, 18, 19 e 20 neurônios na camada oculta para nove entradas e 2, 5, 6, 7, 8, 9, 10, 12, 13, 17, 18 neurônios na camada oculta para 29 entradas.

A seguir têm-se as matrizes de confusão para os conjuntos de testes, para cada resultado. Observe-se que aqui, o conjunto de testes ficou com 26 padrões.

		Classificação prevista	
		G1	G2
Classificação real	G1	6	0
	G2	0	20

QUADRO 4.9 – MATRIZ DE CONFUSÃO TESTE IV (1 e 2,3,4)

		Classificação prevista	
		G1	G2
Classificação real	G1	5	0
	G2	1	20

QUADRO 4.10 – MATRIZ DE CONFUSÃO TESTE IV (2 e 1,3,4)

		Classificação prevista	
		G1	G2
Classificação real	G1	10	3
	G2	2	11

QUADRO 4.11 – MATRIZ DE CONFUSÃO TESTE IV (3 e 1,2,4)

		Classificação prevista	
		G1	G2
Classificação real	G1	22	0
	G2	3	1

QUADRO 4.12 – MATRIZ DE CONFUSÃO TESTE IV(4 e 1,2,3)

O quadro 4.12 mostra a matriz de confusão para a 4ª. rede, com nove entradas e, nos casos de 29 entradas, com 2, 5, 6, 7, 8, 10, 12, 13, 17, 18 neurônios na camada oculta.

No quadro 4.13, a seguir, tem-se a matriz para a mesma rede citada, porém com 29 entradas e nove neurônios na camada oculta

		Classificação prevista	
		G1	G2
Classificação real	G1	20	2
	G2	1	3

QUADRO 4.13 – MATRIZ DE CONFUSÃO TESTE IV (e)

A tabela 4.6, a seguir, mostra o valor da Taxa Aparente de Erro para o Teste IV.

TABELA 4.6 – TAXA APARENTE DE ERRO TESTE IV

QUADRO	APER
4.9	0,000
4.10	0,0384
4.11	0,1923
4.12	0,1153
4.13	0,1153

### 5.3 FUNÇÃO DISCRIMINANTE LINEAR DE FISHER (FDLF)

Para realizar o treinamento da FDLF foi utilizado o software *Matlab 7.0.4*. Foi feito um programa para o caso de duas populações e um outro para o caso de três ou mais populações. E, assim, foram aqui desenvolvidos os testes V e VI.

#### 5.3.1 Teste V

Neste teste têm-se os quatro grupos (“Caso leve”, “Caso leve a moderado”, “Caso moderado” e “Caso grave”) para serem discriminados, sendo que os resultados encontrados foram 57,14% para nove entradas e 28,57% para 29 entradas. Assim como nos testes I e II (onde têm-se os quatro grupos), foram separados 72 padrões para treinamento e 28 para teste.

A matriz de Confusão para o conjunto de testes (onde o número de padrões é 28), do melhor resultado (nove entradas) está no quadro 4.14.

		Classificação prevista			
		CL	CLM	CM	CG
Classificação real	CL	6	0	0	0
	CLM	2	3	0	0
	CM	1	5	4	3
	CG	0	0	1	3

QUADRO 4.14 – MATRIZ DE CONFUSÃO TESTE V

A Taxa Aparente de Erro para esse teste é igual a 0,4285, ou seja, 42,85% dos casos foram classificados incorretamente.

#### 5.3.2 Teste VI

Neste teste foram utilizados dois conjuntos a serem discriminados, sendo que foram realizados oito testes. Em dois deles, um conjunto é formado pelos

padrões da 1ª classe e o outro, pelo restante, com nove e depois com 29 entradas. Nos outros dois testes, um conjunto é formado pelos padrões da 2ª classe e o outro, pelo restante, também com nove e 29 entradas, e assim sucessivamente para o terceiro e quarto, conforme tabela 4.5. Nesse teste, assim como nos testes III e IV (onde têm-se dois conjuntos a serem discriminados), foram separados 74 padrões para treinamento e 26 para teste.

TABELA 4.7 – TESTE VI

Entrada	Grupos	% Acerto
9	1 e 2,3,4	<b>84,62</b>
29	1 e 2,3,4	-
9	2 e 1,3,4	53,85
29	2 e 1,3,4	-
9	3 e 1,2,4	69,23
29	3 e 1,2,4	-
9	4 e 1,2,3	80,77
29	4 e 1,2,3	-

O melhor resultado foi obtido para nove entradas, onde um conjunto é formado pelos padrões da 1ª classe e o outro, pelas demais três classes. Para 29 entradas não se encontrou resultado, pois há uma matriz singular nos dados.

A matriz de Confusão do melhor resultado (nove entradas, grupos 1 e 2,3,4) está no quadro 4.15.

		Classificação prevista	
		G1	G2
Classificação real	G1	6	0
	G2	4	16

QUADRO 4.15 – MATRIZ DE CONFUSÃO TESTE VI

A Taxa Aparente de Erro para esse teste é igual a 0,1538, ou seja, 15,38% dos casos foram classificados incorretamente.

#### 5.4 ANÁLISE DOS RESULTADOS

A tabela 4.8, a seguir, mostra os melhores resultados de todos os treinamentos realizados, especificando a topologia, as classes, o número de



neurônios nas camadas de entrada, escondida e de saída, e o percentual de acerto em cada simulação.

TABELA 4.8 – MELHORES RESULTADOS (RNAs: Testes I; II; III e IV; FDLF Testes V e VI)

TESTE	CLASSES	NEURÔNIOS			PERCENTUAL ACERTO
		ENTRADA	OCULTA	SAÍDA	
I	1; 2; 3; 4	29	6	1	75
II	1; 2; 3; 4	9	16	4	60,71
III	1 e 2,3,4	9	16	1	96,15
III	2 e 1,3,4	29	12	1	88,46
III	3 e 1,2,4	29	7, 9, 15	1	57,69
III	4 e 1,2,3	9	0	1	92,31
IV	1 e 2,3,4	29	0	2	100
IV	2 e 1,3,4	9	5	2	96,15
IV	3 e 1,2,4	9	12	2	80,77
IV	4 e 1,2,3	9 e 29	4, 10, 11, 13, 18, 19, 20 e 2, 5, 6, 7, 8, 9, 10, 12, 13, 17, 18	2	88,46
V	1; 2; 3; 4	9	-	4	57,14
VI	1 e 2,3,4	9	-	2	84,62
VI	2 e 1,3,4	9	-	2	53,85
VI	3 e 1,2,4	9	-	2	69,23
VI	4 e 1,2,3	9	-	2	80,77

Verifica-se que os melhores resultados foram obtidos no teste IV em que há dois conjuntos a serem discriminados, com um conjunto formado pelos padrões da 1ª classe, e o outro pelo restante; e assim sucessivamente para as demais classes. Neste teste IV, pode-se notar que o para o conjunto formado pelos padrões da 3ª classe e pelos padrões das demais, o desempenho da rede foi inferior (80,77% para este teste). Isso pode ser explicado analisando o gráfico contido na figura 4.1, onde os padrões estão plotados através da FDLF. Pode-se notar nesta figura que os grupos 1, 2 e 4 estão bem separados dos demais, enquanto o grupo 3 não.

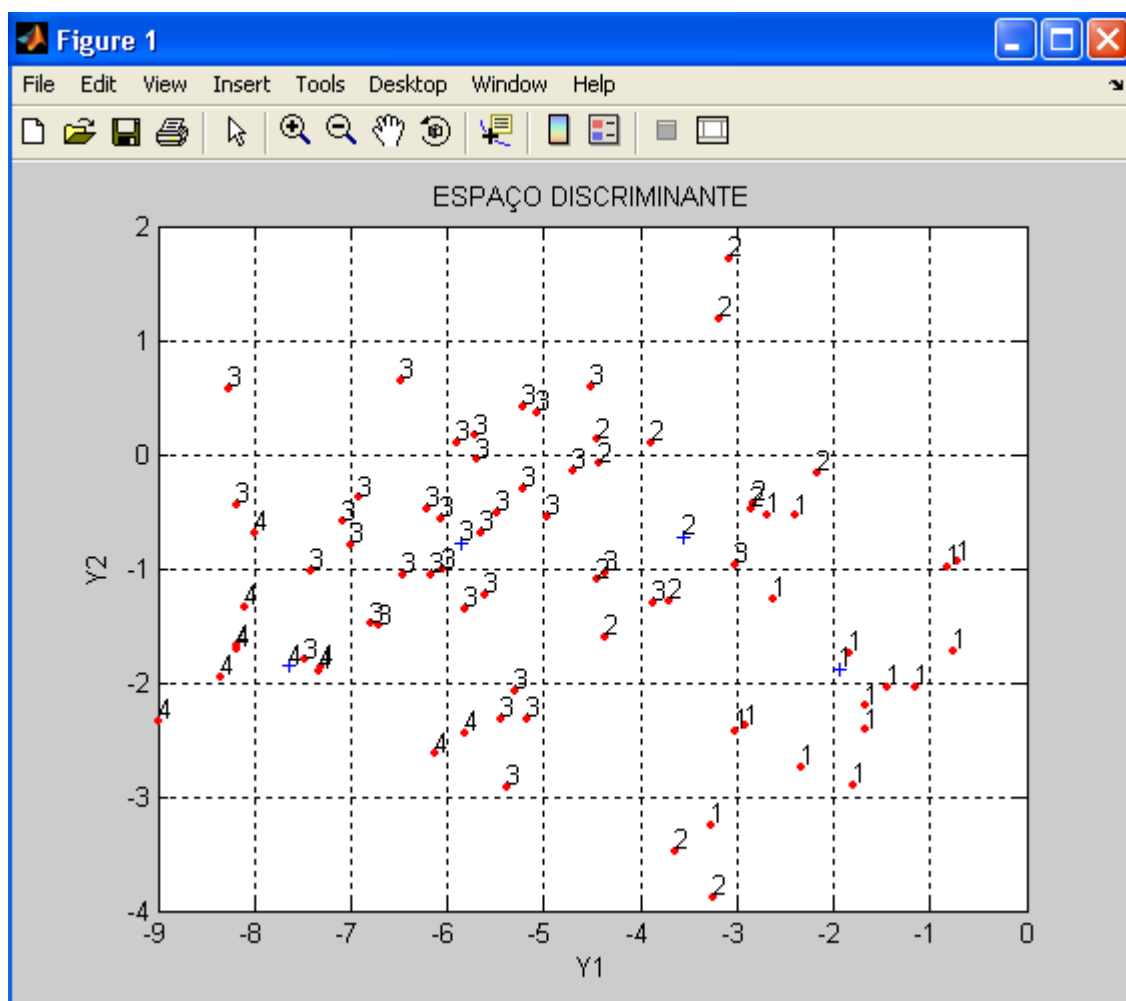


FIGURA 4.1 – ESPAÇO DISCRIMINANTE

## CAPÍTULO VI

### 6 CONSIDERAÇÕES FINAIS

#### 6.1 CONCLUSÕES

Este trabalho teve como objetivo comparar as técnicas RNAs e FDLF na classificação de 100 docentes de um colégio de Curitiba-PR, quanto a gravidade dos distúrbios vocais, com o objetivo de se obter um respaldo adicional na classificação da voz de novos docentes.

Com dados fornecidos pela fonoaudióloga da escola, foram aplicadas as duas técnicas. Como pode-se observar na tabela 4.8, as RNAs apresentaram um melhor desempenho no teste IV, ou seja, ao se ter dois conjuntos a serem discriminados: a 1ª. classe e as demais; a 2ª. classe e as demais; 3ª. classe e as demais e 4ª. classe e as demais. Para cada uma destas quatro situações, os pesos foram devidamente registrados. Assim, ao se apresentar um novo padrão (docente) com as suas respectivas variáveis (nove; valores dos critérios), que serão devidamente codificadas (gerando 29 valores), teremos que avaliar cada uma destas quatro situações para este docente, verificando em qual delas o seu percentual será maior. Tal situação, que fornecer o maior percentual será, então, a classe deste novo padrão, pelas RNAs.

Já a FDLF, como pode-se verificar na tabela 4.8, também apresentou um melhor desempenho ao se trabalhar com dois conjuntos a serem discriminados, assim como as RNAs. Assim, para classificar a voz de um novo docente, segundo a FDLF, basta “alimentar” as expressões (3.32.a) e (3.32.b) para a 1ª. situação; depois para a 2ª. ; 3ª. e 4ª. situações, verificando em qual delas o resultado se apresentará melhor. Tal situação, então, será a classe deste novo padrão, pela FDLF.

Deste modo, os resultados aqui obtidos através das duas técnicas (RNAs e FDLF) poderão ser utilizados na classificação da voz de docentes do colégio de

forma periódica (de tempos em tempos), assim como na de docentes que vierem a ser contratados, diagnosticando se os mesmos poderão aumentar, manter ou diminuir sua carga horária de trabalho.

Vale enfatizar que essas ferramentas podem ser utilizadas pelo especialista para auxiliá-lo nas tomadas de decisão, como um respaldo adicional, sendo que nunca porém poderão substituir o especialista.

## 6.2 SUGESTÕES PARA TRABALHOS FUTUROS

A partir dos resultados obtidos, as sugestões para trabalhos futuros são:

- Pesquisar ocorrência de problemas vocais nos docentes, verificando as principais causas. Poderiam ser utilizados dados do ambiente de trabalho como, por exemplo, quantidade de alunos por sala de aula, ar condicionado, acústica, utilização de giz e outros;
- Resolução do mesmo problema utilizando outras técnicas para o Reconhecimento de Padrões como, por exemplo, árvores de decisão, regressão linear múltipla, fazendo comparativos sempre objetivando obter a técnica com melhor desempenho;
- Usar casos adicionais de docentes para o problema aqui abordado, comparando os resultados com os aqui obtidos, averiguando se as técnicas apresentam (ou não) um melhor desempenho;
- Trabalhar com uma variação maior na quantidade de neurônios da camada oculta para o problema aqui apresentado;
- Desenvolver uma interface onde a escola colocaria os atributos (variáveis) do novo docente e o programa faria a classificação automática do mesmo facilitando, assim, o uso das técnicas apresentadas pelo usuário (escola);
- Utilizar a Análise dos Componentes Principais, com o objetivo de refinar ainda mais os resultados obtidos, como no trabalho de PAVANELLI (2007), pois o mesmo concluiu que essa ferramenta melhorou o desempenho da RNA.

- Calcular as médias e desvio padrão para cada teste e comparar esses valores.

## REFERÊNCIAS

ALECRIM, E. *Redes Neurais Artificiais*. Disponível em <http://www.infowester.com/redesneurais.php>. Acesso em: 26 de maio de 2009. InfoWester, 2004.

ALVES, V. *Avaliação de imóveis urbanos baseada em Métodos Estatísticos Multivariados*. Curitiba, 2005. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

AMORIM, S. N. M. C. *Distúrbio vocal e estresse: o efeito do trabalho na saúde de professores/as do ensino fundamental de Goiânia*. Goiás, 2006. Dissertação de Mestrado – Programa de Pós-Graduação em Ciências Ambientais e Saúde, Universidade Católica de Goiás.

BALDIN, M. S. Z. *Redes Neurais na Orientação Profissional*. Curitiba, 2002. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

BAPTISTELLA, M. *O Uso de Redes Neurais e Regressão Linear Múltipla na Engenharia de Avaliações: Determinação dos Valores Venais de Imóveis Urbanos*. Curitiba, 2005. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

BARRETO, G. A. *Redes Neurais Artificiais: Funções de Base Radial*. Disponível em [http://www.deti.ufc.br/~guilherme/TI016/slides\\_RBF.pdf](http://www.deti.ufc.br/~guilherme/TI016/slides_RBF.pdf). Acesso em 29 de dezembro de 2009. Ceará, CE: 2008.

BEHLAU, M. e PONTES, P. *Avaliação e Tratamento das Disfonias*. São Paulo, SP: Lovise, 1995.

BRAULIO, S. N. *Proposta de uma Metodologia para a Avaliação de Imóveis Urbanos Baseado em Métodos Estatísticos Multivariados*. Curitiba, 2005. Dissertação de

Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

FAYYAD, U.M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. *Advances in Knowledge Discovery & Data Mining*. AAAI/MIT, 1996.

FERRO, M; LEE, H. D. *O Processo de KDD - Knowledge Discovery in Database para Aplicações na Medicina*. Foz do Iguaçu, PR. Seminc 2001.

FIDELIS, M.V.; LOPES, H.S.; FREITAS, A.A. *Um Algoritmo Genético para Descobrir Regras de Classificação em Data Mining*. Anais do XIX Congresso Nacional da Sociedade Brasileira de Computação, vol. IV, p. 17-29, 2000.

FREITAS, A.A. *Uma Introdução a Data Mining*. *Informática Brasileira em Análise*. CESAR - Centro de Estudos e Sistemas Avançados do Recife. Ano II, n. 32, mai./jun. 2000.

GUIMARÃES, I. A; NETO, A. C. *Reconhecimento de Padrões: comparação de Métodos Multivariados e Redes Neurais*. Disponível em <http://publica.fesppr.br/index.php/rnti/article/viewFile/v1n1ART5/89>. Acesso em 28 de dezembro de 2009.

GUIMARÃES, S. M. A. *Voz – Importante instrumento de comunicação*. Departamento Saúde Escolar e Ocupacional. Bom Jesus. Curitiba, PR, 2001.

HAYKIN, S. *Redes Neurais – Princípios e prática*. 2 ed. Porto Alegre RS: Bookman, 2001.

JARDIM, R. *Voz, trabalho docente e qualidade de vida*. Belo Horizonte, 2006. Dissertação de Mestrado – Programa de Pós-Graduação em Saúde Pública da Universidade Federal de Minas Gerais.

JOHNSON, R.A.; WICHERN, D.W. *Applied Multivariate Statistical Analysis*. New Jersey, Prentice-Hall, inc., 1998.

KÓVACS, Z. L. *Redes Neurais Artificiais: fundamentos e aplicações, um texto básico*. São Paulo: Acadêmica, 1996.

Laboratório Nacional de Computação Científica. *Tutorial de Redes Neurais: aplicações em Bioinformática*. Disponível em <http://www.lncc.br/~labinfo/tutorialRN>. Acesso em 28 de dezembro de 2009.

LANDIM, F. *Análise Discriminante*. Disponível em <http://www.im.ufrj.br/~flavia/mad484/discrim.ppt>. Acesso em: 27 de maio de 2009. Instituto de Matemática, 2008.

LAWDER, J. M. R. *As interferências dos distúrbios vocais no trabalho docente no ensino fundamental*. Disponível em <http://www.utp.br>. Acesso em: 25 de maio de 2009. Universidade Tuiuti do Paraná. PR, 2009.

LEMOS, E. P. *Análise de Crédito Bancário com uso de Data Mining: Redes Neurais e Árvores de Decisão*. Curitiba, 2003. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

MACHADO, M. A. S; SOUZA, R. C; TORRES JR, R. G. *Previsão de séries temporais de falhas em manutenção industrial usando Redes Neurais*. Disponível em [http://www.uff.br/engevista/2\\_7Engevista03.pdf](http://www.uff.br/engevista/2_7Engevista03.pdf). Acesso em 29 de dezembro de 2009. Rio de Janeiro, RJ: 2005.

MARQUES, J. *Distúrbios vocais de professores*. Disponível em <http://cienciahoje.uol.com.br/128587>. Acesso em: 25 de maio de 2009. Ciência Hoje. RJ, 2008.

MARTINS, R. X. *Redes Neurais Artificiais, o aprendizado reproduzido em máquinas*. Disponível em <http://rxmartins.pro.br/engconh/rna.pdf>. Acesso em: 21 de junho de 2010. UNIS-MG, 2002.



MEDEIROS, A. M. *Disfonia e condições de trabalho das professoras da rede municipal de ensino de Belo Horizonte*. Belo Horizonte, 2006. Dissertação de Mestrado – Programa de Pós-Graduação em Saúde Pública, Universidade Federal de Minas Gerais.

MENDES, K. B. *O Uso de Redes Neurais Artificiais no Diagnóstico Preditivo dos Tipos mais Frequentes de Cefaléia*. Curitiba, 2008. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

OLIVEIRA, C. *Redes Neurais*. Disponível em <http://www.lem.ep.usp.br/Pef411/~Cristiano%20Oliveira/CristianoOliveira/Paginas/RedesNeurais.htm>. Acesso em 29 de dezembro de 2009. São Paulo, SP: 2002.

PASSOS, E. e GOLDSCHMIDT, R. *Data Mining: um guia prático*. Rio de Janeiro, RJ: Elsevier, 2005.

PAVANELLI, G. *Análise do Tempo de Duração de Processos Trabalhistas Utilizando Redes Neurais Artificiais como Apoio à Tomada de Decisões*. Curitiba, 2007. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

PETERMANN, R. J. *Modelo de Mineração de Dados para Classificação de Clientes em Telecomunicações*. Porto Alegre, 2006. Dissertação de Mestrado – Programa de Pós-Graduação em Engenharia Elétrica da Pontifícia Universidade Católica do Rio Grande do Sul.

PINHO, S. M. R. *Fundamentos em Fonoaudiologia tratando os distúrbios da voz*. Rio de Janeiro, RJ: Guanabara Koogan S.A., 1998.

PINHO, S. M. R. *Manual de higiene vocal para profissionais da voz*. Carapicuíba, SP: Pró-Fono, 1997.

RANGEL, L. *Mundo em sincronia*. Disponível em <http://cienciahoje.uol.com.br/noticias/2010/06/mundo-em-sincronia>. Acesso em 24 de junho de 2010.

Relatório do 3º. Consenso Nacional sobre Voz Profissional, “Voz e Trabalho: uma questão de saúde e direito do trabalhador”, (organizado pela Academia Brasileira de Laringologia e Voz, São Paulo, SP), Rio de Janeiro, RJ, 2004.

RUSSELL, S. J e NORVIG, P. *Inteligência Artificial*. Rio de Janeiro: Elsevier, 2004.

SAITO, M. M. C. K. *Sistemas Inteligentes em Controle e Automação de Processos*. Rio de Janeiro, RJ: Ciência Moderna LTDA, 2004.

SMIDERLE, A. *Data Mining como Suporte à Tomada de Decisões - uma Aplicação no Diagnóstico Médico*. XXXVI Simpósio Brasileiro de Pesquisa Operacional, p. 96-107, 2004.

STEINER, M.T.A. *Uma Metodologia para o Reconhecimento de Padrões Multivariados com Resposta Dicotômica*. Tese de Doutorado em Engenharia de Produção, UFSC, Florianópolis, SC, 1995.

STEINER, M.T.A. *Redes Neurais*. Notas de aula apresentadas na disciplina Aplicações de Pesquisa Operacional, Universidade Federal do Paraná, 2008.

TATAI, V. K. *Técnicas de Sistemas Inteligentes Aplicadas ao Desenvolvimento de Jogos de Computador*. Dissertação de Mestrado – Faculdade de Engenharia Elétrica e de Computação, UNICAMP: 2000.

TATIBANA, C. Y; KAETSU, D. Y. *Uma Introdução às Redes Neurais*. Disponível em <http://www.din.uem.br/~ia/neurais/>. Acesso em: 25 de maio de 2009.

TEIXEIRA, L. L. *O uso de técnicas de Estatística Multivariada no prognóstico de desistência de alunos em IES privadas: um estudo de caso na cidade de Foz do*

*Iguaçu – PR*. Curitiba, 2006. Dissertação de Mestrado – Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

TODESCO, J. L, *Reconhecimento de Padrões usando Rede Neuronal Artificial com uma função de Base Radial: uma aplicação na classificação de cromossomos humanos*. Florianópolis, 1995. Tese de Doutorado – Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina. Disponível em <http://www.eps.ufsc.br/teses/todesco/indice/index.html#index>. Acesso em 01 de julho de 2010.

TONSIG, S. L. *Redes Neurais Artificiais Multicamadas e o Algoritmo Backpropagation*. 2000. Disponível em <http://funk.on.br/esantos/doutorado/INTELIG%CANCIA%20ARTIFICIAL/T%C9CNICAS/REDES%20NEURASIS/Redes%20Neurais%20Artificiais%20Multicamadas/Backp.PDF> . Acesso em 15 de junho de 2010.

WIKIPÉDIA. Extração de Conhecimento. Disponível em [http://pt.wikipedia.org/wiki/Extra%C3%A7%C3%A3o\\_de\\_conhecimento](http://pt.wikipedia.org/wiki/Extra%C3%A7%C3%A3o_de_conhecimento). Acesso em 26 de maio de 2009.

WIKIPÉDIA. *Rede Neural*. Disponível em [http://pt.wikipedia.org/wiki/Rede\\_neural](http://pt.wikipedia.org/wiki/Rede_neural). Acesso em 26 de maio de 2009.

WITTEN, I. H; FRANK, E. *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed. Morgan Kaufmann, 2005.

ZUBEN. F. J. V; ATTUX, R. R. F. *Redes Neurais com Funções de Ativação de Base Radial*. Disponível em [ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia353\\_1s07/topico9\\_07.pdf](ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia353_1s07/topico9_07.pdf). acesso em 29 de dezembro de 2009.

## **APÊNDICES**

## APÊNDICE I DADOS ORIGINAIS

Levantamento de dados de professores para classificação de gravidade de distúrbios vocais													
Casos	Critérios										Pontuação	Classificação	Opinião fonoaudiológica
	1	2	3	4	5	6	7	8	9	10			
Caso 1	0	1	0	0	0	0	1	2	0	1	5	leve	leve
Caso 2	0	0	0	0	0	0	2	1	0	1	4	leve	leve
Caso 3	3 a	1	0	0	0	0	4	1	0	1	10	leve	leve
Caso 4	0	0	0	0	0	0	4	1	0	1	6	leve	leve
Caso 5	0	0	0	0	0	2	4	1	0	0	7	leve	leve
Caso 6	6 ra	0	0	3	0	0	4	1	0	1	15	leve a moderado	moderado
Caso 7	3 a	0	0	3	0	0	4	1	0	1	12	leve a moderado	leve a moderado
Caso 8	0	0	1	3	3	4	4	1	0	0	16	Moderado	grave
Caso 9	3 r	1	5	5	5	0	4	2	0	1	26	Moderado	grave
Caso 10	0	0	1	8	8	4	4	1	0	1	27	Moderado	moderado
Caso 11	6 ra	0	3	5	3	0	4	2	0	1	24	Moderado	grave
Caso 12	3 a	0	3	5	3	0	4	1	0	1	20	Moderado	moderado
Caso 13	6 ra	0	1	5	3	0	4	1	0	1	21	Moderado	moderado
Caso 14	3 a	0	0	5	5	4	1	2	0	1	21	Moderado	moderado
Caso 15	3 a	0	0	5	5	2	2	1	0	1	19	Moderado	moderado
Caso 16	0	2	5	0	3	0	4	1	0	1	16	Moderado	moderado
Caso 17	3 r	0	0	3	0	0	2	1	0	1	10	Leve	leve
Caso 18	0	0	1	5	3	0	5	1	0	1	16	Moderado	moderado
Caso 19	0	1	5	5	3	0	1	1	0	1	17	Moderado	moderado
Caso 20	0	0	1	5	3	0	1	1	0	1	12	leve a moderado	leve a moderado
Caso 21	3 r	0	5	5	3	0	2	2	0	1	21	Moderado	moderada
Caso 22	0	0	1	5	5	0	2	2	0	1	16	Moderado	leve a moderado
Caso 23	0	0	5	5	3	0	4	2	0	1	20	Moderado	moderado
Caso 24	0	1	0	0	0	0	1	1	0	1	4	Leve	leve
Caso 25	3 a	1	5	5	3	0	1	1	0	1	20	Moderado	moderado
Caso 26	0	0	0	5	5	0	1	1	0	1	13	leve a moderado	leve a moderado
Caso 27	3 a	0	1	5	3	0	4	2	0	1	19	Moderado	moderado
Caso 28	6 ra	1	5	5	3	0	2	1	0	1	24	Moderado	moderado
Caso 29	0	0	0	5	3	0	2	2	0	0	12	leve a moderado	moderado
Caso 30	3 r	0	5	5	3	0	4	2	0	1	23	Moderado	moderado
Caso 31	0	0	3	5	3	0	2	1	0	1	15	leve a moderado	moderado
Caso 32	0	0	1	0	0	0	2	2	0	1	6	Leve	leve
Caso 33	3 a	1	5	5	3	4	4	1	0	1	27	Grave	grave
Caso 34	6 ra	1	5	5	3	0	4	1	0	1	26	Grave	grave
Caso 35	6 ra	0	0	3	0	0	4	1	0	1	15	leve a moderado	leve a moderado
Caso 36	3 a	0	0	5	3	0	5	1	0	1	18	Moderado	moderado

<b>Caso 37</b>	0	2	5	5	3	0	1	2	0	1	19	Moderado	moderado
<b>Caso 38</b>	3 a	0	5	5	3	0	4	2	0	1	23	Moderado	grave
<b>Caso 39</b>	6 ra	0	0	3	3	0	4	2	0	1	19	Moderado	moderado
<b>Caso 40</b>	3 a	0	0	3	0	0	1	2	0	1	10	Leve	leve
<b>Caso 41</b>	0	1	5	5	3	0	1	1	0	1	17	Moderado	moderado
<b>Caso 42</b>	0	0	1	0	3	0	1	2	0	1	8	Leve	leve
<b>Caso 43</b>	3 a	1	5	5	3	0	4	1	0	1	23	Moderado	moderado
<b>Caso 44</b>	3 a	0	1	3	0	2	2	1	0	1	13	leve a moderado	moderado
<b>Caso 45</b>	3 a	0	0	0	0	3	1	1	0	1	9	Leve	moderada
<b>Caso 46</b>	0	0	0	3	0	2	2	1	0	1	9	Leve	moderada
<b>Caso 47</b>	3 a	0	5	5	5	0	5	1	0	0	24	Moderado	moderado
<b>Caso 48</b>	6 ra	1	5	5	3	0	4	2	0	1	27	Grave	grave
<b>Caso 49</b>	0	2	3	5	3	0	4	1	0	0	18	Moderado	moderado
<b>Caso 50</b>	0	0	1	5	3	0	2	1	0	0	12	leve a moderado	moderado
<b>Caso 51</b>	0	0	0	3	3	0	2	1	0	0	9	Leve	moderado
<b>Caso 52</b>	3 a	1	0	3	0	0	1	2	0	0	10	Leve	leve
<b>Caso 53</b>	3 a	1	0	3	0	0	1	1	0	0	9	Leve	leve
<b>Caso 54</b>	0	0	5	5	3	0	1	1	0	1	15	leve a moderado	moderado
<b>Caso 55</b>	0	0	1	5	3	0	1	1	0	1	11	leve a moderado	moderado
<b>Caso 56</b>	0	0	0	3	0	0	1	1	0	1	5	Leve	leve
<b>Caso 57</b>	3 a	1	5	3	0	0	1	1	0	1	14	leve a moderado	leve a moderado
<b>Caso 58</b>	0	1	5	5	3	0	4	1	0	1	19	Moderado	moderado
<b>Caso 59</b>	0	0	1	5	3	0	4	2	0	1	15	leve a moderado	moderado
<b>Caso 60</b>	0	0	5	5	3	0	4	2	0	1	19	Moderado	moderado
<b>Caso 61</b>	3 a	1	0	5	5	0	1	2	0	1	17	Moderado	leve a moderado
<b>Caso 62</b>	0	1	5	5	3	0	2	1	0	1	17	Moderado	moderado
<b>Caso 63</b>	0	2	5	5	3	0	4	1	0	1	20	Moderado	grave
<b>Caso 64</b>	3 t	0	0	3	3	2	4	2	0	0	17	Moderado	moderado
<b>Caso 65</b>	3 r	1	3	8	3	0	2	2	0	0	22	Moderado	grave
<b>Caso 66</b>	0	1	1	8	3	0	2	2	0	1	18	Moderado	moderado
<b>Caso 67</b>	0	1	0	3	0	0	4	1	0	1	10	Leve	leve
<b>Caso 68</b>	6 ra	1	0	8	5	0	4	2	0	1	27	Grave	grave
<b>Caso 69</b>	6 ra	1	0	8	5	0	4	1	0	0	25	Moderado	moderado
<b>Caso 70</b>	6 ra	1	0	3	3	0	1	2	0	0	16	Moderado	leve a moderado
<b>Caso 71</b>	0	1	0	0	0	0	1	2	0	0	4	Leve	leve
<b>Caso 72</b>	3 r	0	1	5	5	4	2	2	0	1	23	Moderado	moderado
<b>Caso 73</b>	6 rt	0	1	8	5	0	1	2	0	1	24	Moderado	moderado
<b>Caso 74</b>	0	1	5	5	3	0	1	2	0	1	17	Moderado	moderado
<b>Caso 75</b>	3 r	1	0	5	5	0	4	1	0	1	19	Moderado	moderado
<b>Caso 76</b>	0	1	5	5	3	0	2	2	0	1	19	Moderado	moderado
<b>Caso 77</b>	0	0	1	5	3	0	1	2	0	0	12	leve a moderado	leve a moderado
<b>Caso 78</b>	3 r	1	0	8	3	0	4	2	0	1	22	Moderado	moderado
<b>Caso 79</b>	0	1	0	0	0	0	2	2	0	1	6	Leve	leve
<b>Caso 80</b>	0	0	1	3	5	3	1	1	0	1	15	leve a moderado	leve a moderado

<b>Caso 81</b>	0	1	5	3	0	0	2	1	0	1	14	leve a moderado	leve a moderado
<b>Caso 82</b>	0	1	0	5	3	0	4	1	0	1	15	leve a moderado	moderada
<b>Caso 83</b>	6 ra	1	5	3	3	0	4	1	0	1	24	Moderado	grave
<b>Caso 84</b>	6 ra	0	0	3	3	0	5	2	0	1	20	Moderado	grave
<b>Caso 85</b>	3 r	1	3	5	5	0	4	1	0	1	23	Moderado	moderado
<b>Caso 86</b>	0	1	0	5	5	0	2	2	0	1	16	Moderado	leve a moderado
<b>Caso 87</b>	3 a	0	0	3	0	0	1	1	0	1	9	Leve	leve a moderado
<b>Caso 88</b>	3 a	1	0	0	0	0	2	2	0	1	8	Leve	leve
<b>Caso 89</b>	3 a	1	5	0	3	0	4	1	0	1	18	Moderado	moderado
<b>Caso 90</b>	6 ra	0	0	3	0	0	1	1	0	1	12	leve a moderado	leve a moderado
<b>Caso 91</b>	3 a	0	0	0	0	0	5	1	0	1	10	Leve	leve
<b>Caso 92</b>	0	1	0	3	0	0	2	1	0	1	8	Leve	leve
<b>Caso 93</b>	6 ra	0	0	0	0	0	4	2	0	1	13	leve a moderado	leve a moderado
<b>Caso 94</b>	0	0	1	3	3	0	1	2	0	1	11	leve a moderado	leve a moderado
<b>Caso 95</b>	3 a	0	0	0	0	3	1	2	0	0	9	Leve	leve a moderado
<b>Caso 96</b>	3 a	0	5	5	5	4	2	2	0	1	27	Grave	grave
<b>Caso 97</b>	6 ra	0	1	5	3	0	1	2	0	1	19	Moderado	moderado
<b>Caso 98</b>	3 a	0	0	0	0	0	1	2	0	1	7	Leve	leve
<b>Caso 99</b>	0	0	0	3	0	0	4	1	0	1	9	Leve	moderado
<b>Caso 100</b>	0	0	0	0	0	0	1	2	0	1	4	Leve	leve

Legenda:

3a – refere-se a variável Atopia/ ou suspeita de atopia

3r – refere-se a variável RGE/ ou suspeita de RGE

6ra – refere-se as variáveis RGE e Atopia

3t – refere-se a variável tabagismo

6rt– refere-se as variáveis RGE e tabagismo

## APÊNDICE II

### 9 ENTRADAS

Leve=[0 1 0 0 0 1 2 1;0 0 0 0 0 2 1 1;3 1 0 0 0 0 4 1 1;0 0 0 0 0 0 4 1 1;0 0 0 0 0 2 4 1 0;3 0 0 3 0 0 2 1 1;0 1 0 0 0 0 1 1 1;0 0 1 0 0 0 2 2 1;3 0 0 3 0 0 1 2 1;0 0 1 0 3 0 1 2 1;3 1 0 3 0 0 1 2 0;3 1 0 3 0 0 1 1 0;0 0 0 3 0 0 1 1 1;0 1 0 3 0 0 4 1 1;0 1 0 0 0 0 1 2 0;0 1 0 0 0 0 2 2 1;3 1 0 0 0 0 2 2 1;3 0 0 0 0 0 5 1 1;0 1 0 3 0 0 2 1 1;3 0 0 0 0 0 1 2 1;0 0 0 0 0 0 1 2 1];

Leve a moderado: [3 0 0 3 0 0 4 1 1;0 0 1 5 3 0 1 1 1;0 0 1 5 5 0 2 2 1;0 0 0 5 5 0 1 1 1;6 0 0 3 0 0 4 1 1;3 1 5 3 0 0 1 1 1;3 1 0 5 5 0 1 2 1;6 1 0 3 3 0 1 2 0;0 0 1 5 3 0 1 2 0;0 0 1 3 5 3 1 1 1;0 1 5 3 0 0 2 1 1;0 1 0 5 5 0 2 2 1;3 0 0 3 0 0 1 1 1;6 0 0 3 0 0 1 1 1;6 0 0 0 0 0 4 2 1;0 0 1 3 3 0 1 2 1;3 0 0 0 0 3 1 2 0];

Moderado: [6 0 0 3 0 0 4 1 1;0 0 1 8 8 4 4 1 1;3 0 3 5 3 0 4 1 1;6 0 1 5 3 0 4 1 1;3 0 0 5 5 4 1 2 1;3 0 0 5 5 2 2 1 1;0 2 5 0 3 0 4 1 1;0 0 1 5 3 0 5 1 1;0 1 5 5 3 0 1 1 1;3 0 5 5 3 0 2 2 1;0 0 5 5 3 0 4 2 1;3 1 5 5 3 0 1 1 1;3 0 1 5 3 0 4 2 1;6 1 5 5 3 0 2 1 1;0 0 0 5 3 0 2 2 0;3 0 5 5 3 0 4 2 1;0 0 3 5 3 0 2 1 1;3 0 0 5 3 0 5 1 1;0 2 5 5 3 0 1 2 1;6 0 0 3 3 0 4 2 1;0 1 5 5 3 0 1 1 1;3 1 5 5 3 0 4 1 1;3 0 1 3 0 2 2 1 1;3 0 0 0 3 1 1 1 1;0 0 0 3 0 2 2 1 1;3 0 5 5 5 0 5 1 0;0 2 3 5 3 0 4 1 0;0 0 1 5 3 0 2 1 0;0 0 0 3 3 0 2 1 0;0 0 5 5 3 0 1 1 1;0 0 1 5 3 0 1 1 1;0 1 5 5 3 0 4 1 1;0 0 1 5 3 0 4 2 1;0 0 5 5 3 0 4 2 1;0 1 5 5 3 0 2 1 1;3 0 0 3 3 2 4 2 0;0 1 1 8 3 0 2 2 1;6 1 0 8 5 0 4 1 0;3 0 1 5 5 4 2 2 1;6 0 1 8 5 0 1 2 1;0 1 5 5 3 0 1 2 1;3 1 0 5 5 0 4 1 1;0 1 5 5 3 0 2 2 1;3 1 0 8 3 0 4 2 1;0 1 0 5 3 0 4 1 1;3 1 3 5 5 0 4 1 1;3 1 5 0 3 0 4 1 1;6 0 1 5 3 0 1 2 1;0 0 0 3 0 0 4 1 1];

Grave= [0 0 1 3 3 4 4 1 0;3 1 5 5 5 0 4 2 1;6 0 3 5 3 0 4 2 1;3 1 5 5 3 4 4 1 1;6 1 5 5 3 0 4 1 1;3 0 5 5 3 0 4 2 1;6 1 5 5 3 0 4 2 1;0 2 5 5 3 0 4 1 1;3 1 3 8 3 0 2 2 0;6 1 0 8 5 0 4 2 1;6 1 5 3 3 0 4 1 1;6 0 0 3 3 0 5 2 1;3 0 5 5 5 4 2 2 1];





1 1 0; 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 0 0 0 1 0 1; 1 0 0 0 1 0 0 0 1 0 0 0  
0 1 0 1 0 1 0 0 0 1 0 0 0 0 1 0 1; 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 0 1 0  
1; 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1; 1 0 0 0 1 0 0 0 0 0 1 0 0 1  
0 1 0 1 0 0 0 0 0 1 0 1 0 0 1; 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 1;  
0 0 0 1 1 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0; 1 0 0 0 0 1 0 0 1 0 0 0 1 1 0 1  
0 1 0 0 0 0 1 0 0 1 0 0 1; 0 1 1 0 0 1 0 1 0 0 0 0 1 1 0 0 1 1 0 0 0 0 0 1 0 0 1 1 0; 0 1  
0 0 1 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 0 1 0 0 1; 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0 1 1  
0 0 0 1 0 0 0 1 0 0 1; 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 1 0 0 0 1 0 0 1; 0 1 0 0  
0 1 0 1 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 1 0 0 1 0 1; 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0 0  
0 0 1 0 0 1 0 0 1; 0 1 0 0 0 1 0 1 0 0 0 0 1 1 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1; 1 0 0 0 0 1  
0 1 0 0 0 0 0 1 0 1 0 1 0 0 0 0 0 1 0 0 1 0 1; 0 1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0 0 0 0  
0 1 0 0 1 0 1; 0 0 1 0 0 1 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 0 0 1 0 0 1 0 1; 0 1 1 0 1 0 0 0  
1 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 1 0 0 1; 1 0 0 0 1 0 0 1 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 1  
0 0 1 0 1];  
Grave=[1 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 1 1 0; 0 1 0 0 0 1 0 0 0 0 1 0  
0 1 0 0 1 1 0 0 0 0 0 1 0 1 0 0 1; 0 1 1 0 1 0 0 0 0 1 0 0 0 1 0 1 0 1 0 1 0 0 0 0 0 1 0 1 0 0  
1; 0 0 1 0 0 1 0 0 0 0 1 0 0 1 0 1 0 0 1 0 0 0 0 1 0 0 1 0 1; 0 1 1 0 0 1 0 0 0 0 1 0 0 1  
0 1 0 1 0 0 0 0 0 1 0 0 1 0 1; 0 0 1 0 1 0 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1;  
0 1 1 0 0 1 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 0 0 1; 1 0 0 0 0 0 1 0 0 0 1 0 0 1 0 1  
0 1 0 0 0 0 0 1 0 0 1 0 1; 0 1 0 0 1 0 0 0 0 1 0 0 1 1 0 1 0 1 0 0 0 0 1 0 0 1 0 1 0; 0 1  
1 0 0 1 0 1 0 0 0 0 1 1 0 0 1 1 0 0 0 0 0 1 0 1 0 0 1; 0 1 1 0 0 1 0 0 0 0 1 0 1 0 0 1 0 1  
0 0 0 0 0 1 0 0 1 0 1; 0 1 1 0 1 0 0 1 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0 1 1 0 0 1; 0 0 1 0  
1 0 0 0 0 0 1 0 0 1 0 0 1 0 1 0 0 0 1 0 0 1 0 0 1];



```
saida=[1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0
0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0 0; 1 0 0
0; 1 0 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0;
0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 1 0 0; 0 0 1 0; 0 0 1 0; 0 0
1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1
0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0;
0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1
0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1
0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 1 0; 0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1;
0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1; 0 0 0 1];
```