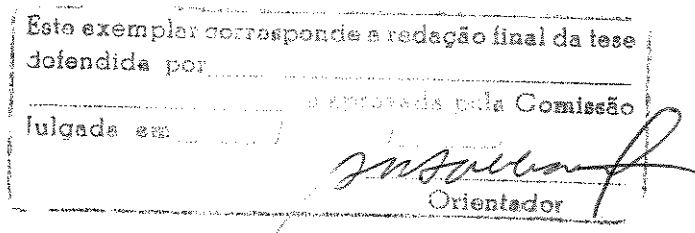


Universidade Estadual de Campinas  
Faculdade de Engenharia Elétrica  
Departamento de Engenharia Biomédica

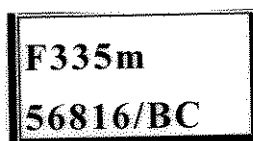
**Um Modelo Preditivo Não-Linear de  
Mortalidade em Pacientes Internados  
em Unidades de Terapia Intensiva,  
utilizando Redes Neurais Artificiais.**

Autor: Paulo Felipe Junior  
Orientador: Renato Marcos Endrizzi Sabbatini



Tese apresentada à Faculdade de Engenharia Elétrica da Universidade Estadual de Campinas como parte dos requisitos para obtenção do Título de Mestre em Engenharia Elétrica.

Janeiro de 1995



UNICAMP  
BIBLIOTECA CENTRAL  
SEÇÃO CIRCULANTE

UNIDADE	BC
Nº CHAMADA	IL/UNICAMP
	F-335m
V	EX
TOMBO BC	56816
PROC.	16/1174109
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	11.100
DATA	19/11/2004
Nº CPE	

CM0019414B-6

BIB ID. 309729

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

F335m

Felipe Junior, Paulo

Um Modelo preditivo não-linear de mortalidade em pacientes internados em unidades terapia intensiva utilizando redes neurais artificiais / Paulo Felipe Junior.- Campinas, SP: [s.n.], 1995.

Orientador: Renato Marcos Endrizzi Sabbatini  
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Prognostico. 2. Rede neurais (computação). 3. Mortalidade. 4. Unidade de tratamento intensivo. I. Sabbatini, Renato Marcos Endrizzi. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

UNICAMP  
BIBLIOTECA CENTRAL  
SEÇÃO CIRCULANTE

---

*“O homem não é nada mais que  
seus planos, ele existe somente à medida  
que se realiza; ele é entretanto nada  
mais do que um conjunto dos seus atos;  
nada mais que sua própria vida”*

*Sartre*

# Agradecimentos

---

Inicialmente, gostaria de agradecer ao Prof. Dr. Renato Marcos Endrizzi Sabbatini, meu orientador, por ter me aberto as portas do mundo da informática biomédica e ter me acompanhado nos meus primeiros passos no que considero a aventura da minha vida.

Aos Professores e colegas do Departamento de Engenharia Biomédica, principalmente os Profs. Alberto Cliquet Jr., Sérgio Santos Mühlen e Eduardo Tavares Costa por terem aceito mais um médico em seu Departamento e a mim, ensinado tanto.

Ao Prof. Dr. Renato Giuseppe Giovanni Terzi por ter acreditado na seriedade deste trabalho e ter aberto as portas da Unidade de Terapia Intensiva do Hospital de Clínicas da Unicamp para o desenvolvimento do mesmo.

Um agradecimento especial aos amigos Paulo Marcondes Carvalho Jr., Valeria Lacerda de Carvalho e Goytá Fernandes Villela Jr. por terem me auxiliado e me apoiado nos momentos mais difíceis durante o desenvolvimento e concretização desta dissertação.

Aos amigos do Núcleo de Informática Biomédica, Dirce Cesar, Alexandre Giovanni Martins, Jaime de Oliveira Ilha, Ronaldo Fonseca Pereira e Maria Lúcia A. Botelho que sempre estiveram do meu lado nos momentos felizes e nos momentos difíceis que se intercalaram durante este período.

Aos meus pais, Paulo Felipe Sobrinho e Dayse Salvadori Felipe, pela formação moral e educação esmerada que recebi, base do ser humano que sou hoje.

Aos meus sogros Blás Alonso Seguro e Maria de Lourdes Monis Seguro que acreditaram em mim e abriram as portas de sua casa para que grande parte deste trabalho fosse desenvolvido.

---

A minha cunhada Prof. Dra. Neusa Alonso que, apesar de toda a sua ocupação, encontrou tempo para me auxiliar na montagem e correção deste trabalho.

Finalmente, e não menos importante, a minha esposa Noemia Alonso Felipe que desde o momento que resolvi adotar uma nova especialidade, abandonando o exercício da clínica médica, sempre acreditou em mim e sempre me apoiou. A ela, a quem devo muitas horas de atenção e dedicação, que foram derivadas para elaboração deste trabalho, dedico esta dissertação de mestrado.

# Abreviaturas

---

<b>APACHE</b>	Acute Physiology and Cronic Health Evaluation
<b>AUC</b>	Area Under Curve
<b>BAM</b>	Bidirecional Associative Memory
<b>BASIC</b>	Beginner’s All-purpose Symbolic Instruction Code
<b><math>\chi^2</math></b>	Qui-quadrado
<b>DBF</b>	Data Base File
<b>ECG</b>	Eletrocardiograma
<b>FORTTRAN</b>	Formula Translation
<b>GI</b>	Gastrointestinal
<b>IBM</b>	International Business Machine
<b>IBM-PC</b>	International Business Machine Personal Computer
<b>ICC</b>	Insuficiência Cardíaca Congestiva
<b>Mb</b>	Megabyte
<b>MHz</b>	Mega Hertz ( $10^6$ Hertz)
<b>MPM</b>	Mortality Probability Model
<b>MS-DOS</b>	Microsoft Disk Operating System
<b>NYHA</b>	New York Heart Association
<b>PAM</b>	Pressão Arterial Média
<b>PC-AT</b>	Personnal Computer Advanced Tecnology
<b>RAM</b>	Random Access Memory
<b>RNA</b>	Redes Neurais Artificiais
<b>ROC</b>	Receiving Operator Characteristic
<b>SVGA</b>	Super Video Graphics Array
<b>UTI</b>	Unidade de Terapia Intensiva

# Índice Geral

---

Agradecimentos .....	iii
Abreviaturas.....	v
Índice Geral.....	vi
Índice de Figuras .....	ix
Índice de Tabelas .....	x
Resumo.....	xi
Abstract .....	xii
 1. Introdução .....	 14
1.1 Introdução .....	14
1.2 Equações Lineares x Redes Neurais .....	15
1.2.1 Introdução .....	15
1.2.2 O Modelo Linear em Medicina .....	16
1.3 O Sistema APACHE II .....	21
1.3.1 Introdução .....	21
1.3.2 O modelo do Sistema APACHE II.....	25
1.4 REDES NEURAIS ARTIFICIAIS .....	30
1.4.1 Introdução .....	30
1.4.2 Aprendizado Supervisionado.....	35
1.4.3 Princípios do Aprendizado Supervisionado .....	37
1.4.4 Ativação.....	44

1.4.5	<i>Implementação</i> .....	45
1.4.6	<i>Discussão</i> .....	47
<b>1.5</b>	<b>Redes Neurais: Aplicações na Medicina Intensiva</b> .....	<b>50</b>
1.5.1	<i>O computador nas Unidades de Terapia Intensiva</i> .....	50
1.5.2	<i>Redes Neurais Artificiais na UTI</i> .....	53
<b>1.6</b>	<b>Objetivos</b> .....	<b>56</b>
<b>2.</b>	<b>Materiais</b> .....	<b>59</b>
2.1	<b>Hardware</b> .....	<b>59</b>
2.2	<b>Software</b> .....	<b>59</b>
2.3	<b>Dados</b> .....	<b>61</b>
<b>3.</b>	<b>Métodos</b> .....	<b>62</b>
3.1	<b>Análise Estatística dos Dados</b> .....	<b>63</b>
3.2	<b>Variáveis Seleccionadas</b> .....	<b>64</b>
3.3	<b>Preparação dos arquivos de Treinamento e Teste</b> .....	<b>77</b>
3.4	<b>Avaliação do Desempenho da Rede Neural</b> .....	<b>78</b>
3.5	<b>Avaliação da Contribuição das Variáveis</b> .....	<b>81</b>
3.6	<b>Construção das Redes Neurais</b> .....	<b>82</b>
<b>4.</b>	<b>Resultados</b> .....	<b>85</b>
4.1	<b>Contribuição individual das variáveis de entrada</b> .....	<b>96</b>
<b>5.</b>	<b>Discussão</b> .....	<b>100</b>
<b>6.</b>	<b>Conclusão</b> .....	<b>110</b>
6.1	<b>Sugestões para Pesquisas Adicionais</b> .....	<b>111</b>



<b>7. Bibliografia.....</b>	<b>114</b>
<b>8. Apêndice A.....</b>	<b>124</b>
8.1 Descrição do programa Neuronet.....	124
8.2 Parâmetros de Definição e Treinamento.....	126
8.3 Características Técnicas .....	129
8.4 Desempenho.....	130
<b>9. Apêndice B.....</b>	<b>131</b>

# Índice de Figuras

---

<b>Figura 1.1(a / b):</b> Princípio da separabilidade de dois padrões por algoritmos de classificação automática, utilizando um espaço bidimensional simples.....	19
<b>Figura 1.2:</b> APACHE II – Sistema de classificação da severidade de uma doença.....	26
<b>Figura 1.3:</b> Modelo teórico do neurônio artificial. ....	31
<b>Figura 1.4:</b> Arquitetura de uma rede neural multicamadas (perceptron).....	35
<b>Figura 1.5:</b> Superfície de erro – mínimos locais e globais.....	37
<b>Figura 1.6:</b> A função de Transferência Sigmoidal.....	40
<b>Figura 3.1:</b> Pesos utilizados pelo APACHE II para a variação de temperatura retal.....	76
<b>Figura 3.2:</b> Método de Cálculo dos índices de desempenho de um sistema prognóstico.....	80
<b>Figura 4.1:</b> Desempenho das Redes Neurais (grupo de treinamento).....	89
<b>Figura 4.2:</b> Desempenho das Redes Neurais (grupo de teste) .....	90
<b>Figura 4.3:</b> Curvas de Treinamento das Redes Neurais.....	91
<b>Figura 4.4:</b> Curvas de aprendizado da Rede (30-5-1) para valores de Taxa de Aprendizado ( $\alpha$ –alfa) iguais a 0,5, 1,0, 1,5 e 2,0.b.....	92
<b>Figura 4.5:</b> Distribuição dos valores de saída das redes neurais treinadas (conjunto de teste).....	93
<b>Figura 4.6:</b> Distribuição dos valores de saída das redes neurais treinadas (conjunto de teste).....	94
<b>Figura 4.7:</b> Curvas ROC das Redes Neurais e do Índice APACHE II (conjunto de teste).....	95
<b>Figura 4.8:</b> Contribuição individual das variáveis de entrada para os experimentos realizados. ....	97

# Índice de Tabelas

---

<b>Tabela 1.1:</b> Valores atribuídos para as faixas etárias no cálculo do valor APS.....	27
<b>Tabela 1.2:</b> Valores da Variável Coeficiente Diagnóstico para casos não cirúrgicos.....	28
<b>Tabela 1.3:</b> Valores da Variável Coeficiente Diagnóstico para casos cirúrgicos. ....	29
<b>Tabela 3.1:</b> Variáveis selecionadas dentro do Banco de Dados da UTI utilizadas como Variáveis de Entrada para as Redes Neurais.....	65
<b>Tabela 3.2:</b> Valores de referência para o sódio sérico .....	71
<b>Tabela 3.3:</b> Valores de referência para o potássio sérico.....	71
<b>Tabela 3.4:</b> Valores de referência para a creatinina sérica.....	72
<b>Tabela 3.5:</b> Variação de referência dos valores do hematócrito .....	73
<b>Tabela 3.6:</b> Variação de referência da leucometria.....	73
<b>Tabela 3.7:</b> Escala de Coma de Glasgow.....	74
<b>Tabela 3.8:</b> Sistema de Classificação de Severidade da Variação de uma Doença, para a variável biológica Temperatura Corporal.....	75
<b>Tabela 4.1:</b> Resumo dos resultados obtidos (redes neurais e teste APACHE II).....	87
<b>Tabela 4.2:</b> Desempenho das Redes Neurais e do Sistema APACHE II na classificação do prognóstico de todos os registros dos 382 pacientes. ....	88
<b>Tabela 4.3:</b> Magnitude da média absoluta de saída dos experimentos realizados.....	97
<b>Tabela 4.4:</b> Variáveis que apresentaram o mesmo comportamento em todos os experimentos realizados.....	98
<b>Tabela 4.5:</b> Variáveis que apresentaram comportamento antagônico em pelo menos um dos experimentos realizados. ....	98
<b>Tabela 9.1:</b> ANÁLISE DE DETEÇÃO: r1nn20tr.OUT - CURVA ROC.....	131
<b>Tabela 9.2:</b> ANÁLISE DE DETEÇÃO: r1nn20ts.OUT - CURVA ROC.....	132
<b>Tabela 9.3:</b> ANÁLISE DE DETEÇÃO: r1nn15tr.OUT - CURVA ROC.....	133
<b>Tabela 9.4:</b> ANÁLISE DE DETEÇÃO: r1nn15ts.OUT - CURVA ROC.....	134
<b>Tabela 9.5:</b> ANÁLISE DE DETEÇÃO: r1nn10tr.OUT - CURVA ROC.....	135
<b>Tabela 9.6:</b> ANÁLISE DE DETEÇÃO: r1nn10ts.OUT - CURVA ROC.....	136
<b>Tabela 9.7:</b> ANÁLISE DE DETEÇÃO: r1nn5tr.OUT - CURVA ROC.....	137
<b>Tabela 9.8:</b> ANÁLISE DE DETEÇÃO: r1nn5ts.OUT - CURVA ROC.....	138
<b>Tabela 9.9:</b> ANÁLISE DE DETEÇÃO: r1nn5tr.OUT - CURVA ROC.....	139
<b>Tabela 9.10:</b> ANÁLISE DE DETEÇÃO: r1nn5ts.OUT - CURVA ROC.....	140
<b>Tabela 9.11:</b> ANÁLISE DE DETEÇÃO: r1nn5tt.OUT - CURVA ROC .....	141
<b>Tabela 9.12:</b> ANÁLISE DE DETEÇÃO: r1nn10tt.OUT - CURVA ROC .....	142
<b>Tabela 9.13:</b> ANÁLISE DE DETEÇÃO: r1nn15tt.OUT - CURVA ROC .....	143
<b>Tabela 9.14:</b> ANÁLISE DE DETEÇÃO: r1nn20tt.OUT - CURVA ROC .....	144
<b>Tabela 9.15:</b> ANÁLISE DE DETEÇÃO: APACHE II - CURVA ROC - (Grupo de Treino + Grupo de Teste).....	145
<b>Tabela 9.16:</b> ANÁLISE DE DETEÇÃO: APACHE II - CURVA ROC - Classificação do Grupo de Teste.....	146

## Resumo

---

Diversos sistemas foram desenvolvidos com o intuito de estimar a probabilidade de morte em Unidades de Terapia Intensiva, tais como o APACHE e o MPM, que são baseados em regressão logística. No presente trabalho, foi estudado o potencial das Redes Neurais Artificiais para a execução da mesma tarefa. Para tal, diversos experimentos foram realizados utilizando, como base, um *perceptron* de três camadas com regra de aprendizado por *backpropagation* onde o número de neurodos da camada intermediária foi de 5, 10, 15 e 20 elementos. Foram utilizados 282 registros de pacientes para o treinamento das redes neurais e 100 registros para o seu teste. Estes dados foram obtidos a partir de um banco de dados da UTI do HC-Unicamp, coletados no período de 1988 e 1992, utilizados previamente para calcular os Índices do APACHE II. Todas as redes apresentaram convergência dentro de 500 ciclos de treinamento sendo que o Erro Global Final, em todos os casos, foi inferior a 0,025. Para avaliar a performance de cada experimento foram utilizados, como índices de desempenho, os valores de acurácia, sensibilidade, especificidade, eficiência, preditividade positiva e negativa e o cálculo da área sob a Curva ROC. Para o conjunto de dados de treino, todas as redes foram capazes de prever o prognóstico do paciente com acurácias variando entre 97,6% (rede com 20 e 10 neurodos intermediários) e 98,6% (rede com 5 neurodos intermediários). No conjunto de dados de teste, os valores de acurácia variaram entre 76% (rede com 15 neurodos intermediários) e 81% (rede com 5 neurodos intermediários). Comparando os resultados obtidos entre os experimentos, concluímos que a melhor rede foi a de 5 neurodos intermediários que apresentou na classificação do grupo de teste uma acurácia de 81%, uma sensibilidade de 64%, uma especificidade de 86,7% e uma área sob a Curva ROC de 0,7871 e na classificação de todos os 382 registros, uma acurácia de 98,43%, uma sensibilidade de 97,22%, uma especificidade de 98,71% e uma área sob a Curva ROC de 0,9732. Este desempenho foi considerado superior aos resultados obtidos pelo Índice APACHE II que, para a mesma amostra, apresentou uma sensibilidade de 81,98%, uma especificidade de 68,75% e uma área sob a Curva ROC de 0,7673. Podemos concluir que o desempenho das Redes Neurais na execução desta tarefa foi muito bom porque o número de variáveis e sua variabilidade eram muito grandes, a associação com o prognóstico não era grande para muitas variáveis, as interações terapêuticas não foram consideradas e a casuística utilizada apresentava uma grande variedade de etiologias. Além disso, as Redes Neurais mostraram uma vantagem de obter seus parâmetros a partir de dados regionais ao invés de modelos universais, como é o caso do APACHE II.

# Abstract

---

Several systems have been devised with the purpose of estimating the probability of death in intensive care units (ICUs). These systems, such as APACHE and MPM, take the approach of logistic regression analysis. In this work, we have studied instead the applicability of artificial neural networks in performing the same task. Thus, we have conducted some experiments, using three-layer perceptrons with backpropagation learning rule, where the neurode count in the intermediate layer was either 5, 10, 15 or 20 elements. 282 patient records were used for training the neural networks, and a further 100 for testing them. The data originated from a database maintained by the ICU of the University Clinics Hospital, State University of Campinas, Brazil; they are from the 1988-92 period, and had been previously used for calculating APACHE II indexes. All networks converged within 500 cycles, with a final global error under 0.025 in all cases. In order to evaluate the experiments, for each of them the accuracy, sensitivity, specificity, efficiency, positive and negative predictability, as well as the calculated area under the ROC curve were used as performance indexes. For the training data set, all networks were capable of predicting the patient's prognosis, with the accuracy ranging from 97.6% (networks with 20 and 10 intermediate neurodes) to 98.6% (networks with 5 intermediate neurodes). For the test data set, accuracy has varied between 76% (networks with 15 intermediate neurodes) and 81% (with 5 intermediate neurodes). Upon comparing the results of all experiments, we conclude that the best network was that with 5 intermediate neurodes, which reached, in the test group, an 81% accuracy, a 64% sensitivity, an 86.7% specificity and a value of 0.7871 for the area under the ROC curve. When classifying all 382 records, it showed a 98.43% accuracy, a 64% sensitivity, an 86.7% specificity and an area under the ROC curve of 0.9732. This overall performance can be considered better than that achieved by the APACHE II index, which yielded, for the same sample, an 81.98% sensitivity, a 68.75% specificity, and an area under the ROC curve of 0.7673. We conclude that the performance of neural networks was very good, since the number of variables and their range of variation were both very large, the relationship between variables and prognosis was weak in many cases, therapeutic interactions were not considered, and there were many different etiologies in the case set. Furthermore, neural networks had the advantage of processing parameters originating from local data, instead of the universal models used by APACHE II.

---

## ***Introdução***

# 1. Introdução

---

## 1.1 Introdução

Nas últimas três décadas, a medicina, principalmente a medicina hospitalar, sofreu dramáticas alterações em seu desenvolvimento. Uma profissão que previamente enfatizava o diagnóstico como o auge da atitude médica, passou a dar especial atenção ao conceito terapêutico. Atualmente, tratamentos multidisciplinares complexos são empregados em estados avançados das doenças. Estes tratamentos, na sua maioria, são realizados dentro da Unidade de Terapia Intensiva (UTI). A UTI é a área hospitalar onde pacientes com alto risco de morte são assistidos por equipes multidisciplinares altamente treinadas, lançando mão da mais alta tecnologia para desempenhar a sua função. Devido a estas características, a rapidez das mudanças observadas e o crescente investimento nestes serviços de custo elevado, tem-se exigido indicações precisas para internação e determinação do período necessário de pacientes na UTI. Uma das melhores formas de se conseguir este objetivo é através da previsão estimada da alta, bem como das condições do paciente neste momento. Com esta proposta, diversos índices preditivos para risco de óbito foram desenvolvidos utilizando, principalmente, regressões logísticas. Dentre os principais podemos citar o APACHE (*Acute Physiologic and Chronic Health Evaluation*) e o MPM (*Mortality Probability Model*) (Seneff *et al.*, 1990). Apesar dos bons resultados obtidos com estes índices, sobretudo com o APACHE II, observa-se que em países em desenvolvimento, como o Brasil, a preditividade deste índice fica aquém do esperado.

Tal acontecimento decorre de fatos tais como rapidez com que um paciente grave é tratado, estado fisiológico do paciente quando dá entrada na UTI, condições pregressas do paciente, tais como desnutrição, moléstias endêmicas, etc. Portanto, a predição do estado de gravidade de um paciente em países como o nosso é de grande importância devido aos gastos deste serviço especializado, bem como a importância de se poder determinar, com antecedência, possíveis agravamentos na saúde do paciente para que medidas preventivas possam ser tomadas.

## **1.2 Equações Lineares x Redes Neurais**

### ***1.2.1 Introdução***

O prognóstico, ou a determinação da evolução futura da doença de um paciente, é um passo importante no processo de atenção médica, pois condiciona, entre outras coisas, as condutas curativas e medidas preventivas a serem tomadas ao longo do seguimento da doença.

Do ponto de vista da teoria da decisão médica, a elaboração de um prognóstico é bastante diferente, em muitos aspectos, da elaboração de um diagnóstico. Enquanto a maioria dos diagnósticos envolve relações causa-efeito relativamente bem definidas (embasadas pela etiologia da doença, pela fisiopatologia de suas manifestações, etc.), o prognóstico é quase sempre um raciocínio do tipo aproximativo ou estatístico, ou seja, caracteriza o que se chama de um modelo fraco (Wiederhold & Perreault, 1990). Estes são modelos preditivos nos quais existem muitas variáveis, cada qual com um pequeno efeito sobre a evolução do paciente. Nenhuma variável pode explicar uma parcela grande do efeito, e mesmo depois de se tentar identificar todas as variáveis



possivelmente correlacionadas, existe ainda uma variabilidade residual não explicada de certa magnitude. Esse fato explicaria, portanto, porque uma grande dose de experiência clínica, baseada em cuidadosa observação de dados empíricos e objetivos, é necessária para que um médico atinja um grau razoável de proficiência prognóstica.

---

### ***1.2.2 O Modelo Linear em Medicina***

Como o prognóstico é normalmente influenciado por um aglomerado de anormalidades encontradas em uma ou mais variáveis mensuráveis, as técnicas de classificação e de reconhecimento de padrões multivariados representam uma abordagem útil ao prognóstico quantitativo. Estas técnicas permitem aumentar o poder preditivo de um conjunto de variáveis, portanto é cada vez maior o número de pesquisas clínicas que utilizam o chamado modelo estatístico multivariado no estudo dessas interdependências. Tipicamente, essas pesquisas realizam estudos epidemiológicos prospectivos, nos quais um conjunto de variáveis padronizadas é coletado em um grupo controlado de pacientes (as variáveis independentes), e depois sua evolução é seguida ao longo de um período determinado de observação, ao final do qual se mede uma variável dependente, que pode ser um risco de doença, um fator de qualidade de vida, o período de tempo livre de sintomas, a ocorrência de doenças ou de óbito, etc. (Hulley & Cummings, 1988).

O modelo estatístico multivariado mais utilizado em Medicina ainda é o modelo linear, que busca um valor quantitativo de predição através de uma combinação polinomial dos fatores causais putativos:

$$y = \sum_{i=1}^n (b_i x_i) + e_i$$

onde:

---

$y =$	variável dependente (o prognóstico, ou efeito)
$x_i =$	a $i$ ésima variável independente (variável preditora ou causal)
$b_i =$	coeficiente linear correspondente à $i$ ésima variável independente (magnitude dos efeitos do vetor $x$ sobre $y$ )
$e =$	o erro, ou a variância de $y$ que não pode ser explicada pela variância em $x$ .

Diversas técnicas estatísticas, como a análise discriminante linear, a análise de regressão linear múltipla e a análise da regressão logística, têm sido extensamente empregadas para realizar tais tipos de estudos, atualmente denominados estudos de modelos de estrutura causal, no prognóstico médico. Elas são baseadas em modelos multidimensionais, ou seja, os padrões podem ser geometricamente representados como pontos em um espaço cartesiano multidimensional, onde os eixos são variáveis mensuráveis ou preditivas.

Algumas técnicas, como a análises de regressão multivariada, buscam ajustar um hiperplano (um plano em várias dimensões) ao aglomerado observado de pontos, generalizando o modelo de regressão bivariado, no qual se busca determinar, através da técnica dos mínimos quadrados, a reta que minimiza a distância média entre ela e os pontos  $[x,y]$ . O modelo linear simples não pode ser utilizado diretamente quando o prognóstico se refere a mortalidade, uma vez que a variável dependente só pode ter dois valores. Nesse caso, o método a ser utilizado é o da regressão logística múltipla, que modela a probabilidade ou risco de morte,  $p$ , através da equação modificada:

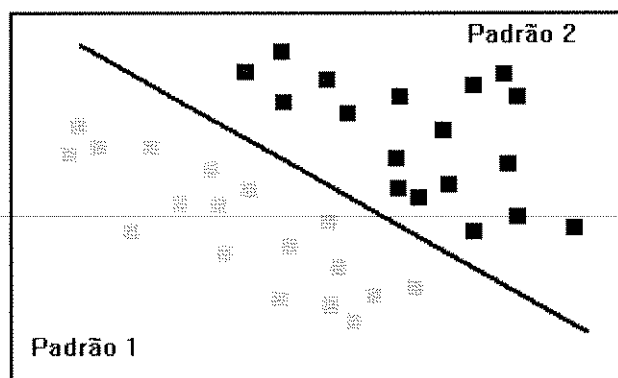
$$y = \ln\left(\frac{p}{1-p}\right) = \sum_{i=1}^n (b_i x_i) + e_i$$

Essa equação pode ser resolvida para  $p$  da seguinte forma:

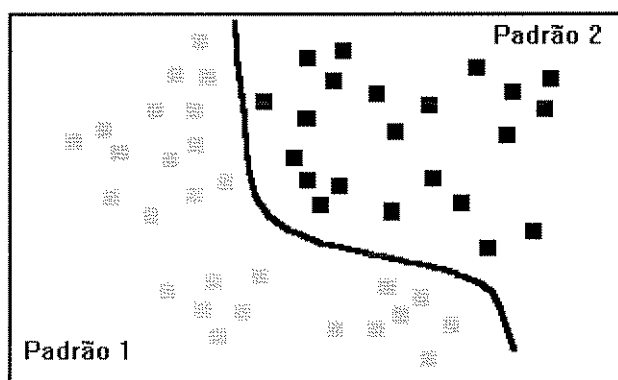
$$p = (1 + e^{-y})^{-1}$$

A vantagem da regressão logística é que as variáveis independentes podem ser de qualquer natureza, nominais ou contínuas. A equação final não é única, e depende dos critérios estatísticos usados para determinar o modelo. Um exemplo clássico da utilização da regressão logística no prognóstico de mortalidade é o trabalho de Knaus e colaboradores (1985), que desenvolveram um modelo para prever a mortalidade de pacientes em estado crítico.

Já a técnica de análise discriminante linear (Overall & Klett, 1972) tem por objetivo achar um hiperplano que segregue da melhor maneira possível pontos representativos das classes de pacientes (por exemplo, sobreviventes e não sobreviventes) em duas regiões distintas do espaço amostral (Figura 1.1-A). A análise discriminante linear determina esse hiperplano de modo a minimizar a variância dentro das classes e maximizar a variância entre as classes. Os cálculos necessários para a realização de análises estatísticas lineares são normalmente simples, e são baseados em procedimentos de álgebra matricial numérica.



A  
Separabilidade  
linear



B  
Separabilidade  
não-linear

**Figura 1.1(a / b):** Princípio da separabilidade de dois padrões por algoritmos de classificação automática, utilizando um espaço bidimensional simples.

Cada padrão na amostra é descrito geometricamente como um ponto no espaço cartesiano, onde cada eixo corresponde a uma variável descritora. **A:** a separabilidade linear é possível quando é achada uma reta que consiga separar uma classe da outra. **B:** a separabilidade não-linear implica na descoberta de uma curva com o mesmo propósito.

Entretanto, vários problemas interessantes na classificação de padrões médicos não podem ser abordados pelos métodos lineares, porque a separabilidade das classes de pacientes no espaço amostral só pode ser alcançada por superfícies de decisão mais complexas, não lineares (Figura 1.1-B). Quando isto acontece, podem ser

desenvolvidos algoritmos numéricos convencionais para a classificação de padrões não lineares, mas são normalmente muito ineficientes do ponto de vista computacional. A maioria deles utiliza técnicas iterativas complexas, que se assemelham ao aprendizado não supervisionado (Avanzolini *et al.*, 1990).

Existem também outros problemas do modelo estatístico linear que prejudicam sua aplicabilidade (Wiederhold & Perreault, 1990):

- interdependência entre as variáveis preditoras;
- especificação ou codificação inadequada das variáveis;
- ocorrência de valores omissos nas variáveis preditoras;
- variáveis distribuídas de forma não paramétrica;
- número pequeno de casos.

Várias investigações têm mostrado que problemas como estes, amplamente presente na maioria das pesquisas que visam desenvolver melhores modelos preditivos de apoio ao prognóstico médico e de estratificação de riscos de saúde, podem ser tratados mais adequadamente por meio de modelos conexionistas não lineares. As redes neurais artificiais podem ser usadas, entre outras coisas, para encontrar mapas quantitativos aproximados entre padrões de entrada e saída. Isto é conseguido através de uma técnica iterativa não linear, de aprendizagem através de exemplos, ou seja, apresentando-se à rede neural os dados de pacientes com evolução conhecida no processo patológico (por exemplo, morte ou sobrevivência depois de um dado tempo de observação). Usando-se um algoritmo de aprendizado apropriado, a rede é capaz de prever a evolução após ter alcançado um erro mínimo de classificação, sempre que as variáveis de entrada (preditivas) sejam relacionadas de uma maneira ordenada às variáveis a serem prognosticadas. Esta é uma propriedade interessante de um certo tipo de redes neurais artificiais de múltiplas camadas, que

conseguem realizar uma separação não intuitiva, complexa e não linear entre as classes de pacientes.

Assim, as redes neurais parecem apresentar uma capacidade ímpar de encontrar salientes ocultos no espaço de entrada, onde pouca coisa pode ser visualizada por métodos estatísticos convencionais, para não mencionar por decisões puramente humanas. Até o presente, os modelos conexionistas têm sido usados para prever a evolução das coronariopatias (Jayaweers *et al.*, 1993) e de miocardiopatias (Sabbatini *et al.*, 1993<sup>a</sup>) de pacientes em estado crítico sob cuidados intensivos (Felipe *et al.*, 1992; Tu & Guerriere, 1993), de ocorrência de infecções em feridas (Hudson *et al.*, 1992), de aparecimento de diabetes (Smith *et al.*, 1988), do prognóstico do câncer de mama (Ravdin & Clark, 1992) e da recorrência de câncer de cólon (Ferrer *et al.*, 1990), entre outros. Em algumas aplicações recentes na decisão médica, como no diagnóstico de infarto agudo no miocárdio, as redes neurais artificiais têm mostrado eficiência impressionante em relação a métodos quantitativos convencionais e também ao julgamento humano (Baxt, 1991 a e b). Como pretendemos demonstrar nesta dissertação as redes neurais artificiais representam uma abordagem extremamente interessante para a implementação de sistemas de apoio à decisão médica na área do prognóstico de sobrevida.

## **1.3 O Sistema APACHE II**

### ***1.3.1 Introdução***

O sistema de classificação APACHE II é uma versão revisada do sistema protótipo APACHE (*Acute Physiology and Chronic Health System Evaluation*). A

base do desenvolvimento do APACHE foi a hipótese que a severidade de doenças agudas poderiam ser medidas pela quantificação do grau de anormalidade de múltiplas variáveis fisiológicas. Este enfoque foi utilizado devido ao fato de que uma das maiores funções das unidades de terapia intensiva é detectar e tratar alterações fisiológicas agudas que possam comprometer a vida do paciente. Desta forma, um sistema que pretende classificar o grau de severidade de uma patologia, precisaria ser baseado em medidas fisiológicas objetivas sendo, o mais possível, independente das terapias adotadas.

Um estudo sobre a influência de diversas variáveis clínicas e laboratoriais na evolução dos pacientes foi utilizado para escolher e avaliar o peso de cada variável fisiológica. Este processo, foi seguido pelas sugestões de Gustafson *et al.*(1981) para a construção de escalas de severidade baseadas em pesos atribuídos para faixas de valores das variáveis. O sistema APACHE original utiliza 34 variáveis fisiológicas, sendo que a soma dos pesos destas variáveis constitui o valor APS (*Acute Physiology Score*).

O sistema de pesos foi baseado em uma escala que variava de 0 a 4, como mostra a Figura 1.2 .

O valor APS é obtido pela soma dos pesos relativos aos piores valores observados nas variáveis fisiológicas dimensionadas, dentro das primeiras 24 horas após a admissão do paciente na UTI. O período de 24 horas foi escolhido porque, frequentemente, garante que todos os valores das variáveis pertinentes estejam disponíveis, bem como estes valores sejam legitimados, individualmente, pelo julgamento clínico da equipe médica. Devido ao fato das doenças crônicas reduzirem a probabilidade de sobrevivência dos pacientes no decorrer de doenças agudas, o sistema APACHE original incorpora quatro letras (A,B,C e D) que correspondem à

faixa de variação entre saúde excelente (A) ao comprometimento crônico severo de sistemas orgânicos (D). Um incremento no valor APS está associado a um aumento do risco de óbito. Entretanto, o sistema APACHE original é extremamente complexo, necessitando um grande número de informações clínicas e laboratoriais que, em muitos hospitais, não fazem parte da sua rotina de procedimentos. O surgimento do APACHE II foi o resultado dos esforços de simplificação do método tornando-o clinicamente mais útil e estatisticamente mais acurado. Sua essência consiste na utilização de julgamentos clínicos e relações fisiológicas documentadas para a seleção das variáveis e a determinação de seus pesos e limiares de normalidade e anormalidade. O número de variáveis medidas diminuiu de 34 para 12. Medidas pouco frequentes de variáveis fisiológicas, tais como, osmolaridade sérica e nível de ácido láctico, foram excluídos, bem como variáveis potencialmente redundantes. As variáveis escolhidas foram:

1. Temperatura Retal
2. Pressão Arterial Média
3. Frequência Cardíaca
4. Frequência Respiratória
5. Oxigenação
6. pH Arterial
7. Sódio Sérico
8. Potássio Sérico
9. Creatinina Sérica
10. Hematócrito
11. Contagem de Leucócitos
12. Escala de Coma de Glasgow



O registro dos valores das variáveis continuou sendo o mesmo utilizado pelo APACHE original, ou seja, o pior valor medido de cada variável dentro das primeiras 24 horas após a admissão do paciente na UTI. Diferente do sistema original, o APACHE II exige que os valores das 12 variáveis sejam registrados para que seu valor possa ser calculado. Isto elimina o problema de valores não registrados, ou não mensurados, serem considerados como normais. Embora a medida dos valores dos gases arteriais possa ser inapropriada para alguns pacientes, a exclusão destes valores não é encorajada, sendo sua medida não registrada apenas em casos onde um julgamento clínico severo sugere que os resultados possam estar dentro dos limites normais.

Devido ao fato da idade avançada e problemas de saúde crônicos refletirem em uma diminuição da reserva fisiológica, estes fatores foram incluídos na medida do Índice APACHE II. A idade cronológica é um fator bem documentado de risco de óbito em doenças agudas, independentemente do grau de severidade. Observando as alterações fisiológicas agudas e a idade cronológica, constatou-se que três das quatro classificações de comprometimento orgânico crônico (B,C e D) estavam associadas com altas taxas de mortalidade. Entretanto, somente a classe que representava o maior comprometimento sistêmico (D) possuía uma marcada influência em relação ao prognóstico do paciente. Constatou-se também que as admissões não cirúrgicas e cirúrgicas de emergência, possuíam um risco potencial para óbito maior que as admissões cirúrgicas de caráter eletivo. Isto decorre, provavelmente, do fato de pacientes com patologias crônicas severas não serem considerados candidatos às cirurgias eletivas. Portanto, no APACHE II, as admissões não cirúrgicas ou cirúrgicas de emergência, com uma disfunção crônica severa de órgãos sistêmicos recebem, para

o cálculo do APS, 5 pontos adicionais, enquanto que as admissões cirúrgicas eletivas recebem apenas 2 pontos.

---

### ***1.3.2 O modelo do Sistema APACHE II***

O valor APS (*Acute Physiology Score*) é determinado pela soma dos pontos referentes aos dados das variáveis fisiológicas (A), idade (B) e doenças crônicas (C).

O sistema pode ser observado na Figura 1.2.

Para calcular o risco de óbito, a seguinte equação é utilizada:

$$RISCO = \frac{1}{1 + (EXP(-Y))}$$

Onde:  $Y = -3,517$   
+ 0,146(APS)  
+ 0,603 (somente se for caso cirúrgico de emergência)  
+ coeficiente diagnóstico (Tabelas 1.2 e 1.3)

De acordo com a Figura 1.2 os valores obtidos das 12 variáveis fisiológicas, descritas acima, recebem valores que variam de 0, para a faixa de normalidade, até 4, quando os desvios são considerados máximos.

Os valores atribuídos às idades cronológicas são descritos na Tabela 1.1.

(A)Variável Fisiológica	+4	+3	+2	+1	0	+1	+2	+3	+4
Temp. Retal (°C)	≥ 41	40,9 – 39		38,9 – 38,5	38,4 – 36	35,9 – 34	33,9 – 32	31,9 – 30	≤ 29,9
PAM	≥ 160	159 – 130	129 – 110		109 – 70		69 – 50		≤ 49
Freq. Resp.	≥ 50	49 – 35		34 – 25	24 – 12	11 – 10	9 – 6		< 5
Freq. Card.	≥ 180	179 – 140	139 – 110		109 – 70		69 – 55	54 – 40	< 39
Oxig. (mmHg)									
a) FIO <sub>2</sub> ≥ 0.05 → D(A-a)O <sub>2</sub> *	≥ 500	499 – 350	349 – 200		< 200				
b) FIO <sub>2</sub> < 0.05 → PaO <sub>2</sub>					> 70	70 – 61		60 – 55	< 55
pH arterial	≥ 7,7	7,69 – 7,6		7,59 – 7,50	7,49 – 7,33		7,32 – 7,25	7,24 – 7,15	< 7,15
Na <sup>+</sup> Sérico (mmol/l)	≥ 180	179 – 160	159 – 155	154 – 150	149 – 130		129 – 120	119 – 111	< 110
K <sup>+</sup> Sérico (mmol/l)	≥ 7	6,9 – 6,0		5,9 – 5,5	5,4 – 3,5	3,4 – 3,0	2,9 – 2,5		< 2,5
Creatinina (mg%)	≥ 3,5	3,4 – 2,0	1,9 – 1,5		1,4 – 0,6		< 0,6		
Hematócrito (%)	≥ 60		59,9 – 50	49,9 – 46,0	45,9 – 30		29,9 – 20		< 20
Leucóc (1000/mm <sup>3</sup> )	≥ 40		39,9 – 20	19,9 – 15,0	14,9 – 3,0		2,9 – 1,0		< 1,0

Escala de Glasgow (pontos):

Abertura Ocular	Espontânea	4	Resposta Verbal	Orientado	5	Resp. Motora	Verbal	6
	Verbal	3		Confuso	4		Loc. Dor	5
	Dor	2		Inapropriada	3		Flex. Normal	4
	S/Resposta	1		Sons	2		Decorticação	3
				S/Resposta	1		Decerebração	2
							o S/Resposta	

(A) Total APS (Acute Physiology Scores) : \_\_\_\_\_ (soma dos pesos das variáveis)

$$^*D(A-a)O_2 = (710 \times FIO_2) - \frac{PaCO_2}{0,8} - PaO_2$$

**(C) Pontos de Doença Crônica**

APACHE II Score  
(A) + (B) + (C) = \_\_\_\_\_

Na presença de doenças crônicas ou de comprometimento imunológico, é utilizado o critério descrito na Figura 1.2, ou seja, atribui-se 5 pontos para os casos não cirúrgicos e cirúrgicos de emergência e 2 pontos para os casos cirúrgicos eletivos.

**Tabela 1.1:** Valores atribuídos para as faixas etárias no cálculo do valor APS.

Idade	Pontos
<= 44	0
45 – 54	2
55 – 64	3
65 – 74	5
>= 75	6

A variável coeficiente diagnóstico, utilizada na equação acima, corresponde a um valor proporcional à severidade da patologia da qual o paciente é portador. O sistema APACHE II classifica 29 entidades não cirúrgicas (Tabela 1.2) e 21 cirúrgicas (Tabela 1.3).

**Tabela 1.2:** Valores da Variável Coeficiente Diagnóstico para casos não cirúrgicos.

Patologias		Pesos
Falência ou Insuficiência Respiratória devido a:		
	Asma/Alergia	-2,108
	Doença Pulmonar Obstrutiva Crônica	-0,367
	Edema Pulmonar (não Cardiogênico)	-0,251
	Parada Respiratória	-0,168
	Embolo Pulmonar	-0,128
	Infecção	0,000
Falência ou Insuficiência Cardiovascular devido a:		
	Hipertensão	-1,798
	Distúrbios do Ritmo Cardíaco	-1,368
	Insuficiência Cardíaca Congestiva	-0,424
	Choque Hemorrágico/Hipovolemia	0,493
	Doenças das Artérias Coronárias	-0,191
	Sepsis	0,113
	Parada Cardíaca	0,393
	Choque Cardiogênico	-0,259
	Aneurisma Dissecante da Aorta	0,731
	Torácica/Abdominal	
Traumáticas:		
	Traumas Múltiplos	-1,228
	Trauma Cefálico	-0,517
Neurológicas:		
	Doenças Convulsivas	-1,228
	Hipertensão Craniana	0,723
Outras:		
	Overdose de drogas	-3,353
	Cetoacidose Diabética	-1,507
	Sangramento Gastrointestinal	0,334
<ul style="list-style-type: none"> <li>Caso a patologia vigente não se enquadre em nenhum dos grupos acima citados, deve-se considerar qual o maior sistema orgânico vital que é o responsável pela admissão na UTI e classificá-lo em dos grupos abaixo:</li> </ul>		
Sistema:		
	Metabólico/Renal	-0,885
	Respiratório	-0,890
	Neurológico	-0,759
	Cardiovascular	0,470
	Gastrointestinal	0,501

**Tabela 1.3:** Valores da Variável Coeficiente Diagnóstico para casos cirúrgicos.

<b>Pós-operatório</b>	<b>Pesos</b>
Múltiplos Traumas	-1,684
Admissões devido doenças cardiovasculares crônicas	-1,376
Cirurgias Vasculares Periféricas	-1,315
Cirurgias de Válvulas Cardíacas	-1,261
Craniotomia devido a neoplasia	-1,245
Cirurgia Renal devido a neoplasia	-1,204
Transplante Renal	-1,042
Trauma Cefálico	-0,955
Cirurgia Torácica devido a neoplasia	-0,802
Craniotomia devido a hipertensão craniana	-0,788
Laminectomia e outras cirurgias da coluna espinal	-0,699
Choque hemorrágico	-0,682
Sangramento gastrointestinal	-0,617
Cirurgia gastrointestinal devido a neoplasia	-0,248
Insuficiência respiratória pós-cirúrgica	-0,140
Perfuração/Obstrução gastrointestinal	0,060
<ul style="list-style-type: none"> <li>• Para pacientes em pós-operatório admitidos na UTI devido a sepsis ou infarto do miocárdio, usar os pesos correspondentes aos pacientes não cirúrgicos.</li> <li>• Caso o pós-cirúrgico vigente não se enquadre em nenhum dos itens acima citados, deve-se considerar qual o maior sistema orgânico vital que é o responsável pela admissão pós-cirúrgica na UTI e classificá-lo em dos grupos abaixo:</li> </ul>	
Neurológico	-1,150
Cardiovascular	-0,797
Respiratório	-0,610
Gastrointestinal	-0,613
Metabólico/Renal	-0,196

## 1.4 REDES NEURAIS ARTIFICIAIS

### 1.4.1 Introdução

Uma rede neural artificial consiste dos seguintes elementos básicos (Simpson, 1990):

- Elementos de processamento (“neurodos” ou “nodos”), também referidos como nodos (Figura 1.3); interconectados entre si na forma de uma topologia ou rede. Um nodo recebe sinais binários ou contínuos provindos do meio exterior ou de outros nodos da rede, através de conexões diretas (“sinapses”), processa estes sinais de alguma forma e envia o produto do processamento para outros nodos ou para o exterior. Os sinais de entrada formam um vetor  $A = a_1, a_2 \dots a_i \dots a_n$ , onde  $a_i$  é o sinal ou nível de atividade da  $i$ -ésima conexão de entrada. Tipicamente existem dois tipos de conexões: conexões excitatórias, ou positivas, e conexões inibitórias, ou negativas, que tendem a aumentar ou diminuir, respectivamente, a ativação de um nodo. Associado a cada conexão  $a_j$  entre dois nodos existe um valor ajustável denominado peso; formando portanto outro vetor  $W_j = w_{1j}, w_{2j} \dots w_{ij} \dots w_{nj}$ , onde  $w_{ij}$  corresponde ao peso da conexão entre os nodos  $a_i$  e  $b_j$ . No modelo generalizado pode existir também um parâmetro extra  $\Theta_j$ , modulado pelo peso  $w_{\Theta j}$ , que corresponde a um limiar a ser excedido para haver ativação do nodo. Para calcular o valor de saída  $b_j$  do nodo, realiza-se a soma ponderada de todos os  $a_i w_{ij}$ , e o resultado é aplicado a uma função de ativação, que usualmente é não linear (sigmóide, por exemplo), e que equivale ao limiar de disparo do neurônio;



- 31



Uma das formas mais comuns de utilização das redes neurais artificiais é como um classificador ou reconhecedor de padrões. Um conjunto de sinais  $I = i_1, i_2 \dots i_j \dots i_n$  é aplicado à camada de entrada com  $n$  nodos. O padrão de ativação desta camada é propagado para as camadas restantes, até produzirem um padrão de saída  $O = o_1, o_2 \dots o_j \dots o_m$ , com  $m$  nodos. Obtemos assim um mapeamento de  $I$  em  $O$ , que pode ter diversas aplicações:

- em uma rede heteroassociativa, a natureza dos padrões de entrada e saída são diferentes. Por exemplo,  $I$  pode ser uma sequência de amostras tomadas de um sinal biológico, como um ECG, enquanto  $O$  representaria as classes diagnósticas (interpretação do ECG). Em uma outra aplicação médica, o padrão de entrada teria um nodo binário associado a cada sintoma possível ( $I$  é o vetor de sintomas), e o padrão de saída teria um nodo binário associado a cada diagnóstico possível a partir daquele conjunto de sintomas ( $O$  é o vetor de diagnósticos). Desta forma a rede é capaz de classificar padrões, interpretar sinais e imagens, realizar diagnósticos, etc.
- em uma rede autoassociativa, os padrões de entrada e saída têm a mesma natureza. Por exemplo,  $I$  pode ser uma matriz de pixels correspondentes a uma imagem bidimensional: os valores de ativação de um nodo seria proporcional ao nível de cinza de um pixel da imagem. Na saída,  $O$  seria também uma matriz de pixels: a rede poderia realizar funções de manipulação de imagens, como compressão, detecção de bordas, filtragem, reconstituição de imagens incompletas ou degradadas, etc.

O aprendizado é outro conceito muito importante na caracterização de redes neurais artificiais. Aprendizado é definido como uma modificação realizada na matriz

de pesos  $W$  de uma rede, no sentido de otimizar um mapeamento dos padrões de entrada e saída. Embora em casos simples fosse possível ajustar a matriz  $W$  manualmente, de modo a obter um mapeamento único, é desejável que este processo se dê de forma auto-organizada, ou seja, a rede, através de algum algoritmo apropriado, auto-ajusta os seus pesos durante a progressão do aprendizado. No aprendizado de tipo supervisionado, uma série de pares de padrões de entrada e saída são apresentados simultaneamente à rede, em cada iteração do processo de treinamento. A rede computa então uma função de discrepância ou erro entre o padrão de saída desejado e o obtido naquela iteração, e ajusta os pesos das sinapses em função da magnitude do erro. Existem redes neurais artificiais capazes de aprender continuamente, ou seja, a cada introdução de um exemplo não classificável, a rede efetua um aprendizado *on-line*, adaptando rapidamente seus pesos de modo a acomodar a nova classe. Mais comumente, entretanto, as redes neurais apresentam aprendizado *off-line*, ou seja, a fase de aprendizado é separada da fase de classificação, e cada novo exemplo introduzido leva a um reaprendizado total. Um grande número de algoritmos de aprendizado foi e tem sido desenvolvido para redes neurais artificiais. Num dos algoritmos de aprendizado mais conhecidos, a retropropagação (*backpropagation*), uma quantidade delta proporcional ao erro é aplicada retroativamente a cada um dos pesos, a partir da camada de saída até a camada de entrada (Rumelhart *et al.*, 1986). Existem diversos teoremas que estudam as condições de convergência de uma rede deste tipo, ou seja, a redução gradual da função de erro em direção a um mínimo global. Embora seja muito utilizada, a retropropagação, é menos eficiente em velocidade de aprendizado, do que outros algoritmos publicados na literatura (Kinser *et al.*, 1990). Outros algoritmos, como o do aprendizado hebbiano (Brown *et al.*, 1990), procuram imitar os mecanismos

sinápticos de aprendizado que ocorreriam em redes neurais biológicas, tais como a potenciação pós-sináptica retardada.

Outro tipo de aprendizado possível seria o *não supervisionado*. Em redes capazes deste tipo de aprendizado, os exemplos são apresentados apenas na camada de entrada da rede, e ela se auto-organiza de modo a conseguir uma separação máxima entre as classes naturalmente existentes nos exemplos. Desta forma, a rede se comporta como um algoritmo de aglomeração (*cluster analysis*) do tipo divisivo (Lippman, 1987). Um exemplo deste tipo de rede é o mapa auto-organizado, desenvolvido por Kohonen (1984).

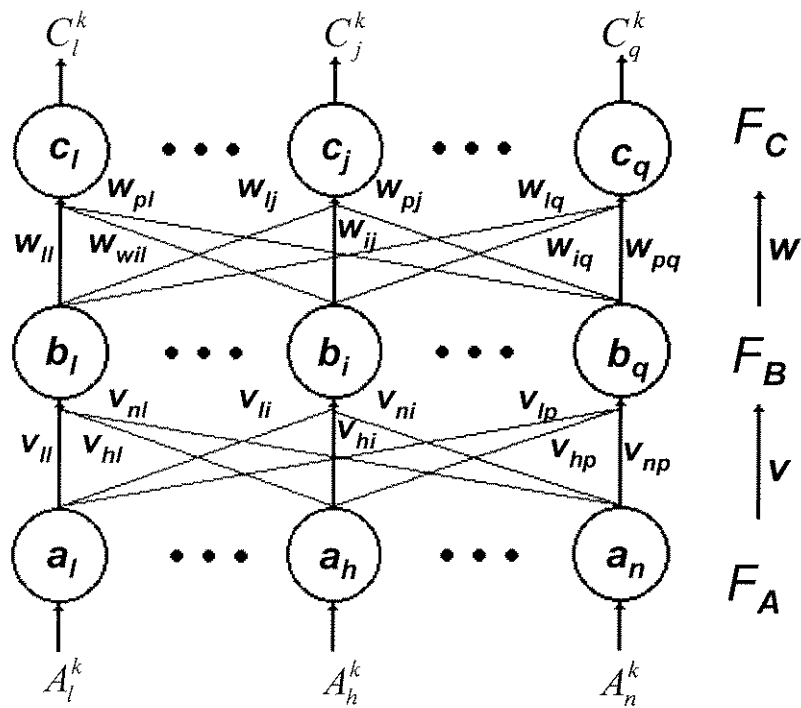
Finalmente, um importante parâmetro de operação de uma rede neural é a forma como os padrões aprendidos, ou incorporados à rede são recuperados. Se a arquitetura da rede é de tipo ativação progressiva (*feedforward*), ao se apresentar o padrão a ser classificado à camada de entrada *I*, as camadas sucessivas são ativadas por meio da propagação do sinal através das conexões ponderadas, até determinar um padrão de ativação final na camada de saída *O*, que corresponderá à classificação. Os perceptrons são bons exemplos de redes deste tipo. De outro lado, as arquiteturas de rede com ativação regressiva (*feedback*), ao se apresentar um padrão *I* de entrada, a própria camada devolve o padrão de saída, *O*. A rede denominada Memória Associativa Bidirecional (BAM) é um exemplo bem conhecido deste tipo de rede (Kosko, 1988).

As redes neurais artificiais podem ser utilizadas atualmente para implementar uma grande quantidade de funções e tarefas de processamento computacional e numérico.

### 1.4.2 Aprendizizado Supervisionado

Como exemplo de uma rede neural artificial capaz de aprendizado supervisionado, apresentaremos um dos algoritmos mais utilizados atualmente devido a sua capacidade de classificação, que é o da retropropagação, sendo este o algoritmo utilizado nesta pesquisa.

O algoritmo básico de aprendizado de retropropagação é empregado em uma rede neural artificial com uma topologia do tipo perceptron de três camadas, heteroassociativo, com conexões de alimentação direta (*feedforward*) da camada de entrada  $F_A$  para a camada intermediária  $F_B$  e desta para a camada de saída  $F_C$  (Figura 1.4).

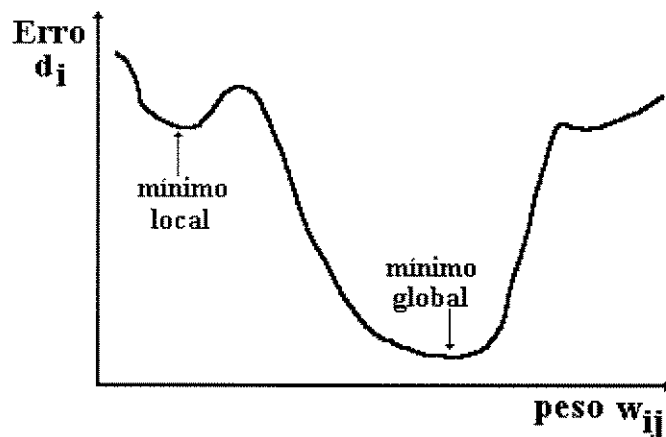


**Figura 1.4:** Arquitetura de uma rede neural multicamadas (perceptron)

O perceptron de três camadas é capaz de armazenar pares de padrões arbitrários  $(A_k, C_k)$ ,  $k = 1, 2, \dots, m$ , onde o par de padrão  $k$  é representado por dois vetores,  $A_k = (a_1^k \dots a_n^k)$  de entrada, e  $C_k = (c_1^k \dots c_q^k)$  de saída. A retropropagação aprende de forma supervisionada (através de exemplos), fora de linha (*off line*), em tempo discreto, e utiliza um método de descida de gradiente por correção de erro, ou seja, o algoritmo de codificação executa um mapeamento entrada-saída através da minimização de uma função de custo qualquer. Em um perceptron de duas camadas, se representarmos o erro em função de todas as combinações possíveis de pesos teremos uma figura denominada hiperparabolóide, que tem apenas um mínimo possível. Deste modo, o gradiente aponta sempre para baixo, em direção ao valor mínimo citado. A função de custo é minimizada realizando-se iterativamente ajustes nos pesos sinápticos de acordo com o erro, ou diferença, entre os valores desejados e obtidos na camada de saída  $F_C$ , ao se aplicar um par  $(A_k, C_k)$  na rede. A função de custo a ser minimizada normalmente é o erro quadrático (o quadrado da diferença entre os dois valores) acumulado para todos os padrões do conjunto de treinamento. Outras funções de custo podem ser utilizadas, mas independentemente disto, o procedimento de ajuste de pesos é realizado através do cálculo da mudança da função de custo com respeito à mudança em cada peso (método do delta). O processo de redução gradativa de erro que acompanha a minimização se denomina convergência. À medida que a rede aprende, o valor do erro converge para um valor estável, normalmente irreduzível. Para os perceptrons de duas camadas, pode ser provado matematicamente que a convergência é garantida. O mesmo não acontece para o perceptron de três camadas ou mais, que tem uma superfície muito complexa, que

pode ter mínimos locais, além do mínimo global (Figura 1.5) O processo de aprendizagem prossegue até que algum critério seja estabelecido, como por exemplo, um valor mínimo de erro global, ou uma diferença sucessiva mínima entre erros calculados para cada interação.

A rememoração é feita também pelo método de alimentação direta.



**Figura 1.5:** Superfície de erro – mínimos locais e globais

### ***1.4.3 Princípios do Aprendizado Supervisionado***

Antes de se utilizar o conjunto de treinamento, os valores de  $A_k$  e  $C_k$  devem ser todos normalizados em uma faixa adequada, por exemplo, convertidos para valores que fiquem entre 0 e 1, inclusive. Uma forma de realizar esta normalização é aplicar a seguinte fórmula para cada elemento  $A_k$  do vetor:

$$A_k = \frac{A_k - A_{\min}}{A_{\max} - A_{\min}}$$

onde  $A_{\min}$  é o valor mínimo encontrado ou determinado em  $A_k$  e  $A_{\max}$  é o valor máximo encontrado ou determinado em  $A_k$ . A mesma fórmula se aplicaria para o conjunto de saída  $C$ . Outras funções de normalização e transformação (por exemplo, logaritmo) podem ser utilizadas.

---

A normalização é importante, porque se um ou mais nodos tiverem valores de ativação muito maiores ou menores do que os demais, isto poderá produzir valores de erro desproporcionais, e levar a um desbalanço da rede, consequentemente prejudicando a sua convergência.

Os exemplos devem ser escolhidos de forma a representar todos os padrões existentes na possível população de padrões que se deseja classificar. De preferência, coloca-se vários exemplares de cada classificação de saída (por exemplo, vários conjuntos de sintomas associados a um determinado diagnóstico). Os exemplos podem ser sintéticos ou reais, sendo preferível utilizar-se os segundos, caso seja possível. Por exemplo, um conjunto de exemplos pode ser derivado aleatoriamente de um conjunto maior de pares de padrões entrada-saída, e utilizado para treinar a rede até a convergência. Em seguida, os pares de padrões restantes podem ser utilizados para testar se a rede é capaz de classificar corretamente casos conhecidos.

O algoritmo básico de treinamento da rede consiste em aplicar um padrão  $A_k$  na camada de entrada, e, em seguida, propagar os valores de ativação da camada de entrada até a camada de saída. Os valores de ativação assim obtidos na camada de saída são comparados com os valores contidos em  $C_k$  (desejados), e uma função de erro é computada para cada nodo de saída. Uma parte deste erro (chamado delta) é utilizada então para modificar os valores dos pesos sinápticos entre a camada intermediária e a camada de saída, de tal forma que, quanto maior o erro, maior será a

modificação do peso, sempre no sentido de diminuir o erro para cada par de nodos. Este método de correção por erros é aplicado em seguida aos pesos sinápticos entre a camada de entrada e a camada intermediária, levando, deste modo à uma propagação dos deltas no sentido inverso ao da ativação da rede (daí o nome de retropropagação).

Este procedimento é realizado para cada um dos pares de padrões utilizados para treinar a rede (chamado de conjunto de treinamento), voltando-se a repeti-lo para todo o conjunto diversas vezes, até que o erro global da rede convirja para algum valor mínimo. Esta passagem do algoritmo por todos os exemplos do conjunto de treinamento é chamado de ciclo de treinamento. Dependendo da complexidade e do tamanho dos padrões, o método de aprendizado por retropropagação pode exigir entre milhares e centenas de milhares de ciclos de treinamento. O número de fatores que influenciam a rede se torna muito grande, e as curvas de aprendizado podem variar amplamente entre si.

Na representação do algoritmo abaixo, utilizamos a notação convencionada por Simpson (1990).

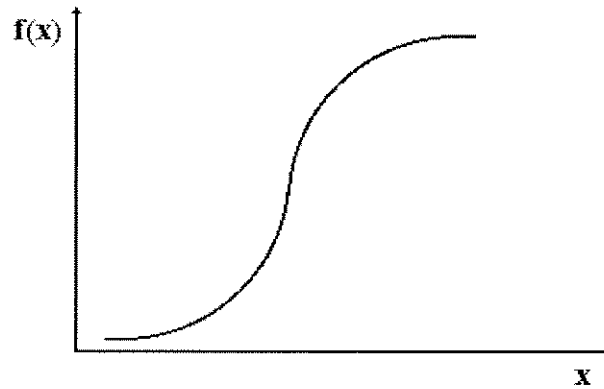
1. Atribua valores aleatórios entre -1 e +1 a todos os pesos sinápticos  $v_{hi}$ , das conexões entre as camadas  $F_A$  a  $F_B$ ; aos pesos sinápticos  $w_{ij}$  das conexões entre as camadas  $F_B$  a  $F_C$ ; a cada limiar  $\Theta_j$  da camada  $F_B$ ; e a cada  $\Gamma_j$  da camada  $F_C$  (vide Figura 1.4). Os valores limiares  $\Theta$  e  $\Gamma$  dos nodos das camadas de saída e intermediária, respectivamente, são normalmente constantes que regulam a sensibilidade de um nodo a variações em suas funções de ativação
2. Para cada par de padrões  $(A_k, C_k)$ ,  $k = 1, 2 \dots m$  existente no conjunto de treinamento (exemplos), realize o seguinte:



- a) Transfira os valores do vetor correspondente ao padrão de entrada  $A_k$ ,  $k = 1, 2, \dots, m$  aos nodos da camada de entrada  $F_A$  (um valor do vetor para cada um dos  $m$  nodos).
- b) Calcule os novos valores de ativação  $b_i$  dos nodos da camada intermediária  $F_B$ , através do seguinte método:

$$b_i = f\left(\sum_{h=1}^n (a_h v_{hi}) + \Theta_i\right)$$

para todos os nodos de  $F_B$ ,  $i = 1, 2, \dots, p$ , onde  $b_i$  é o valor de ativação do nodo  $i$  da camada intermediária  $F_B$ ,  $\Theta_i$  é o valor do limiar  $i$  do nodo da camada  $F_B$ , e  $f(\cdot)$  é a função de limiar sigmóide  $f(x) = (1 + e^{-x})^{-1}$  (Figura 1.6).



**Figura 1.6:** A função de Transferência Sigmoidal

- c) Calcule os novos valores de ativação  $c_j$  dos nodos da camada de saída  $FC$ , através da seguinte equação:

$$c_j = f\left(\sum_{i=1}^p (b_i w_{ij}) + \Gamma_j\right)$$

para todos os nodos de  $FC$ ,  $j = 1, 2 \dots q$ , onde  $c_j$  é o valor de ativação do nodo  $j$  da camada de saída  $FC$ ,  $\Gamma_j$  é o valor do limiar  $j$  do nodo da camada  $FC$ , e  $f(\cdot)$  é a função de limiar sigmóide.

- d) Calcule os valores de discrepância (erro)  $d_j$  entre os valores obtidos e desejados dos valores dos nodos da camada de saída  $FC$ , usando a seguinte equação:

$$d_j = c_j(1 - c_j)(c_j^k - c_j)$$

para todos os nodos de  $FC$ ,  $j = 1, 2 \dots q$ , e onde  $c_j^k$  é o valor do nodo correspondente ao elemento  $C_k$  do vetor de saída do par de treinamento  $k$ .

- e) Calcule os erros para cada nodo da camada intermediária  $FB$  correspondente a cada  $d_j$ , usando a equação:

$$e_i = b_i(1 - b_i) \sum_{j=1}^q w_{ij} d_j$$

para todos os nodos de  $FB$ ,  $i = 1, 2 \dots p$ , onde  $e_i$  é o erro computado para o nodo  $i$ .

- f) Ajuste os pesos das conexões entre as camadas de saída  $F_C$  e intermediária  $F_B$ , utilizando:

$$\Delta w_{ij} = \alpha b_i d_j$$

e

$$w_{ij} = w_{ij} + \Delta w_{ij}$$

para todos os nodos da camada intermediária  $F_B$ ,  $i = 1, 2 \dots p$  e todos os nodos da camada de saída  $F_C$ ,  $j = 1, 2 \dots q$ ; onde  $\Delta w_{ij}$  é o valor da mudança a ser realizada da conexão entre o nodo  $i$  da camada intermediária  $F_B$  e o nodo  $j$  da camada de saída  $F_C$  (chamado delta); e  $\alpha$  é uma constante positiva que controla a taxa de aprendizado (quanto maior  $\alpha$ , maior será a mudança  $\Delta$ , portanto mais acentuado será o gradiente e mais rápido o aprendizado, daí o significado deste parâmetro). Note, portanto, que o valor do delta para cada sinapse equivale ao produto do erro do nodo de saída pelo valor de ativação do nodo intermediário, pela taxa de aprendizado.

- g) Ajuste também os limiares dos nodos  $F_C$  usando o delta:

$$\Delta \Gamma_j = \alpha d_j$$

para todos os nodos da camada  $F_C$ ,  $j = 1, 2 \dots q$ .

- h) Ajuste agora os pesos sinápticos das conexões entre os nodos da camada de entrada  $F_A$  e da camada intermediária  $F_B$ , utilizando o delta:

$$\Delta v_{hi} = \beta a h_{ei}$$

e

$$v_{hi} = v_{hi} + \Delta v_{hi}$$

para todos os nodos da camada de entrada  $F_A$ ,  $h = 1, 2 \dots n$  e todos os nodos da camada intermediária  $F_B$ ,  $i = 1, 2 \dots p$ ; onde  $\Delta v_{ij}$  é o valor da mudança a ser realizada da conexão entre o nodo  $h$  da camada intermediária  $F_A$  e o nodo  $j$  da camada intermediária  $F_B$ ; e  $\beta$  é uma constante positiva semelhante à  $\alpha$ , que controla a taxa de aprendizado desta camada. Normalmente,  $\beta$  é feito igual a  $\alpha$ , mas o algoritmo permite que sejam diferentes, o que pode ser interessante em certos casos.

- i) Ajuste também os limiares dos nodos  $F_B$  usando o delta:

$$\Delta \Theta_i = \beta e_i$$

para todos os nodos da camada  $F_B$ ,  $i = 1, 2 \dots n$ .

3. Pegue um novo par de padrões  $(A_k, C_k)$  do conjunto de treinamento, e execute novamente o passo 2, repetindo-o até que os valores de correção de erro  $d_j$  para cada nodo da camada de saída  $F_C$ ,  $j = 1, 2 \dots p$ , para cada exemplo  $k = 1, 2 \dots m$  seja suficientemente baixo. Uma função global de erro para cada ciclo de treinamento pode ser computada com a seguinte equação:

$$D = \sum_{k=1}^n \sum_{j=1}^q d_j^k$$

que pode ser utilizada para fins de critério de parada do treinamento. Esta parada pode ser efetuada por intervenção do operador, através do acompanhamento visual dos valores globais  $D$  a cada 100 ou 1000 ciclos de treinamento (um gráfico de evolução do  $D$  em função do número de ciclos de treinamento); ou dos valores locais  $d_j$  de erro, através da exibição de um histograma indicando a distribuição dos valores de  $d_j$  a cada número fixo de ciclos. Alternativamente, a parada pode ser efetuada automaticamente, quando um dos critérios discutidos acima for atingido.

#### 1.4.4 Ativação

A chamada, ou rememoração de um padrão armazenado ocorre propagando-se um padrão de entrada  $A$  à camada de entrada  $F_A$  do perceptron assim treinado, até a camada de saída  $F_C$ , utilizando-se os conjuntos de pesos  $v_{hi}$  e  $w_{ij}$  estabilizados após convergência:

Da camada de entrada  $F_A$  para a camada intermediária  $F_B$ :

$$b_i = f\left(\sum_{k=1}^n (a_k v_{ki}) + \Theta_i\right)$$

para todos os nodos de  $F_B$ ,  $i = 1, 2 \dots p$ , onde  $b_i$  é o valor de ativação do nodo  $i$  da camada intermediária  $F_B$ .

E da camada intermediária  $F_B$  para a camada de saída  $F_C$ :

$$c_j = f\left(\sum_{i=1}^n (b_i w_{ij}) + \Gamma_j\right)$$

para todos os nodos de  $FC$ ,  $j = 1, 2 \dots q$ , onde  $c_j$  é o valor de ativação do nodo  $j$  da camada de saída  $FC$ .

### 1.4.5 Implementação

A codificação de um algoritmo de retropropagação em computador, utilizando-se uma linguagem de programação científica tipo BASIC, FORTRAN, PASCAL, C, etc., não oferece maiores dificuldades.

1. Com relação à forma de representação da rede (perceptron de 3 camadas), normalmente utiliza-se três vetores e duas matrizes. Cada vetor representa o valor de ativação dos nodos de uma camada (por exemplo, o elemento  $A_i$  do vetor  $A$  representa o valor de ativação do  $i$ -ésimo nodo de entrada). Quanto às matrizes, uma é usada para representar os pesos sinápticos entre os nodos da camada de entrada e os nodos da camada intermediária, e outra para representar os pesos sinápticos entre os nodos da camada intermediária e a camada de saída. Um dos elementos dessa matriz, digamos  $W_{ij}$ , corresponde ao peso da conexão entre o nodo  $i$  e o nodo  $j$  (matriz  $W$ ). Dois vetores adicionais são usados para representar os valores limiares  $\Theta$  e  $\Gamma$  para cada nodo das camadas intermediária e de saída.
2. Na configuração de uma nova rede, pede-se ao usuário que forneça diversos parâmetros, tais como número de nodos de entrada, número de nodos intermediários, número de nodos de saída, títulos para as camadas de entrada e de saída, títulos para cada nodo de entrada (por exemplo, nome do sintoma associado

àquele nodo), títulos para cada nodo de saída (por exemplo, nome do diagnóstico associado àquele nodo), etc. Normalmente, estes parâmetros da rede são armazenados em disco para posterior carregamento.

3. No treinamento da rede, o usuário deve fornecer ao programa parâmetros de treinamento, tais como número de exemplos de treinamento, valores da taxa de aprendizado e momento, número de ciclos de treinamento, valor de erro máximo desejado, etc. Deve fornecer também os pares de exemplos, que consistem de dois vetores: um contendo os valores a serem aplicados à camada de entrada (padrão), e outro contendo os valores a serem aplicados à camada de saída (classificação). Estes pares podem estar armazenados previamente em um arquivo de disco, e são carregados por nome, em dois vetores. A seguir, o programa deve realizar a inicialização da rede, que consiste em colocar valores aleatórios, pseudo-aleatórios, ou determinísticos, nas matrizes de pesos sinápticos. Os pesos devem variar entre -1 e +1, mas resultados melhores são obtidos algumas vezes, se estes pesos variarem entre uma fração destes, por exemplo, entre -0,3 e +0,3. Os vetores  $\Theta$  e  $\Gamma$  são também inicializados com um valor constante predeterminado. Finalmente o programa realiza o treinamento propriamente dito, através de duas alças de repetição: uma externa, que controla o número de ciclos através do conjunto de treinamento; e outra interna, que controla a apresentação sucessiva dos exemplos do conjunto de treinamento. Esta pode ser em ordem determinística ou aleatória. Comandos de desvio condicional podem ser colocado no interior da primeira alça para testar se o erro mínimo, ou outro critério qualquer de parada, foi alcançado, interrompendo o treinamento. Da mesma forma, segmentos de programa podem mostrar na tela ou na impressora a progresso do treinamento, como o erro global, a curva de aprendizado, etc. Após determinado número de

ciclos, ou ao se completar o treinamento, o programa pode armazenar o status da rede (matrizes e vetores) em disco novamente.

4. Finalmente, após ter sido obtida convergência do aprendizado, o programa pode ser utilizado para tarefas de classificação de padrões desconhecidos. Neste módulo, um vetor de entrada é fornecido pelo usuário, e a rede ativada para produzir um vetor de saída. Tanto a entrada quanto a saída podem ocorrer via teclado ou via arquivo de dados previamente construído pelo usuário.

#### **1.4.6 Discussão**

O algoritmo de retropropagação não garante que o erro global mínimo seja achado durante o processo de treinamento, mas apenas algum erro local mínimo (Figura 1.5). Este fato pode levar a oscilações do erro global da rede durante o treinamento, que fica aumentando e diminuindo, sem convergir. Pode acontecer, também, que o erro global para a qual a rede converge é inaceitavelmente alto (as classificações resultantes após treinamento serão erradas em um número muito grande de vezes. Em ambos os casos, diz-se que a rede é incapaz de aprender.

Não existe uma solução pré-determinada para resolver estes problemas. Portanto, quando uma rede for incapaz de aprender, deve-se realizar modificações em sua topologia ou parâmetros de aprendizado, de forma tentativa e experimental:

- a) Normalização: verificar se todas as variáveis de entrada estão normalizadas, ou se existe algum erro de digitação que torne o conjunto incoerente com a classificação que deve ser aprendida.
- b) Seleção de variáveis: verificar se existem variáveis irrelevantes para o problema de classificação. Isto pode ser feito através de testes estatísticos de associação e



correlação, como  $\chi^2$  (qui-quadrado), entre cada variável de entrada e a classificação de saída. Assim, podem ser incluídas apenas as variáveis que apresentam um alto grau de associação positiva ou negativa, e retiradas aquelas que não apresentam associação ou correlação significativas.

- c) Número e variabilidade dos exemplos: o aprendizado e/ou efetividade de classificação pode ser prejudicado se o número de exemplos for muito pequeno (não cobrindo todas as possibilidades de classificação), ou se os exemplares de uma mesma classificação forem muito variáveis entre si (confusão), ou se existirem exemplares muito similares entre diferentes classes (ambiguidade). Pode-se, nestes casos, utilizar critérios de eliminação de *outliers*, ou seja, exemplos que estejam provocando confusão ou ambiguidade.
- d) Camada intermediária: verificar se há um número ótimo de nodos na camada intermediária. Não existe uma fórmula para determinar esse número. O teorema da existência de Kolmogorov (1957) diz que a convergência é garantida se o número de nodos intermediários for igual ao número de nodos da camada de entrada (tamanho do padrão), mais um. Entretanto, essa convergência nem sempre acontece ou fica extremamente lenta. O número necessário de nodos na camada intermediária também depende de forma não predizível de outros fatores, como a variabilidade dos exemplos, a ocorrência de separações não-lineares complexas entre padrões, a relação entre o número de nodos de entrada e de saída, entre outros. Algumas pesquisas mostraram também que a efetividade de uma rede pode ser aumentada utilizando-se um procedimento de duplo treinamento: primeiro treina-se a rede a aprender até um valor (não ótimo), em seguida analisa-se a matriz de pesos e retira-se os nodos intermediários que contribuem pouco, ou seja, tem valores de pesos próximos de zero. Em seguida, retreina-se a rede usando este

conjunto diminuído de nodos na camada intermediária. Outra forma mais complexa seria “desconectar” algumas sinapses que contribuem pouco, ao invés de todo o nodo intermediário.

- e) Taxa de aprendizado: os fatores  $\alpha$  e  $\beta$  de aprendizado, usados para transferir parte do delta para a correção dos pesos, são determinantes em duas coisas: a velocidade do aprendizado (tanto maior quanto maiores forem estas taxas), e a ocorrência de mínimos locais. O uso de uma taxa elevada leva a um aprendizado rápido, mas se houver problemas na topologia da rede ou na estrutura dos padrões, poderão ocorrer oscilações do erro global. Nestes casos, é preferível utilizar uma taxa menor, levando a modificações menores dos pesos, com maior estabilidade na evolução da convergência. Alguns algoritmos modificados de retropropagação têm mudanças adaptativas da taxa de aprendizado, ou seja, utilizam taxas de aprendizado maiores no início, levando a quedas grandes de erro, e que são diminuídas de forma proporcional ao número de ciclos de treinamento, proporcionando maior estabilidade em fases avançadas do treinamento.
- f) Limiares e função de transferência: estes são parâmetros que também podem ter influência na capacidade e velocidade de aprendizado. Como mencionado anteriormente, os valores limiares  $\Theta$  e  $\Gamma$  dos nodos das camadas de saída e intermediária, respectivamente, são normalmente constantes que regulam a sensibilidade de um nodo a variações em suas funções de ativação. Embora sejam inicializados aleatoriamente no início do aprendizado, eles podem percorrer gamas mais baixas ou mais elevadas, conforme as características da rede neural empregada. Da mesma forma, é possível utilizar outras funções de transferência além da sigmóide, tais como logística, sigmóide segmentada, senóides ou cossenóides, cossenóides truncadas, etc. O importante é que seja uma função não

linear. A escolha da função também pode ter repercussões sobre a velocidade de aprendizado, principalmente pelo número de operações em ponto flutuante que exigem.

Outro problema da retropropagação é a lentidão do algoritmo, que normalmente exige muitos ciclos de treinamento para convergir. Observando-se a velocidade de convergência nos primeiros ciclos, e o início da tendência à estabilização, pode-se ter uma idéia se a rede está convergindo efetivamente ou não, e se os parâmetros escolhidos são adequados.

## **1.5 Redes Neurais: Aplicações na Medicina Intensiva**

### ***1.5.1 O computador nas Unidades de Terapia Intensiva***

A medicina intensiva exige o registro atualizado, acurado, dinâmico e integrado de um grande volume de informações sobre os pacientes internados, de modo que se possa dar assistência com a maior qualidade possível. Além disso, ocorrem constantemente na UTI complexos problemas de decisão, que vão desde o diagnóstico correto e ágil dos problemas que surgem com os sistemas orgânicos dos pacientes, até a escolha das condutas terapêuticas e de manutenção de vida mais adequadas. O prognóstico clínico acurado e o planejamento das atividades médicas e de enfermagem são também dois outros aspectos de grande relevância para a operação de rotina de uma UTI.

Por esses motivos, os computadores começaram a ser utilizados para a automação de diversas funções da UTI já há um certo tempo; e a abrangência e intensidade dessa aplicação vem crescendo muito ao longo dos anos, principalmente

após o aparecimento de instrumentação biomédica digital e a sua possibilidade de integração com os computadores (Kampmann *et al*, 1991; Clemmer e Gardner, 1992).

Em um sistema totalmente informatizado de UTI, os dados se originam de quatro fontes (Imhoff, 1992): (1) dos dispositivos eletrônicos de monitoração à beira do leito, (2) da rede local da UTI, (3) da rede geral de informação hospitalar, e (4) da entrada manual de dados. Os sistemas mais modernos implementados atualmente enfatizam os seguintes aspectos: (1) operação em tempo real e em linha, (2) maior independência possível da entrada manual de dados, (3) uso de interfaces gráficas de usuário para facilitar a captação e exibição de dados (Wenkebach, Pollwein e Finsterer, 1992), e (4) integração de dados de imagem e de sinais. Em consequência, as aplicações da Informática na UTI estão na vanguarda tecnológica em termos de recursos computacionais de alto desempenho necessários.

Mais recentemente, o surgimento do conceito da estação de trabalho médica (*medical workstation*), que procura integrar em um único terminal de vídeo todas as funções e informações necessárias para o apoio a assistência, teve também repercussão sobre sistemas orientados a esta filosofia na UTI (Higgins *et al*, 1991). Estas estações permitem o acesso, além do registro médico computadorizado, a imagens médicas (Gillespy, Staab e Lawrence, 1992) e sinais biológicos (Collura *et al.*, 1993). Podem dispor, também, de diversos sistemas de cálculo, análise, interpretação e decisão, que auxiliam o médico e a enfermeira em suas complexas atividades diárias dentro de uma UTI.

A utilização de tipos mais avançados de interface entre o computador e o usuário, tais como o uso de alarmes baseados em voz digitalizada (McIntyre e Nelson, 1989), também estão sendo explorados.

A experiência no uso de sistemas automatizados de gerenciamento de dados clínicos na UTI mostra inequivocamente que eles são mais bem aceitos pelo pessoal clínico, auxiliam a coleta de dados de forma mais confiável e apresentam menos erros no registro e na recuperação da informação, quando comparados com o sistema manual (Kari, Ruokonen e Takala, 1990). Um exemplo claro da utilidade dos computadores na UTI é a possibilidade de implementar programas para auxiliar a impressão automática de prescrições médicas, que incorporam mecanismos para aumentar a segurança dos pacientes, tais como a análise de interações medicamentosas, a análise de consistência de dosagens, de prescrições contraditórias ou potencialmente perigosas para uma determinada nosologia, etc. (Halpern, Thompson e Greenstein, 1992). A grande importância e prioridade que se dá atualmente nos países desenvolvidos à intensa utilização da Informática na UTI se refletiu no apoio a programas multinacionais de peso, tal como o projeto INFORM, na Comunidade Econômica Européia (Ambroso *et al.*, 1992).

Um dos aspectos mais notáveis da informatização das UTIs na última década consistiu no uso de técnicas de Inteligência Artificial para a implementação de sistemas de análise e monitoração ditos “inteligentes”, tais como sistemas baseados em conhecimento para interpretação automática de resultados de exames laboratoriais (Schwaiger, Haller e Finsterer, 1992); sistemas especialistas para a avaliação de dados hemodinâmicos (Van Den Heuvel *et al.*, 1990); de apoio ao diagnóstico e monitoração inteligente (Uckun, Dawant e Lindstrom, 1993), de apoio à ventilação artificial (Dojat *et al.*, 1993), e vários outros. As técnicas de Inteligência Artificial, combinadas com técnicas estatísticas e de análise dedutiva de bancos de dados são especialmente úteis para manejar a enorme massa de dados gerados pelos pacientes em uma UTI (Lau e Vincent, 1993), particularmente na área de prognóstico.

Nesta área, as tecnologias baseadas em sistemas conexionistas, especialmente as redes neurais artificiais, tem assumido um valor prognóstico significativo.

---

### ***1.5.2 Redes Neurais Artificiais na UTI***

Existem ainda poucos trabalhos publicados sobre o uso de redes neurais artificiais no prognóstico de pacientes sob cuidados intensivos. As medidas prognósticas mais comumente avaliadas nesses estudos são:

- a) mortalidade durante a internação
- b) severidade da doença
- c) duração da internação na UTI (estadia)
- d) predição de instituição ou suspensão de medidas terapêuticas

Com relação à predição de mortalidade na UTI, que é o mesmo tipo de prognóstico realizado por sistemas consagrados, baseados em modelos estatísticos lineares como o APACHE II, o trabalho relatado na presente dissertação foi o primeiro a ser publicado e divulgado em comunicação preliminar (Felipe *et al.*, 1992).

Posteriormente em outro trabalho sobre o mesmo tema, (Doig *et al.*, 1993) relataram o desenvolvimento de um modelo preditivo de mortalidade na UTI, utilizando redes neurais artificiais do tipo retropropagação, o qual foi comparado quanto ao desempenho com regressão logística multivariada (o método usado pelo APACHE). Os dados de entrada consistiram de 15 variáveis registradas no terceiro dia, provenientes de 422 pacientes consecutivos com tempo de estadia superior a 72 horas. Foram usadas as mesmas variáveis que as empregadas pelo método APACHE II (Knaus *et al.*, 1985). A base de dados completa foi usada para desenvolver os modelos de regressão e de redes neurais. Em seguida, a base de exemplos foi dividida

aleatoriamente em um conjunto de treinamento com 284 pacientes e um conjunto de teste com 138 pacientes. Usando a base de dados completa, a rede neural apresentou um desempenho claramente superior ao da regressão logística (sensibilidade de 1,000 e especificidade de 0,997, versus 0,525 e 0,966, respectivamente, com área sob a curva ROC de 0,9993 e 0,9259, respectivamente). Ambos os métodos, entretanto, tiveram desempenho semelhante no conjunto de teste (área sob a curva de 0,82).

Chen *et al.* (1992) utilizaram uma abordagem semelhante para prever quantitativamente o grau de severidade da doença de pacientes internados em UTI cirúrgica e unidade coronariana. Foi usada uma rede neural do tipo perceptron de três camadas, com algoritmo de treinamento por retropropagação. Os autores usaram 15 variáveis de entrada, as mesmas utilizadas para o cálculo do APACHE II, sendo: 12 variáveis fisiológicas (temperatura, pressão arterial média, frequência cardíaca, frequência respiratória, oxigenação, pH arterial, sódio, potássio e creatinina séricas, hematócrito, número de leucócitos e escore de coma de Glasgow), a idade do paciente, o seu estado fisiológico e a presença de problemas de saúde crônicos. Como variável de saída, foi usada uma escala de severidade de 0 a 10, estimada como a média dos escores atribuídos por dois médicos). Como conjunto de treinamento foram usados os dados de 35 pacientes, e como conjunto de teste, os dados de 10 pacientes. O coeficiente de correlação linear entre os valores estimados pelos médicos e os valores dados pela rede foi de 0,8, considerado como alto pelos autores.

Tu e Guerriere (1993) investigaram o uso de redes neurais artificiais para prever a duração da estadia na UTI de pacientes que passaram por cirurgia cardíaca. A predição acurada do tempo de estadia na UTI é importante, devido ao alto custo desse tipo de internação, e à baixa disponibilidade de leitos, o que coloca em risco os pacientes que necessitam ser internados, e ao fato de que quanto maior a estadia de um

paciente na UTI, maior é o risco de infecções, intercorrências e mortalidade. Segundo os autores, um instrumento preditivo do tempo de estadia poderia auxiliar uma melhor programação das admissões, altas e planejamento do pessoal. A rede neural foi treinada com uma base de dados com 713 pacientes e 15 variáveis de entrada. O parâmetro predito era a estratificação dos pacientes em três grupo de risco para estadia prolongada: baixa, intermediária e alta. Estadia prolongada era definida como sendo maior do que dois dias. A base de dados de teste consistiu de 696 pacientes. O desempenho da rede foi avaliado através da curva ROC, tendo atingido 0,7094 +/- 0,224 no conjunto de treinamento e 0,6960 +/- 0,0227 no conjunto de teste. Os autores concluíram pela utilidade das redes neurais na predição do tempo de estadia na UTI.

Ainda em relação à predição de tempo de internação, Buchman *et al.* (1994) realizaram uma comparação entre redes neurais artificiais e um modelo de regressão logística linear múltipla no prognóstico da cronicidade de pacientes internados na UTI. Foi utilizada uma base de dados de 191 pacientes com internação superior a 3 dias e que sobreviveram pelo menos durante mais quatro dias. Cronicidade foi definida como uma estadia superior a 7 dias. A comparação mostrou que as redes neurais foram capazes de predizer a cronicidade com maior acurácia do que o modelo logístico linear. Essa acurácia, entretanto, mostrou uma degradação em função do tempo de estadia (quanto maior o período de internação, pior era o desempenho da rede).

Outra área de predição onde as redes neurais podem ser usadas na UTI é relativa a decisões sobre instituição e retirada de medidas terapêuticas, com base nos indicadores fisiológicos e clínicos mostrados pelo paciente. Ashutosh *et al.* (1992) investigaram o uso de redes neurais, em comparação com análise discriminante linear, para elaborar um sistema semi-automático de apoio a decisão à suspensão do apoio



respiratório por ventilador na UTI, utilizando como base de dados 25 pacientes idosos dependentes desse apoio. Os dados de entrada consistiram de medidas de rotina de força inspiratória negativa, volume respiratório voluntário, etc., seguidas de um ensaio de suspensão do ventilador, cujo resultado (sucesso ou falha) era a saída esperada para a rede. O modelo desenvolvido foi testado com nove ensaios em seis pacientes. Os autores relataram acurácias entre 90 a 100 %, tanto para a rede neural quanto para o sistema de análise discriminante e concluíram que ambos podem ser úteis para a tarefa em questão.

## 1.6 Objetivos

Infelizmente, não é possível determinar através de regras quais condições tornam um paciente mais propício ao óbito, tão pouco esperar que uma equação matemática genérica consiga obter um resultado prognóstico significativo em todo o mundo, sem levar em consideração peculiaridades regionais e intercontinentais.

Entretanto, acreditamos que existam certos padrões clínicos, dentro de grupos populacionais, que possam ser determinantes para a predição da evolução do paciente. Neste sentido, chamamos atenção para as Redes Neurais Artificiais que são sistemas aptos, por excelência, ao reconhecimento de diversos tipos de padrões. Nosso objetivo é avaliar se as Redes Neurais Artificiais seriam capazes de abstrair um modelo preditivo para óbitos baseado na amostra populacional de uma Unidade de Terapia Intensiva. Desta forma, este modelo poderia ser aplicado em diversas localidades, respeitando as características locais a um custo reduzido. Para tal, pretende-se:

- definir as variáveis ideais para este propósito;

- definir uma melhor topologia de Rede Neural para a realização desta tarefa, procurando dimensionar um número ideal de camadas e elementos processadores por camada;
  - avaliar o desempenho das Redes Neurais Artificiais na predição de óbito e comparar seus resultados com o Índice APACHE II<sup>1</sup>.
- 

---

<sup>1</sup> Embora o Índice APACHE II não apresente um desempenho muito bom em países pobres e em desenvolvimento, esta comparação é feita em virtude da sua utilização em larga escala nas Unidades de Terapia Intensiva destes mesmos países. Portanto, por ser considerado um “*gold standard*” uma comparação dos resultados deste trabalho com os valores obtidos pelo APACHE II, para a mesma amostra, assume uma grande importância.

---

## ***Materiais e Métodos***

58

UNICAMP  
BIBLIOTECA CENTRAL  
SEÇÃO CIRCULANTE

## 2. Materiais

---

### 2.1 Hardware

Na preparação deste trabalho, utilizou-se um microcomputador padrão IBM PC-AT modelo 80486 DX, com velocidade de 33 MHz, 8 Mb de memória RAM, 200 Mb de capacidade do disco rígido, um monitor de vídeo colorido SVGA e uma placa de vídeo SVGA com 1 Mb de memória.

Este microcomputador executava o Sistema Operacional MS-DOS versão 6.2.

### 2.2 Software

Foram utilizados os programas NEURONET, AUDITOR, RANDOM1, RANDOM2, DBF2NET e EPI-INFO descritos a seguir:

- **NEURONET:** Este programa foi desenvolvido no Núcleo de Informática Biomédica da Universidade Estadual de Campinas (UNICAMP) (Sabbatini *et al.*, 1992), com o objetivo de facilitar a aplicação de redes neurais artificiais em Medicina. Para o seu desenvolvimento, foi utilizada a linguagem compilada Turbo BASIC<sup>®</sup>. O programa, atualmente em sua versão 2.09, é um *shell* genérico para implementação de três tipos de redes neurais organizadas em camadas múltiplas sendo operado através de uma interface interativa de usuário baseada em menus. Para a implementação das redes neurais do tipo perceptron de 3 camadas com aprendizado por retropropagação, utilizadas neste trabalho, o Neuronet implementa as equações detalhadas no Item 1.4. Uma descrição mais pormenorizada do Neuronet pode ser encontrada no Apêndice A, desta dissertação.
- **AUDITOR:** Programa desenvolvido em linguagem Clipper<sup>®</sup>, versão Summer'87 (Nantucket, Inc., EUA) com a finalidade de fazer uma crítica aos registros do Banco de Dados da UTI selecionando apenas aqueles que se encontram completos,

com valores não discrepantes e dentro dos critérios assumidos para a pesquisa. Estes critérios são discutidos no *Item 3 – Métodos*.

- **RANDOM1:** Este programa foi desenvolvido para produzir um grupo de casos para treinamento das redes neurais artificiais. Sua função é sortear, aleatoriamente, os casos existentes dentro de Banco de Dados da UTI, gerando um arquivo que poderá ser utilizado para o treinamento das redes neurais, após ser convertido pelo programa DBF2NET. O número de casos do arquivo gerado é determinado pelo usuário que também identifica, no início do programa, um número que será utilizado como semente do algoritmo de aleatorização. Além disso, o sistema converte os valores encontrados no banco de dados da UTI em valores escalonados dentro do intervalo [0,1], bem como decompõe variáveis discretas em diversas variáveis contínuas dentro, também, do intervalo [0,1]. Este programa foi desenvolvido, na linguagem Clipper<sup>®</sup>, versão Summer'87 (Nantucket, Inc., EUA).
- **RANDOM2:** Este programa , também desenvolvido na linguagem Clipper<sup>®</sup>, versão Summer'87 (Nantucket, Inc., EUA), gera um arquivo complemento para o teste das redes neurais, baseando-se naquele criado pelo RANDOM1. Desta forma, temos 2 arquivos distintos, sem casos redundantes, cuja somatória remonta o arquivo original da UTI. Como o RANDOM1, este programa, ao gerar o seu arquivo complemento, também converte os valores encontrados no banco de dados da UTI em valores escalonados dentro do intervalo [0,1] e decompõe variáveis discretas em diversas variáveis contínuas.
- **DBF2NET:** Este programa também foi desenvolvido no Núcleo de Informática Biomédica da Universidade Estadual de Campinas (UNICAMP), na linguagem Clipper<sup>®</sup>, versão Summer'87 (Nantucket, Inc., EUA), para efetuar automaticamente a conversão de arquivos em formato xBASE<sup>2</sup> para arquivos compatíveis com o NEURONET. Com esse programa, o usuário pode escolher os campos de um

---

<sup>2</sup> Padrão de Banco de Dados do tipo relacional compatível com o Banco de Dados dBASE<sup>®</sup> desenvolvido, inicialmente, pela Ashton-Tate, Inc., USA.

arquivo .DBF que irão constituir os nodos de entrada e saída da rede, bem como os demais parâmetros de definição das redes neurais, de forma totalmente interativa.

- **EPI-INFO 5.0:** Sistema utilizado para estatística epidemiológica e bioestatística, tendo sido desenvolvido em linguagem PASCAL, pelo *Center for Control Disease and Prevention*, Atlanta, USA.

## 2.3 Dados

Os dados usados para treinamento e teste das redes neurais foram coletados na UTI do Hospital das Clínicas da Universidade Estadual de Campinas<sup>3</sup>, englobando cerca de 3.000 pacientes internados de 1988 a 1992, com cerca de 120 variáveis registradas durante a permanência dos pacientes na unidade de internação. Estes registros estavam gravados em um Banco de Dados relacional compatível com o padrão xBASE.

---

<sup>3</sup> Os dados armazenados nos bancos de dados da Unidade de Terapia Intensiva do Hospital de Clínicas que foram utilizados dentro deste projeto, sofreram uma comparação prévia entre as informações existentes dentro dos arquivos eletrônicos e as informações constantes nos prontuários médicos da UTI dos 382 pacientes.

### 3. Métodos

---

Como citado anteriormente, os dados utilizados para esta pesquisa foram coletados na UTI do Hospital das Clínicas da Universidade Estadual de Campinas, englobando cerca de 3.000 pacientes internados no período de 1988 a 1992.

Os dados dos pacientes eram registrados diariamente, em um banco de dados relacional compatível com dBASE III+® (Ashton-Tate Corp., EUA), devido ao fato da Unidade de Terapia Intensiva do Hospital de Clínicas da Unicamp fazer parte do estudo multicêntrico do Sistema APACHE II. A coleta e digitação dos dados dentro do sistema da UTI era feita por uma enfermeira de nível superior, com conhecimentos em informática, cuja responsabilidade principal era zelar pela integridade dos dados. Entretanto, apesar das variáveis armazenadas refletirem as necessidades do sistema acima citado, por necessidade do próprio departamento, o banco de dados foi dimensionado de forma a abranger um volume maior de informações do paciente e de sua nosologia.

Para a identificação dos registros de pacientes que iríamos utilizar em nosso trabalho, assumimos algumas diretrizes com o objetivo de depurarmos o nosso universo de dados de forma a evitar problemas de convergência durante o treinamento das redes neurais e, conseqüentemente, sua tarefa de classificação. Os critérios adotados para a seleção dos registros foram os seguintes:

- utilização apenas de registros que possuíssem todas as variáveis preenchidas;

- utilização apenas de registros que não apresentassem variáveis com valores discrepantes;
- utilização de registros que incluíssem todos os tipos de patologias, com exceção de infarto agudo do miocárdio e insuficiência coronariana aguda (que também não são utilizadas para o APACHE II)<sup>4</sup>.

A seleção dos registros, baseada nas diretrizes acima citadas, foi realizada com o auxílio de um software de auditoria de dados especialmente desenvolvido (AUDITOR) que resultou uma casuística final de 382 pacientes. A ocorrência de óbito, ainda durante o período de internação do paciente (com duração não controlada), também era registrada no banco de dados. Embora os procedimentos terapêuticos e de manutenção de vitalidade tivessem sido registrados, esses dados não foram incluídos entre as variáveis preditoras de evolução.

### 3.1 Análise Estatística dos Dados

Em seguida, a matriz de dados foi analisada estatisticamente com o auxílio do programa EPI-INFO 5.0 (desenvolvido pelo *Centers for Disease Control and Prevention*, Atlanta, EUA), o qual foi utilizado para obter as seguintes análises:

- Frequências, médias e desvio-padrão para cada variável<sup>5</sup>;

---

<sup>4</sup> Visando uma posterior comparação com o APACHE II, resolvemos excluir também da nossa amostra os casos de pacientes internados em Unidades de Terapia Intensiva em decorrência de infarto agudo do miocárdio e insuficiência coronariana aguda. Estas patologias não foram incluídas durante o desenvolvimento do Sistema Apache devido ao fato de muitos hospitais possuírem UTI's próprias para doenças coronarianas. Portanto, para evitar qualquer interferência nos dados deste estudo multicêntrico, um dos parâmetros adotados foi a exclusão das patologias acima citadas.

<sup>5</sup> No caso de variáveis discretas, o valor a elas atribuído ao invés de representar uma graduação de severidade, representava uma forma de discriminação entre suas diversas formas de apresentação. Neste



- Medidas de  $\chi^2$  para associação entre cada variável e a mortalidade. Quando a variável era escalar, como idade, ela foi dividida em classes, e testada da mesma forma, usando o  $\chi^2$ .

Os resultados desta análise foram usados para derivar uma visão geral da associação dos dados e da estrutura causal, e para guiar a seleção das variáveis de entrada da rede. Foram selecionadas todas as variáveis associadas com a mortalidade (com uma probabilidade de associação devida ao acaso igual ou menor que 0,05), bem como algumas outras que, embora não associadas significativamente à mortalidade, poderiam ser fatores clínicos relevantes (tais como o sexo do paciente, pressão arterial média, etc.).

## 3.2 Variáveis Selecionadas

Seguindo os princípios acima expostos, obtivemos como resultado um total de 30 variáveis, apresentadas na Tabela 3.1 e explicadas a seguir:

- **Pressão Arterial Média** – Variável do tipo contínua. A Pressão Arterial Média (PAM) pode ser obtida, em mmHg, pela equação:

$$PAM (mmHg) = 0,3(2Pressão Diastólica + Pressão Sistólica)$$

---

caso, elas sofreram uma decomposição em diversas variáveis contínuas do tipo binário. Por exemplo, a variável Ritmo Cardíaco tinha seus valores variando entre 1 e 23, onde o valor “1” era atribuído para ritmo sinusal, o valor “2” para fibrilação atrial, o valor “3” para fibrilação ventricular, e assim por diante. Neste caso, esta variável foi decomposta em diversas outras variáveis binárias, tais como: ritmo cardíaco sinusal (sim/não), ritmo cardíaco fibrilação atrial (sim/não), etc.

**Tabela 3.1:** Variáveis selecionadas dentro do Banco de Dados da UTI utilizadas como Variáveis de Entrada para as Redes Neurais.

Variáveis	Tipo	Correlação	$\chi^2$	p
Pressão Arterial Média	C		1,67	0,4328
Insuficiência Renal	L	+	11,16	0,0008
Peritonite Purulenta	L	+	4,99	0,0254
Rítmo Cardíaco Sinusal	L		2,18	0,1401
Fibrilação Atrial	L	+	4,30	0,0380
Imunodepressão	L	+	3,95	0,0469
Casos Não-Cirúrgicos	L	-	10,38	0,0012
Casos Cirúrgicos Eletivos	L	-	25,72	0,0013
Casos Cirúrgicos de Urgência	L	+	3,90	0,0484
Idade	C		5,06	0,1671
Sexo	L		1,94	0,1631
Coma na Admissão	L	+	11,75	0,0006
Pós-cirúrgico Imediato	L	+	8,84	0,0029
Instabilidade Hemodinâmica	L	+	10,17	0,0014
Coeficiente Diagnóstico	C	+	79,56	0,0000
Temperatura Axilar	C*	+	17,40	0,0006
Frequência Cardíaca	C*	+	22,45	0,0000
Frequência Respiratória	C*		4,76	0,1904
FIO <sub>2</sub>	C*	+	9,21	0,0265
pH Sangüíneo	C*	+	16,75	0,0007
Sódio Sérico	C*	+	8,70	0,0335
Potássio Sérico	C*	+	11,87	0,0007
Creatinina Sérica	C*		5,21	0,0740
Hematócrito	C*		1,69	0,6403
Contagem Global de Leucócitos	C*	+	14,61	0,0021
Índice de Glasgow	C	+	42,27	0,0000
Coma	L	+	29,15	0,0000
Doença Neoplásica	L		1,09	0,2957
Infecção	L	+	11,75	0,0006
Número de Órgãos em Falência	C	+	26,42	0,0000

$\chi^2$  = Estatística de associação da variável com mortalidade

Correlação = (+) positivamente associada, (-) negativamente associada

Tipo = L – variável lógica, C – variável escalar

\* = pior medida obtida durante a estadia na UTI

p = probabilidade de associação aleatória

- **Insuficiência Renal** – Variável discreta (Presença/Ausência). Pode ser definida como sendo o resultado de uma redução abrupta da função renal (Insuficiência Renal Aguda) ou como consequência de um declínio progressivo da taxa de filtração glomerular (Insuficiência Renal Crônica).
- **Peritonite Purulenta** – Variável discreta (Presença/Ausência). O peritônio é uma membrana mesotelial contínua que reveste a cavidade abdominal e recobre as vísceras aí contidas. A peritonite purulenta consiste em uma infecção do peritônio, geralmente, devida a infecções bacterianas com uma consequente formação de pus. Muito comumente, os organismos entéricos penetram na cavidade peritoneal através de um defeito necrótico na parede do intestino, ou de outra víscera abdominal, defeito esse resultante de obstrução, infarto, neoplasia, corpo estranho ou doença inflamatória primária tal como a colite ulcerativa.
- **Rítmo Cardíaco Sinusal** – Variável discreta (Presença/Ausência). O coração é dotado de um sistema especial para gerar impulsos rítmicos que causam a contração do músculo cardíaco. O nodo Sino-Atrial (SA), pequena tira de músculo especializado com cerca de 3mm de largura e 1 cm de comprimento, é o responsável para gerar o impulso excitador rítmico normal. Este impulso é propagado pelas vias internodais até o nodo Átrio-Ventricular, onde o impulso é retardado antes de passar aos ventrículos. A seguir, o impulso desce pelos feixes Átrio-Ventriculares e fibras de Purkinje provocando, então, a contração da musculatura ventricular. Ao processo fisiológico de geração do impulso excitador e consequente contração da musculatura cardíaca, atribuí-se o nome de Rítmo Cardíaco Sinusal.
- **Fibrilação Atrial** – Variável discreta (Presença/Ausência). A fibrilação atrial constitui uma das formas mais frequentes das arritmias atriais.

Eletrocardiograficamente, distingue-se pela atividade atrial desorganizada, tendo uma frequência de 400 a 600 despolarizações por minuto. As ondas P, no traçado eletrocardiográfico, não são identificadas e a atividade atrial manifesta-se apenas por pequenas deflexões irregulares e desiguais.

- **Imunodepressão** – Variável discreta (Presença/Ausência). Reúnem-se sob o rótulo de deficiências imunitárias condições diversas, que possuem como denominador comum as seguintes características principais:
  - a) a reação imunitária é deprimida ou suprimida;
  - b) em geral, falta o caráter de estrita especificidade que distingue a tolerância imunitária;
  - c) são facilitadas as infecções (especialmente aquelas por vírus e eucariotas), a repetição e o agravamento destas, assim como as infecções por germes (procariotas, fungos), usualmente não patógenos (infecções oportunistas);
  - d) é facilitado o sucesso de enxertos, portanto são deprimidos os mecanismos de rejeição destes tecidos, embora as células imunocompetentes não as possam tolerar.
- **Casos Não-Cirúrgicos** – Variável discreta (Sim/Não). Pacientes admitidos no hospital para tratamento de nosologia de caráter clínico.
- **Casos Cirúrgicos Eletivos** – Variável discreta (Sim/Não). Pacientes admitidos no hospital para tratamento cirúrgico pré-programado.
- **Casos Cirúrgicos de Urgência** – Variável discreta (Sim/Não). Paciente admitidos no hospital para tratamento cirúrgico de emergência.
- **Idade** – Variável contínua. Idade do paciente quando da admissão hospitalar em anos completos (Tabela 1.1).

- **Sexo** – Variável discreta (Feminino/Masculino). Sexo do paciente.
- **Coma na Admissão** – Variável discreta (Presença/Ausência). Paciente apresentando-se em estado de coma quando da admissão na Unidade de Terapia Intensiva.
- **Pós-cirúrgico Imediato** – Variável discreta (Sim/Não). Paciente admitido na Unidade de Terapia Intensiva após tratamento cirúrgico, independente deste ter ocorrido em caráter eletivo ou de emergência.
- **Instabilidade Hemodinâmica** – Variável discreta (Presença/Ausência). A circulação sistêmica fisiológica é resultante do equilíbrio de quatro fatores principais, a saber: frequência cardíaca, complacência arterial, resistência vascular periférica e caráter da ejeção cardíaca. Alterações em um ou mais dos fatores citados podem produzir uma instabilidade hemodinâmica, ou seja, uma alteração no fluxo sanguíneo com repercussão sistêmica. Por exemplo: uma frequência cardíaca muito elevada impede o enchimento total do ventrículo antes que a sua sístole ocorra; conseqüentemente, a fração de sangue ejetado na aorta apresenta um volume e uma pressão diminuída. Tal situação determina o estabelecimento da síndrome de baixo débito cardíaco que resulta na diminuição da perfusão dos tecidos.
- **Coeficiente Diagnóstico** – Variável contínua. Valor proporcional atribuído à patologia constatada na admissão do paciente. Para as patologias que apresentam um menor risco de comprometimento à vida humana, são atribuídos valores numéricos baixos. O valor do coeficiente diagnóstico cresce à medida que o risco de vida aumenta para uma determinada patologia (Tabela 1.2 e Tabela 1.3).

- **Temperatura Axilar** – Variável contínua. A temperatura corporal média está em torno de 37 °C. Entretanto, durante um período de 24 horas, a temperatura varia de um ponto baixo no início da manhã aos níveis mais altos entre 16 e 18 horas. A amplitude desta variação diária, também chamada de ciclo circadiano da temperatura, é de cerca de 0,6 °C. Apesar da medida da temperatura na região axilar ser a habitualmente adotada no nosso meio, a medida que exprime com maior fidelidade a temperatura corporal é aquela tomada na região retal. Além disso, esta temperatura costuma apresentar um valor entre 0,2 e 0,3 °C acima da medida realizada na região axilar.
- **Frequência Cardíaca** – Variável contínua. Em um adulto jovem, o coração contrai, em média, 70 vezes por minuto durante o repouso. A frequência é diminuída (bradicardia) durante o sono e aumentada (taquicardia) por emoções, durante o exercício, a febre, e por muitos outros estímulos. Em indivíduos jovens saudáveis, a frequência varia com as fases respiratórias. Este efeito da respiração pode não se manifestar durante a respiração normal, mas aparece durante a respiração profunda.
- **Frequência Respiratória** – Variável contínua. Representada como sendo o número de movimentos respiratórios registrados dentro de um intervalo de 1 minuto. Com relação à frequência respiratória, podemos definir:
  - e) *Normal* - 15 a 21 movimentos por minuto, no adulto.
  - f) *Aumentada* – acima de 21 movimentos por minuto denominamos polipnéia ou taquipnéia. Determinadas alterações na amplitude e na frequência podem determinar estados patológicos específicos, por exemplo: respiração com amplitude aumentada e frequência não muito aumentada pode caracterizar uma acidose fixa ou metabólica.

g) *Diminuída* – bradipnéia, caracterizado por poucas excursões, abaixo de 15, e lentas.

Esta variável armazena a frequência respiratória do paciente independente da respiração ser espontânea, assistida ou controlada.

- **FIO<sub>2</sub>** – Variável contínua. Corresponde ao grau de oxigenação do sangue.
- **pH Sanguíneo** – Variável contínua. Os processos biológicos que ocorrem nas soluções aquosas do organismo, essenciais para mantê-lo vivo, são extremamente sensíveis à concentração de íons hidrogênio. Esta sensibilidade decorre do fato de que tanto os reagentes como os sistemas enzimáticos catalisadores modificam a sua constituição de acordo com a concentração de íons hidrogênio. Pequenas alterações dessa concentração, de seus valores normais, produzem acentuadas modificações na velocidade das reações químicas nas células, deprimindo ou acelerando suas funções. Como existe produção contínua de íons hidrogênio em decorrência dos processos metabólicos, torna-se necessário um mecanismo regulador que mantenha constante essa concentração, que pode ser expressa pelo pH dentro de limites estreitos, compatíveis com a vida. No homem, os limites situam-se entre 6,8 e 7,8. Em condições normais, encontram-se entre 7,35 e 7,45.
- **Sódio Sérico** – Variável contínua. O sódio é um importante elemento na manutenção da homeostase do organismo estando presente em uma série de processos sistêmicos e celulares, tais como mecanismo de funcionamento da Bomba Na<sup>+</sup>/K<sup>+</sup>, potencial de ação de células nervosas, sistema de condução cardíaco e outros. Variações do nível de sódio provocam reações orgânicas distintas quando do seu aumento (hipernatremia), a saber: sede, agitação, irritabilidade, letargia, abalos musculares, convulsões, etc.; ou de sua diminuição

(hiponatremia): apatia, anorexia, náuseas vômitos, cefaléia, hipotermia, etc. Os valores normais do sódio sérico são descritos na Tabela 3.2.

**Tabela 3.2:** Valores de referência para o sódio sérico

<b>Faixa Etária</b>	<b>mmol/l</b>
Recém-nascido	134 – 144
Lactente	139 – 146
Criança	138 – 145
A partir daí	136 – 146

- **Potássio Sérico** – Variável contínua. Como o íon sódio, o potássio tem uma importância muito grande na manutenção do equilíbrio interno do organismo estando presente em uma série de processos biológicos, muitos dos quais se relacionando intimamente com o sódio. Variações na concentração deste íon provocam também importantes reações sistêmicas. Quando do seu excesso (hipercalemia) observa-se o aparecimento de fraqueza, parestesias, arreflexia, paralisia, fibrilação ventricular, etc. Uma concentração baixa de potássio (hipocalemia) pode determinar o aparecimento de íleo-paralítico, alcalose, irritabilidade, náuseas, fraqueza, etc. Os valores normais do sódio sérico são descritos na Tabela 3.3.

**Tabela 3.3:** Valores de referência para o potássio sérico

<b>Faixa Etária</b>	<b>mmol/l</b>
Recém-nascido	3,7 – 5,0
Lactente	4,1 – 5,3
Criança	3,4 – 4,7
A partir daí	3,5 – 5,1



- **Creatinina Sérica<sup>6</sup>** - Variável contínua. A creatinina é um metabólito orgânico formado principalmente pela musculatura esquelética. Esta substância tem sua eliminação do organismo exclusivamente por via renal. Devido a esta característica, a avaliação do seu nível plasmático, ou na urina de 24 horas, funciona como um indicador da integridade renal. Os valores normais deste metabólito pode ser observado na Tabela 3.4.

**Tabela 3.4:** Valores de referência para a creatinina sérica

Faixa Etária		mg/dl
Cordão		0,6 – 1,2
Criança		0,3 – 0,7
Adulto:	M	0,6 – 1,2
	F	0,5 – 1,1

- **Hematócrito** – Variável contínua. Corresponde à porcentagem de papa de hemácias dentro de uma amostra de sangue total.

$$\text{Hematócrito} = 100(\text{Volume de hemácias}/\text{Volume de Sangue Total})$$

O valor normal do hematócrito varia de acordo com o sexo e a idade do paciente (Tabela 3.5).

---

<sup>6</sup> O Laboratório Central do Hospital de Clínicas da Unicamp não realiza exames de creatinina sérica em caráter de emergência, a menos que o valor da uréia sérica esteja fora da sua faixa de normalidade. Desta forma, quando o valor da uréia sérica era normal, atribuíam-se o valor 0 (dentro do normal) para a variável creatinina.

**Tabela 3.5:** Variação de referência dos valores do hematócrito

Idade		Faixa normal do Hematócrito
1 -3d		45 – 67
2m		28 – 42
6 – 12a		35 – 45
12 – 18a	H:	37 – 49
	M:	36 – 46
18 – 49a	H:	41 – 53
	M:	36 – 46

- **Contagem Global de Leucócitos** – Variável contínua. Corresponde à contagem de glóbulos brancos, independente do seu tipo, dentro de uma amostra de sangue total.

$$\text{Contagem Global de Leucócitos} = \text{Número de leucócitos}(1000 \text{ células/mm}^3)$$

O valor normal da Contagem Global de Leucócitos varia de acordo com a idade do paciente (Tabela 3.6).

**Tabela 3.6:** Variação de referência da leucometria

Faixa Etária	x 1000 células/mm <sup>3</sup>
Nascimento	9,0 – 30,0
24h	9,4 – 34,0
1m	5,0 – 19,5
1 – 3a	6,0 – 17,5
4 – 7a	5,5 – 15,5
8 – 13a	4,5 – 13,5
Adulto	4,5 – 11,0

- **Índice de Glasgow** – Variável contínua. Sistema baseado em pontuação utilizado para avaliar o nível da função cerebral. Ele inclui a avaliação da abertura ocular, resposta motora e resposta verbal a estímulos verbais e dolorosos (Tabela 3.7).

**Tabela 3.7:** Escala de Coma de Glasgow

A Escala de Glasgow corresponde à soma de cada parcela dos 3 grupos de sinais descritos abaixo. Desta forma, o melhor valor do Índice corresponde a 15 e o pior, ao valor 3.

Abertura dos olhos	Espontânea	4
	Ao som	3
	À dor	2
	Ausente	1
Resposta motora	Obedece comandos verbais	6
	Localiza a dor	5
	Flexão normal (retirada)	4
	Flexão anormal (decorticação)	3
	Extensão (decerebração)	2
	Ausente	1
Resposta verbal	Orientado	5
	Conversação confusa	4
	Palavras inadequadas	3
	Sons incompreensíveis	2
	Ausente	1

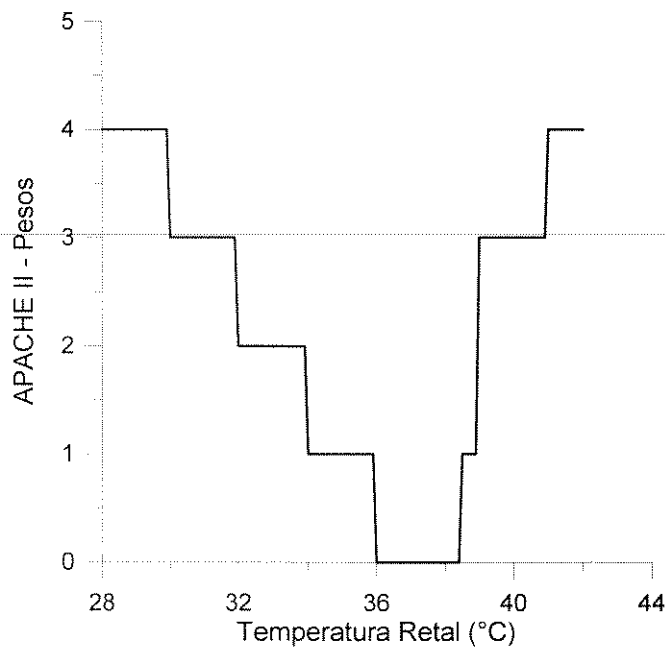
- **Coma** – Variável discreta (Presente/Ausente). Pacientes que entraram em “estado de coma” após a sua admissão na Unidade de Terapia Intensiva.
- **Doença Neoplásica** – Variável discreta (Presente/Ausente). Paciente portador de alguma neoplasia.
- **Infecção** – Variável discreta (Presente/Ausente). Denomina-se infecção à lesão tecidual decorrente de uma contaminação por microorganismos onde são encontrados os 4 sinais patognomônicos da inflamação (dor, calor, hiperemia e edema) acompanhados da formação de pus. Geralmente, a formação de pus continua até que toda a infecção seja suprimida.
- **Número de Órgãos em Falência** – Variável contínua. Número de órgãos que apresentavam acentuado grau de comprometimento funcional.

Os valores das variáveis fisiológicas e de laboratório registradas no banco de dados foram os piores (mais extremos) valores observados durante as primeiras 24 horas de internação. Nestes conjuntos de dados. As variáveis binárias foram codificadas como 0 (zero) para não/negativo/falso e 1 (um) para sim/positivo/verdadeiro. As variáveis escalares foram recalculadas de acordo com os critérios fisiopatológicos definidos por Gustafson *et al.* (1981) e usados por Knaus *et al.* (1985) no sistema APACHE II (*score APS*, ou *Acute Physiological Score*). Essa pontuação é um índice de severidade de doença monotonicamente ascendente no qual cada variável fisiológica é recalculada de acordo com os seus valores extremos considerados preditivos de severidade aumentada da doença. Por exemplo, a pior temperatura corporal registrada é codificada com a seguinte pontuação (Tabela 3.8):

**Tabela 3.8:** Sistema de Classificação de Severidade da Variação de uma Doença, para a variável biológica Temperatura Corporal.

Temperatura em graus C	Pontos
>> 41 ou << 30	4
30,9 a 40,9 ou 30 a 31,9	3
32 a 33,9	2
38,5 a 38,9 ou 34 a 35,9	1
36 a 38,4	0

Nota-se, portanto, que este recálculo atribui o número ordinal zero aos valores fisiológicos na faixa da normalidade, e escores tão maiores quanto mais distante da normalidade for o valor fisiológico, caracterizando dessa forma uma função em forma de U (Figura 3.1).



**Figura 3.1:** Pesos utilizados pelo APACHE II para a variação de temperatura retal.

As medidas de temperatura encontram-se distribuídas no eixo horizontal. O peso de cada ponto representa o grau de severidade da doença em relação à variação da temperatura retal.

As variáveis recalculadas dessa forma foram: temperatura corporal, pressão arterial média, frequência cardíaca, frequência respiratória, pressão parcial do oxigênio, pH arterial, potássio, sódio, creatinina, hematócrito e contagem de leucócitos. A variável idade foi também classificada em faixas (Tabela 1.1).

O coeficiente de diagnóstico é um valor proporcional à patologia constatada na admissão do paciente. O sistema APACHE II classifica 29 patologias cirúrgicas e 21 não cirúrgicas, utilizando valores escalares positivos ou negativos. Para os nossos experimentos, utilizamos a mesma notação que o APACHE II (Tabela 1.2 e Tabela 1.3)

As tabelas utilizadas para a recodificação das variáveis do APS, da doença crônica e do coeficiente de diagnóstico podem ser encontradas em Knaus *et al.* (1985).

### 3.3 Preparação dos arquivos de Treinamento e Teste

A seguir, os 382 casos foram divididos aleatoriamente em dois conjuntos de dados: um **conjunto de dados de treinamento** (282 casos) e um **conjunto de dados de teste** (100 casos), utilizando os programas RANDOM1 e RANDOM2. A determinação do número de registros de cada grupo baseou-se nos exemplos de diversos trabalhos da literatura internacional onde, quanto maior a variabilidade dos casos a serem classificados, maior deveria ser o número de casos no grupo de treinamento. Desta forma, devido à natureza da tarefa a que nos propusemos, determinamos que, aproximadamente,  $\frac{3}{4}$  da nossa amostra seria utilizada para a tarefa de treinamento das redes neurais e  $\frac{1}{4}$  dos casos seriam utilizados para o teste de classificação das redes já treinadas.

Posteriormente, cada um dos arquivos citados foi convertido em um arquivo de dados sequenciais do tipo ASCII, necessário para leitura direta pelos programas de simulação de redes neurais, e contendo apenas as variáveis selecionadas para entrada e saída. Além disso, para uso pela rede neural, todas as variáveis foram padronizadas no intervalo [0,1]. O valor máximo e o valor mínimo usados para padronizar cada variável foram obtidos pela análise univariada de dados de todo o conjunto de dados. As variáveis nominais, quando binárias (como sexo), foram alocadas a um único nodo e padronizadas.

### 3.4 Avaliação do Desempenho da Rede Neural

Como ocorre com qualquer teste diagnóstico novo desenvolvido em Medicina (Owens & Sox, 1990), é necessário avaliar de forma objetiva o poder preditivo das redes neurais artificiais, realizando-se diversas medidas de desempenho com relação à tarefa discriminativa a que se destina. Essa avaliação mais detalhada é fundamental, pois o parâmetro de desempenho mais comumente utilizado, que é a porcentagem de exemplos cuja predição foi realizada corretamente (parâmetro esse denominado **acurácia**), não é suficiente para descrever completamente como o método de decisão se comporta em relação aos falsos positivos (pacientes que tem uma evolução favorável, mas que o método de decisão aponta como desfavorável), e aos falsos negativos (o oposto do anterior).

Assim, devem ser realizadas sempre duas avaliações quantitativas do desempenho das redes neurais artificiais treinadas: uma utilizando o próprio conjunto de dados para treinamento original da rede, e o segundo, utilizando seu conjunto complementar de dados de teste, contendo dados de pacientes que não foram usados para o treinamento. Deste modo, é possível discriminar-se a capacidade de generalização da rede neural, ou seja, o quanto o modelo dos dados formado através do treinamento é capaz de realizar predições corretas com casos nunca vistos antes pela rede.

O método mais comum para avaliação de decisões com dois resultados possíveis (como sobrevida ou óbito) é ilustrado na Figura 3.2. Os resultados obtidos pelo método preditivo (no caso, a rede neural), são comparados com os resultados observados, e uma tabela de contingência de 2 x 2 é construída a partir das estatísticas de esperado/obtido:

- **Acurácia:** a proporção de predições corretas (soma de positivos verdadeiros e negativos verdadeiros dividido pela soma dos positivos verdadeiros, falsos positivos, falsos negativos e negativos verdadeiros);
- **Sensibilidade:** a proporção de positivos verdadeiros, ou seja, a medida da capacidade do método de decisão de prever óbito para aqueles que não sobreviveram na amostra de teste (número de positivos verdadeiros dividido pela soma de positivos verdadeiros e falso negativos) ;
- **Especificidade:** a proporção de negativos verdadeiros, ou seja, a medida da capacidade do método de decisão de prever sobrevida para aqueles que sobreviveram na amostra de teste (número de negativos verdadeiros dividido pela soma de negativos verdadeiros e falsos positivos);
- **Preditividade positiva:** a proporção de positivos verdadeiros em relação a todas as predições positivas (número de positivos verdadeiros dividido pela soma de positivos verdadeiros e falsos positivos); e
- **Preditividade negativa:** a proporção de negativos verdadeiros em relação a todas as predições verdadeiras (número de negativos verdadeiros dividido pela soma de falsos negativos e negativos verdadeiros).

Geralmente, para qualquer tarefa de decisão médica, a sensibilidade e a especificidade variam em direções opostas, ou seja, quando o teste é muito sensível a positivos, isto tende a gerar muitos falsos positivos e, conseqüentemente, a especificidade diminui. O oposto acontece quando o teste tem uma baixa sensibilidade: a especificidade aumenta devido ao grande número de falsos negativos. Em conseqüência, um método de decisão perfeito (100% de sensibilidade e



especificidade) raramente é alcançado, e um balanço entre ambas medidas de desempenho deve ser atingido.

		Evolução Observada		
		Morte	Sobrevida	
Predição	Morte	<b>A</b> Positivo Verdadeiro	<b>B</b> Falso Positivo	<i>Preditividade Positiva:</i> $A/(A+B)$
	Sobrevida	<b>C</b> Falso Negativo	<b>D</b> Negativo Verdadeiro	<i>Preditividade Negativa:</i> $D/(C+D)$
		<i>Sensibilidade:</i> $A/(A+C)$	<i>Especificidade:</i> $D/(D+B)$	<i>Acurácia:</i> $(A+D)/(A+B+C+D)$

**Figura 3.2:** Método de Cálculo dos índices de desempenho de um sistema prognóstico

Como as saídas da rede neural são contínuas por natureza, ou seja, elas produzem um valor de saída no intervalo contínuo [0,1] para cada dado registrado, é necessário definir-se um ponto de corte para contar o número de predições verdadeiras e falsas, visando construir a tabela de contigência 2 x 2. Como esta definição é essencialmente arbitrária, o melhor procedimento é estudar os efeitos de diversos pontos de corte sobre as estatísticas de desempenho (por exemplo, para valores de saída de 0,1, 0,2, 0,3, etc.) Para cada ponto de corte, a sensibilidade e a especificidade são calculadas e colocadas uma em cada eixo de um gráfico bidimensional, produzindo uma curva chamada ROC (*Receiver Operating Characteristic*) (Beck & Schultz, 1986) (Figura 4.7). Uma curva ROC perfeita corresponde a uma linha horizontal no topo do gráfico (0 % de especificidade e 100 % de sensibilidade), o qual percorre até o ponto de 100 % de especificidade e depois desce em direção ao eixo horizontal (100 % de especificidade e 0 % de sensibilidade). Curvas abaixo deste nível indicam métodos de decisão menos perfeitos, mas qualquer curva situada acima

da reta diagonal que atravessa o gráfico entre os pontos [0,0] e [100,100] pode ser considerada como boa (ela corresponde aos testes em que há igualdade entre sensibilidade e especificidade). A área sob a curva (AUC), obtida por integração, é outra medida normalmente usada para comparar diferentes curvas ROC entre si (Hanley & McNeil, 1982). No presente contexto, a AUC pode ser interpretada matematicamente como a probabilidade que, dados dois indivíduos, um que morreu, e outro que sobreviveu, a predição calculada será a maior possível para o indivíduo que morrerá.

Para identificarmos o melhor ponto de corte, dada uma curva ROC qualquer, calcularemos uma nova medida chamada **eficiência**, que é a média aritmética da sensibilidade e especificidade. O  $\chi^2$  calculado para a tabela de contingência 2 x 2 também pode ser utilizado como um indicador da significância estatística da melhor combinação de especificidade e sensibilidade ao longo da curva ROC. As tabelas 2 x 2 foram construídas para cada um dos seguintes pontos de corte de saída: 0,001, 0,01, 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 0,95, 0,99 e 0,999. A seguir, a curva ROC correspondente era obtida, e calculada a área sob a curva.

### 3.5 Avaliação da Contribuição das Variáveis

Uma das maiores críticas feitas com relação às redes neurais artificiais é a dificuldade de identificar a contribuição das variáveis de entrada à decisão, particularmente em redes de múltiplas camadas, onde o grande número de conexões e combinações não lineares de entrada oculta o papel das variáveis individuais de entrada. De fato, a topologia do perceptron determina que a saída (por exemplo, prognóstico estimado de sobrevivência) é o resultado de funções simultâneas de todas

as suas entradas. Em um perceptron de duas camadas, a saída pode ser interpretada como uma combinação linear de todas as entradas, ou seja um polinômio linear. Em contraste, não há nenhuma fórmula matemática de fácil interpretação, resultante da operação de um perceptron de três camadas (Huang & Lippman, 1987).

Para obtermos uma medida aproximada da contribuição de cada variável de entrada ao nodo de saída, utilizaremos o modelo sugerido Harrison *et al.* (1991) e modificado por Sabbatini (1994\*). Este modelo consiste em manter-se todas as entradas constantes (por exemplo, iguais a zero), com exceção de uma delas (que é igualada a 1, por exemplo), sendo que, a seguir, a rede treinada é propagada com este exemplo artificial e a saída é avaliada. Quanto maior for esta saída em relação à saída obtida com todas as entradas iguais a zero, maior será o impacto daquela variável de entrada em particular sobre a saída. Este parâmetro, que chamamos de **Magnitude Relativa de Saída**, pode ser negativo ou positivo, mostrando a direção da contribuição. Uma forma de documentar esses resultados é através de um gráfico, com um polígono ou histograma, mostrando as magnitudes relativas de saída para cada variável.

### 3.6 Construção das Redes Neurais

Após diversos experimentos com várias configurações de redes neurais, selecionamos os seguintes critérios para os experimentos principais:

- Perceptron de três camadas com regra de aprendizado por retropropagação;
- Variáveis de entrada: 30 variáveis, resultando em igual número de nodos de entrada (Tabela 3.6);
- Número de nodos da camada intermediária: 5, 10, 15 e 20;

- Um único nodo de saída (sobrevivência/óbito);
- Função de transferência: sigmoideal;
- Medida de erro: diferença absoluta;
- Limiarização interna (polarização) de 0,0;
- Taxa de aprendizado de 0,5 a 1,5 e momento de 0,0;
- Pesos iniciais: aleatorizados no intervalo [-0,5, +0,5];
- Método de apresentação de exemplos durante o treinamento: aleatório;
- Método de atualização dos pesos: contínuo.

Além da realização dos experimentos de acordo com os critérios acima adotados, foram realizados alguns experimentos variando alguns parâmetros, a saber:

- Método de apresentação de exemplos durante o treinamento: determinístico. Tal experimento pretende avaliar a importância da aleatorização da amostra durante os ciclos de treinamento na obtenção da convergência da rede;
- Variação da Taxa de Aprendizado. O objetivo deste experimento é avaliar se o treinamento das redes neurais sofre alguma influência significativa com taxas de aprendizado diferentes. Para tal, um mesmo experimento é repetido com as seguintes taxas de aprendizado: 0,5, 1,0, 1,5 e 2,0.

Para a realização destes experimentos, utilizamos o *software* Neuronet, desenvolvido pelo Núcleo de Informática Biomédica da Universidade Estadual de Campinas – Unicamp, descrito anteriormente.

---

## ***Resultados***

## 4. Resultados

---

Os resultados dos experimentos realizados com as redes neurais artificiais estão resumidos na Tabela 4.1 e nas figuras Figura 4.1 a Figura 4.3<sup>7</sup>. Para facilitar a identificação e referência às redes neurais elaboradas para este trabalho, doravante, utilizaremos a seguinte notação: (30-5-1), onde o primeiro número representa o número de nodos de entrada, o segundo número representa o número de nodos na camada intermediária da rede (*hidden layer*) e o último representa o número de nodos de saída. No exemplo dado, estamos nos referindo a rede com 30 nodos de entrada, 5 nodos na camada intermediária e 1 nodo de saída.

Todas as redes testadas convergiram dentro de 500 ciclos de treinamento. Uma série de experimentos anteriores realizados para avaliar a importância da aleatorização da amostra de entrada foram submetidas até 45.000 ciclos de treinamento, de modo a testar as propriedades de convergência. Estes experimentos mostraram que, mesmo com um número de ciclos de treinamento maior, o erro global final não foi muito menor que o registrado dentro dos 500 ciclos iniciais e o resultado global dos índices de avaliação prognóstica da rede mostraram-se até inferiores aos obtidos com as redes que tiveram um número de ciclos de treinamento menor. Um exemplo do exposto pode ser observado na comparação dos resultados obtidos na rede (30-5-1) com 500 e 45.000 ciclos de treinamento (Tabela 4.1).

---

<sup>7</sup> Os resultados completos de cada experimento podem ser encontrados no Apêndice B.

Na Figura 4.4 observamos as diferentes curvas de treinamento obtidas, para um mesmo experimento, alterando-se apenas os valores da taxa de aprendizado da rede ( $\alpha$ ). Neste caso específico, estamos avaliando os resultados da rede (30-5-1).

O erro global médio quadrado (RMS) foi menor que 0,025 em todos os casos.

A distribuição dos valores de saída das redes, em todos os casos, mostrou que as soluções convergiam para uma alta proximidade aos valores extremos, de 0 e 1, principalmente, no grupo de treinamento. A distribuição da saída da rede para o grupo de teste mostrou, em contraste com o grupo de treinamento, que 4 a 10% das predições caíam em faixas intermediárias: 0,3 a 0,8 (Figura 4.5 e Figura 4.6).

As estatísticas de desempenho variaram consideravelmente em função do ponto de corte na curva ROC, conforme evidenciado pelo teste do  $\chi^2$  para cada uma das tabelas de contingência 2 x 2 (Figura 4.7 e Apêndice B).

Todas as redes foram capazes de prever a evolução do paciente (óbito/sobrevida), com acurácias variando entre 97,9 % (redes 30-20-1 e 30-10-1) e 98,6% (rede 30-5-1), para o conjunto de dados de treinamento. Para o conjunto de dados de teste, a acurácia variou entre 76% (rede 30-15-1) e 81% (rede 30-5-1).

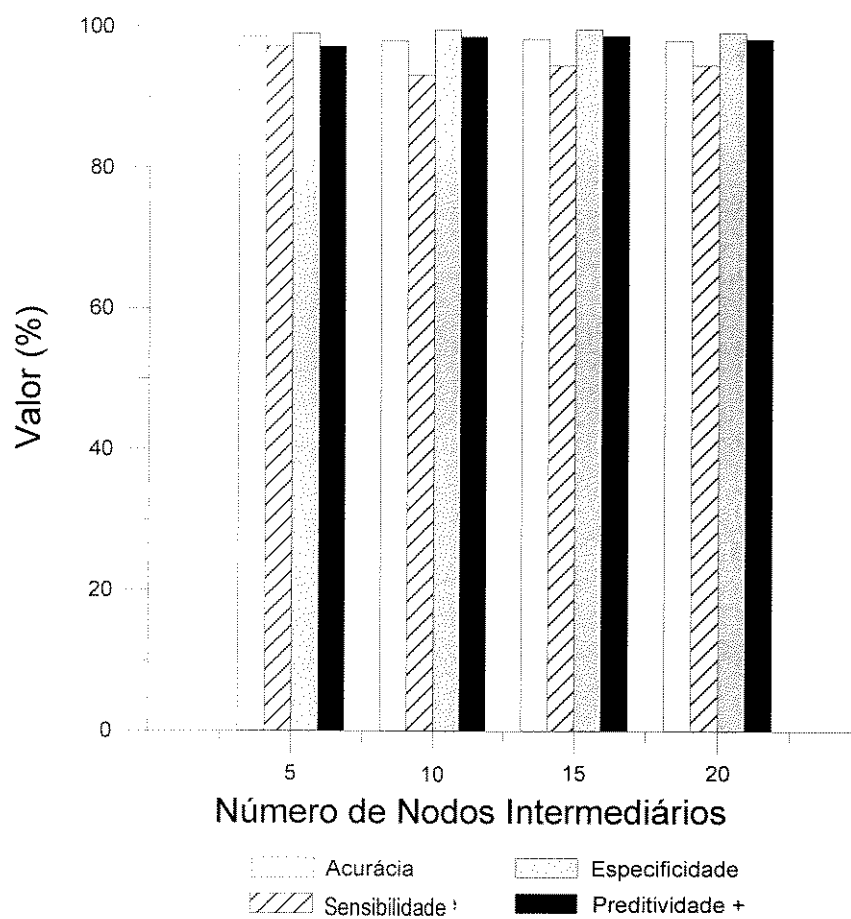
Submetendo as redes já treinadas a um teste de classificação dos 382 pacientes pertencentes ao universo de dados deste trabalho, obtivemos os resultados apresentados na Tabela 4.2.



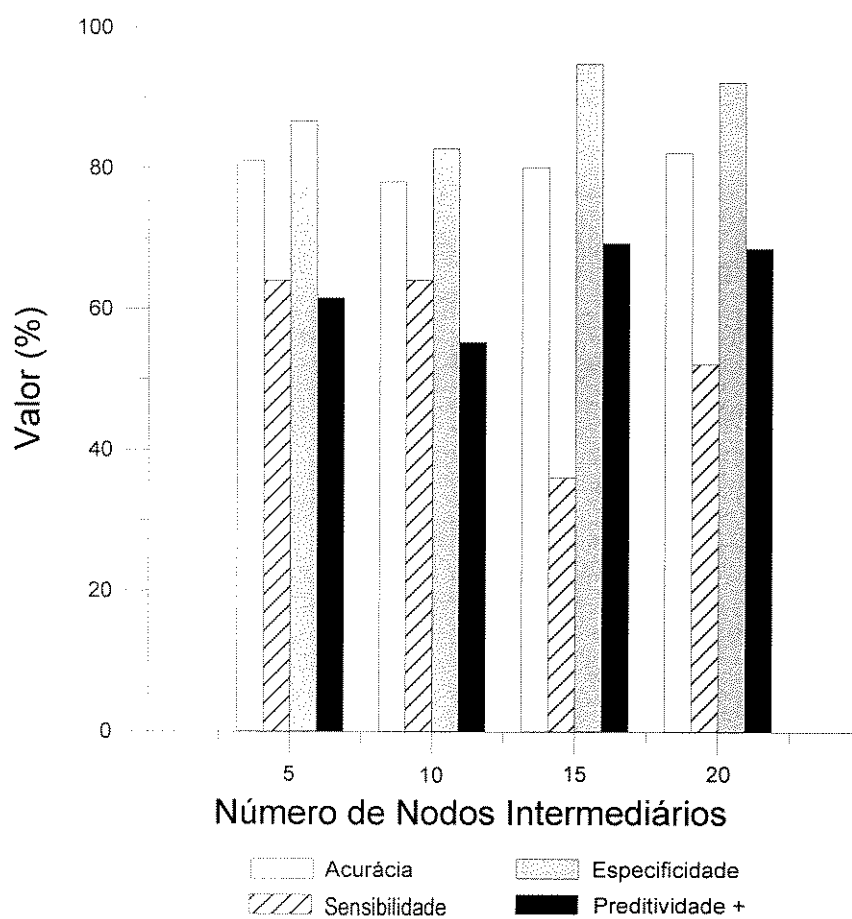


**Tabela 4.2:** Desempenho das Redes Neurais e do Sistema APACHE II na classificação do prognóstico de todos os registros dos 382 pacientes.

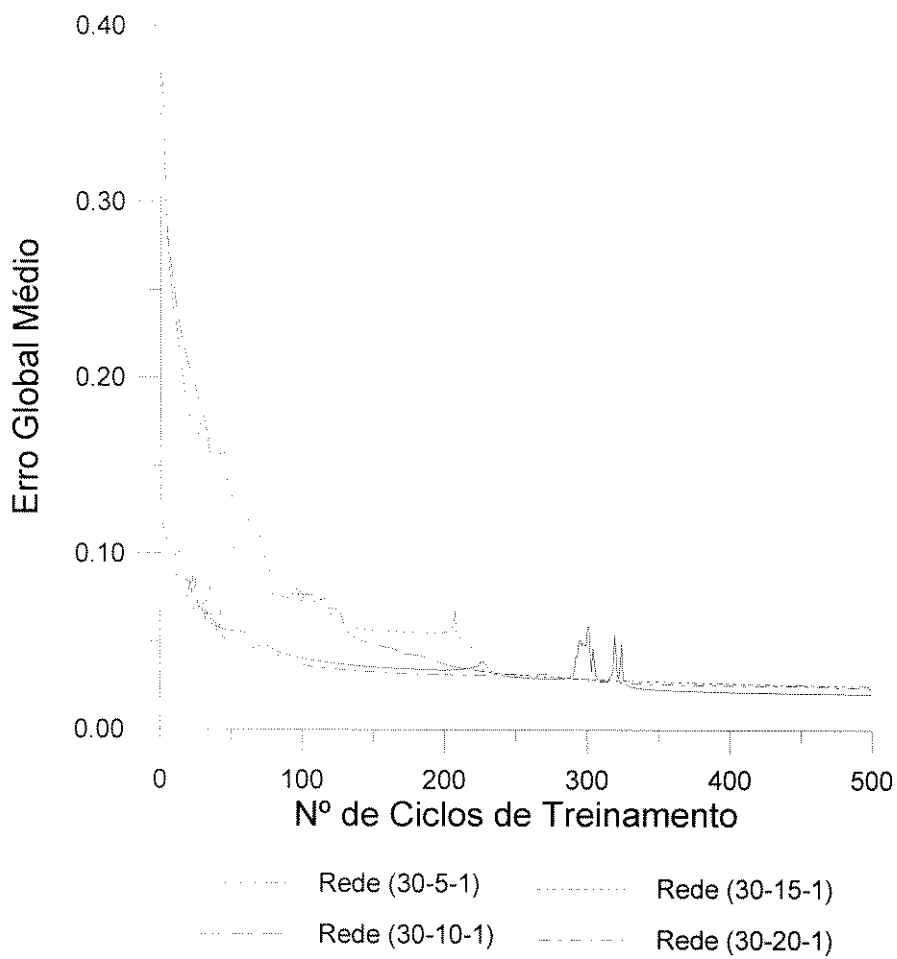
	(30-5-1)	(30-10-1)	(30-15-1)	(30-20-1)	APACHE II
Nº falsos positivos	4	72	66	9	28
Nº falsos negativos	2	7	6	3	52
Taxa de acerto	98,43 %	79,3 %	81,2 %	96,86 %	73,29 %
Sensibilidade	97,22 %	90,3 %	91,7 %	95,83 %	67,01 %
Especificidade	98,71 %	76,8 %	78,7 %	97,10 %	75,43 %
Preditividade +	94,59 %	47,4 %	50,0 %	88,46 %	51,85 %
AUC	0,9732	0,8405	0,8331	0,9673	0,7973



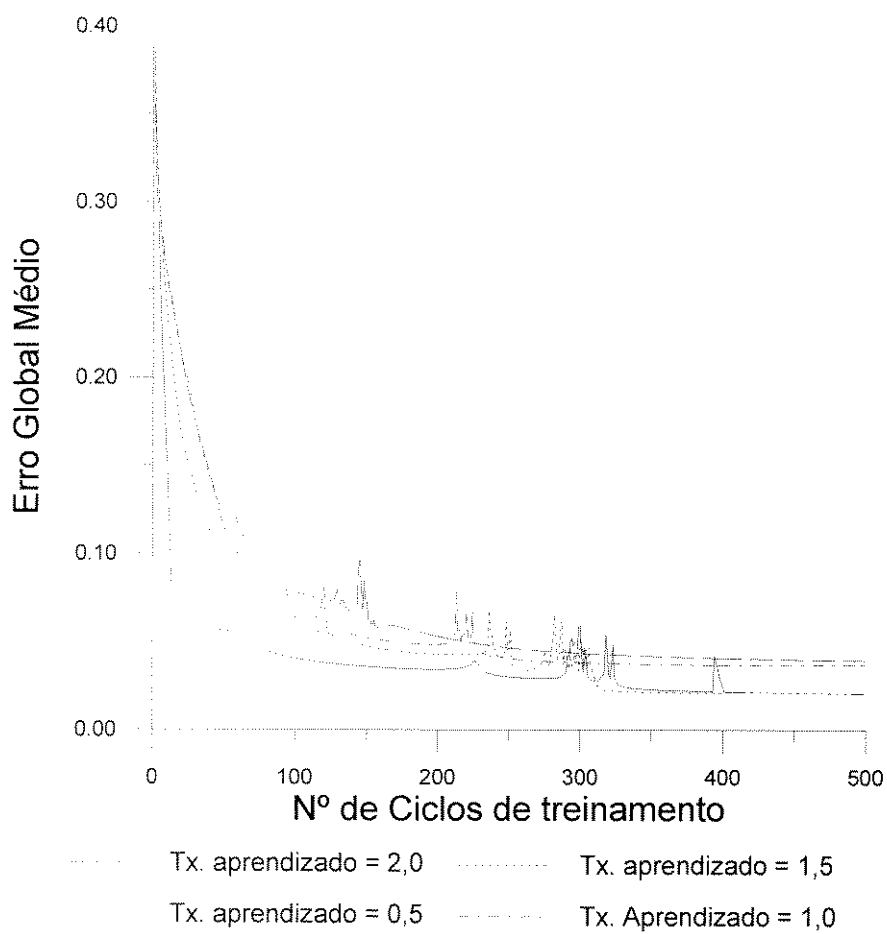
**Figura 4.1:** Desempenho das Redes Neurais (grupo de treinamento)



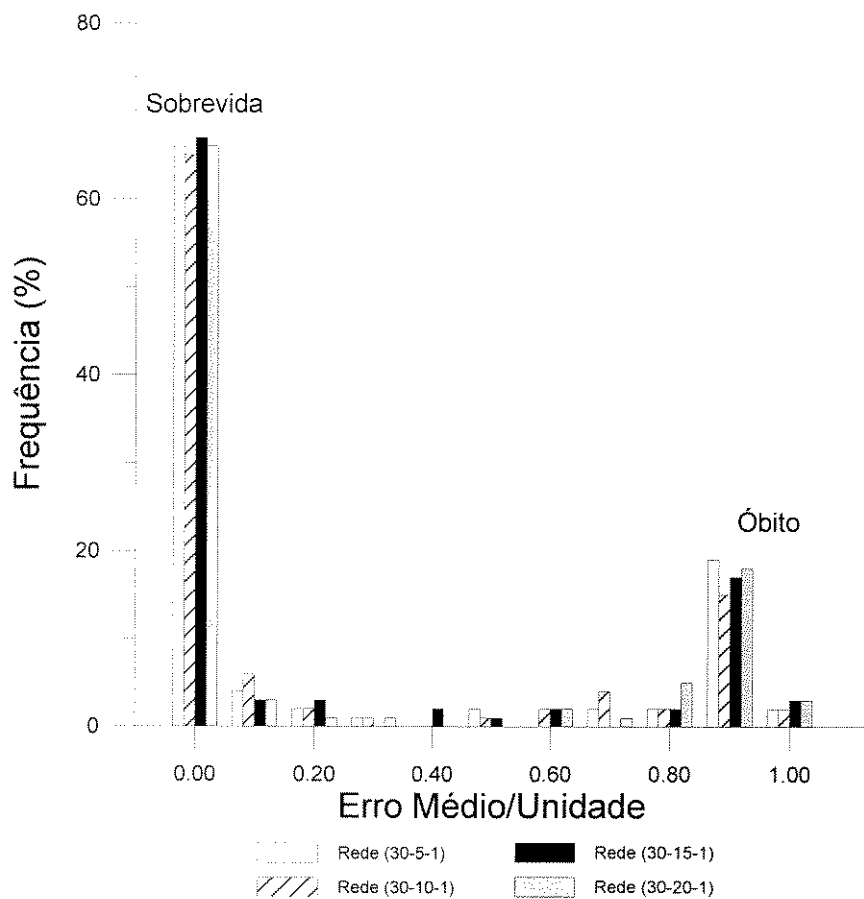
**Figura 4.2:** Desempenho das Redes Neurais (grupo de teste)



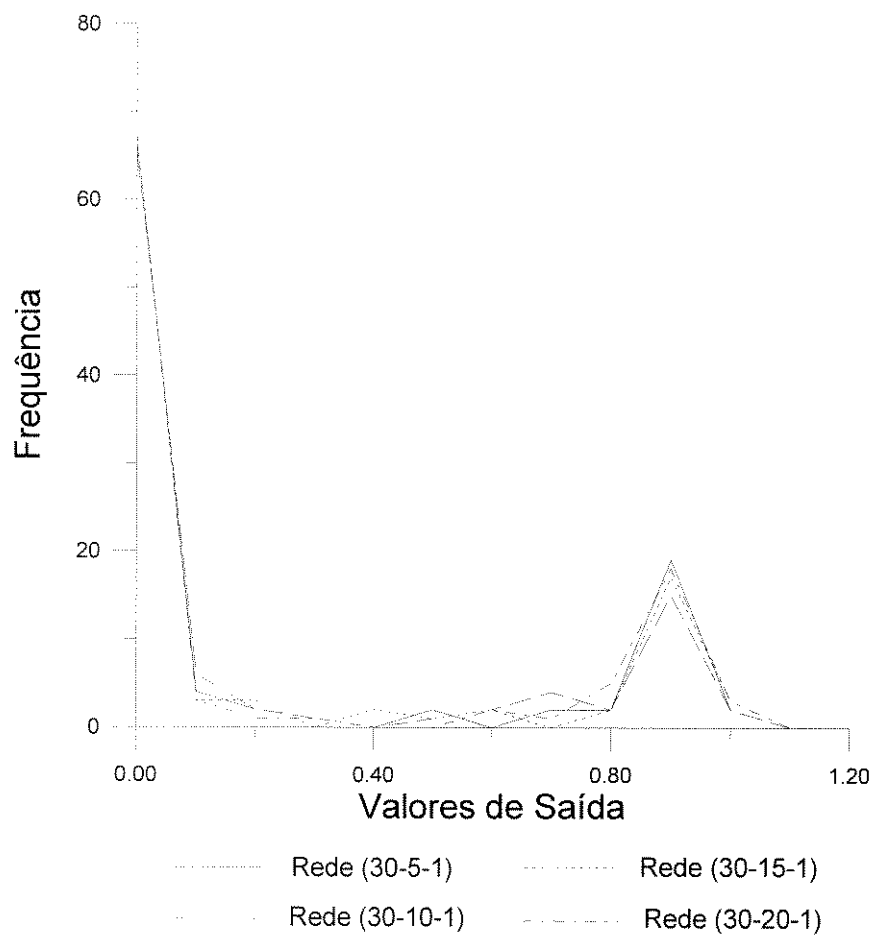
**Figura 4.3:** Curvas de Treinamento das Redes Neurais



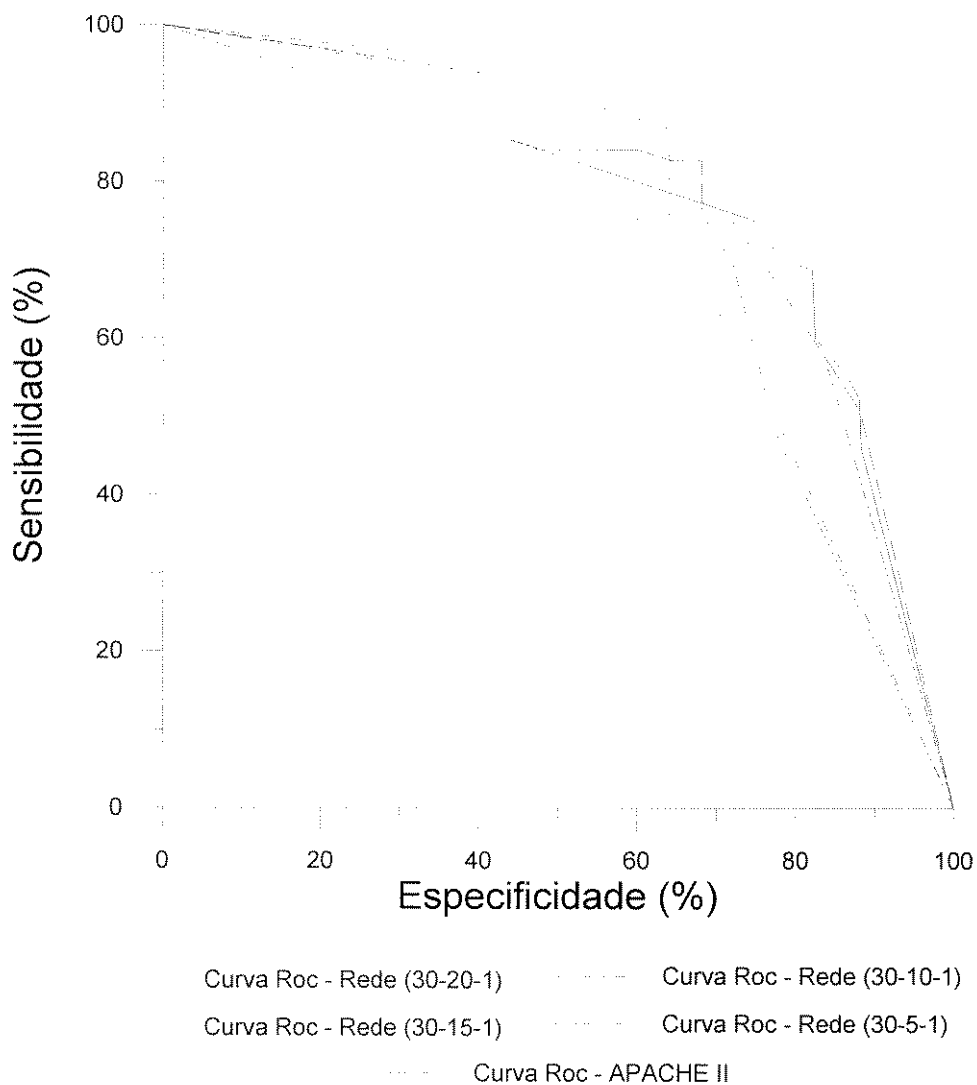
**Figura 4.4:** Curvas de aprendizado da Rede (30-5-1) para valores de Taxa de Aprendizado ( $\alpha$  -alfa) iguais a 0,5, 1,0, 1,5 e 2,0.



**Figura 4.5:** Distribuição dos valores de saída das redes neurais treinadas (conjunto de teste).



**Figura 4.6:** Distribuição dos valores de saída das redes neurais treinadas (conjunto de teste).



**Figura 4.7:** Curvas ROC das Redes Neurais e do Índice APACHE II (conjunto de teste).



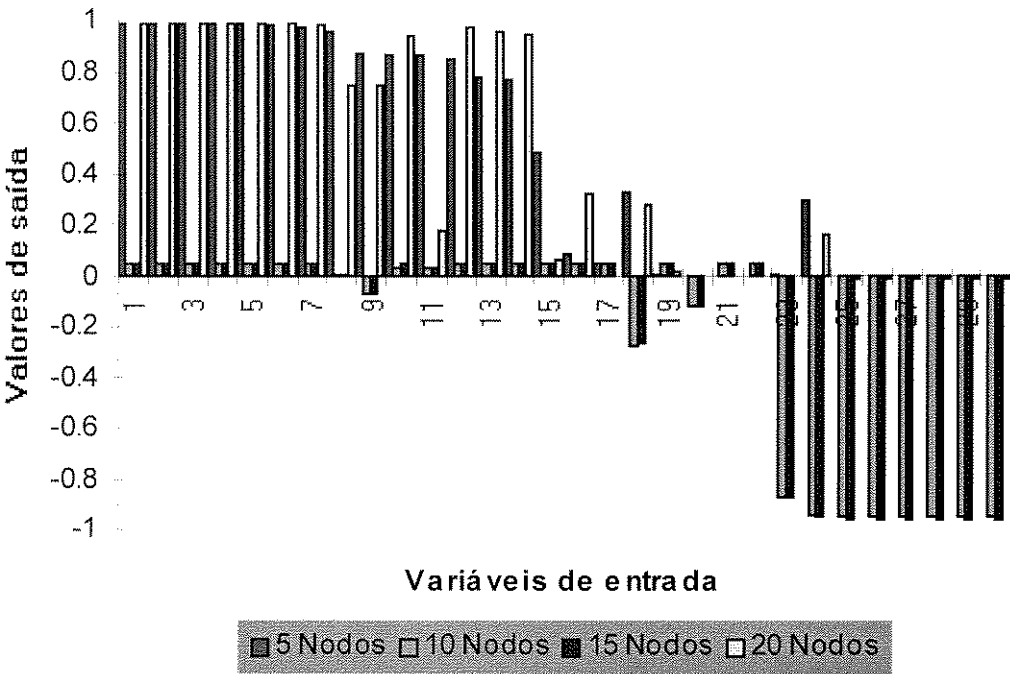
## 4.1 Contribuição individual das variáveis de entrada

Foi realizada uma análise comparativa da influência de cada variável de entrada em cada um dos experimentos, sendo seu resultado mostrado em um gráfico (Figura 4.8). Observa-se que as variáveis de entrada das redes (30-5-1) e (30-20-1) apresentam um comportamento com predominância de pesos de contribuição positiva enquanto que as mesmas possuem um caráter com predominância de pesos de contribuição negativa para as redes (30-10-1) e (30-15-1). Levando em consideração apenas o caráter positivo ou negativo das variáveis de entrada, independente da sua amplitude, constatamos que 22 das 30 variáveis apresentaram o mesmo comportamento para os quatro experimentos (Tabela 4.4), o que denota uma consistência das soluções obtidas. Nas 8 variáveis restantes, um comportamento com direções opostas de influência pode ser observado para as mesmas variáveis porém, somente em experimentos onde apenas se diferenciava o número de nodos na camada intermediária das redes treinadas (Tabela 4.5.).

Das 22 variáveis que apresentaram o mesmo comportamento dentro dos 4 experimentos, apenas 6 tiveram uma influência em direção negativa. A magnitude da média absoluta de saída dos quatro experimentos realizados, quando todas as variáveis de entrada têm seus valores igualados a zero, pode ser observada na Tabela 4.3.

Tabela 4.3: Magnitude da média absoluta de saída dos experimentos realizados

Rede	Média Absoluta de Saída
(30-5-1)	0,0087
(30-10-1)	0,9544
(30-15-1)	0,9531
(30-20-1)	0,0074



VARIÁVEIS DE ENTRADA					
1	Insuficiência Renal	11	Doença Neoplásica	21	Fibrilação Atrial
2	Coefficiente Diagnóstico	12	Potássio Sérico	22	Pressão Arterial Média
3	Temperatura Axilar	13	Frequência Respiratória	23	Hematócrito
4	Idade	14	Instab. Hemodinâmica	24	FIO <sub>2</sub>
5	Coma na Admissão	15	Cont. Global Leucócitos	25	Pós-cirúrgico Imediato
6	Número Órgãos em Falência	16	Coma	26	Índice de Glasgow
7	Sódio Sérico	17	Creatinina Sérica	27	Ritmo Cardíaco Sinusal
8	Imunodepressão	18	Peritonite Purulenta	28	Casos Cirúrgicos Urgência
9	Frequência Cardíaca	19	Infecção	29	Casos Cirúrgicos Eletivos
10	Sexo	20	pH Sangüíneo	30	Casos Não-Cirúrgicos

Figura 4.8: Contribuição individual das variáveis de entrada para os experimentos realizados.

**Tabela 4.4:** Variáveis que apresentaram o mesmo comportamento em todos os experimentos realizados.

VARIÁVEIS	5	10	15	20
Insuficiência Renal	+	+	+	+
Rítmo Cardíaco Sinusal	-	-	-	-
Imunodepressão	+	+	+	+
Casos Não-Cirúrgicos	-	-	-	-
Casos Cirúrgicos Eletivos	-	-	-	-
Casos Cirúrgicos de Urgência	-	-	-	-
Idade	+	+	+	+
Sexo	+	+	+	+
Coma na Admissão	+	+	+	+
Pós-cirúrgico Imediato	-	-	-	-
Instabilidade Hemodinâmica	+	+	+	+
Coeficiente Diagnóstico	+	+	+	+
Temperatura Axilar	+	+	+	+
Frequência Respiratória	+	+	+	+
Sódio Sérico	+	+	+	+
Potássio Sérico	+	+	+	+
Contagem Global de Leucócitos	+	+	+	+
Índice de Glasgow	-	-	-	-
Coma	+	+	+	+
Doença Neoplásica	+	+	+	+
Infecção	+	+	+	+
Número de Órgãos em Falência	+	+	+	+

**Tabela 4.5:** Variáveis que apresentaram comportamento antagônico em pelo menos um dos experimentos realizados.

VARIÁVEIS	5	10	15	20
Pressão Arterial Média	-	+	+	-
Peritonite Purulenta	+	-	-	+
Fibrilação Atrial	-	+	+	-
Frequência Cardíaca	+	-	-	+
FIO <sub>2</sub>	+	-	-	+
Creatinina Sérica	+	+	+	-
pH Sanguíneo	+	-	-	-
Hematócrito	+	-	-	-

---

## *Discussão*

## 5. Discussão

---

Pelo menos dois escores prognósticos quantitativos foram desenvolvidos para estimar a probabilidade de morte para pacientes internados em Unidades de Terapia Intensiva: o APACHE (Knaus *et al.*, 1985) e o MPM (Lemeshow *et al.*, 1985), ambos utilizando a metodologia da regressão logística múltipla para a construção de modelos lineares de um conjunto grande de variáveis colhidas durante a internação. Estes modelos de regressão logística geralmente possuem um desempenho muito bom na predição da mortalidade, relativa a uma determinada patologia, dentro do universo de casos de uma UTI. Porém, seu desempenho é clinicamente inferior quando são utilizados para a predição de mortalidade para pacientes de maneira individual ou em serviços cuja nosologia seja bem diferente da média (Rutledge *et al.*, 1991 e Ferraris *et al.*, 1992).

O primeiro modelo APACHE, desenvolvido por Knaus em 1981, envolvia 34 variáveis. Devido ao alto custo de colher número tão grande de variáveis, posteriormente, no APACHE II (Knaus *et al.*, 1985), elas foram reduzidas para 12 variáveis fisiológicas, mais idade, doença prévia, diagnóstico na admissão e tipo de admissão na UTI. O modelo logístico linear final foi baseado em um banco de dados com milhares de casos coletados em hospitais ao redor do mundo. O bom desempenho desses índices é muito bem documentado na literatura, e eles são utilizados em grande número de UTIs em todo o mundo, tendo a vantagem de facilitar a comparação de dados de diferentes lugares.

No presente trabalho, foi estudado, pela primeira vez na literatura o uso de uma tecnologia essencialmente não-linear, a rede neural artificial, para implementar sistemas de apoio à decisão clínica com o mesmo objetivo, ou seja, predição de mortalidade de pacientes críticos. Por isso, é imediatamente relevante uma discussão sobre o desempenho comparativo da mesma em relação aos métodos baseados em regressão logística, principalmente o APACHE.

É difícil afirmar qual das redes neurais teve melhor desempenho, baseado apenas nos critérios de sensibilidade e especificidade. Isto ocorre devido ao fato da sensibilidade variar em direção oposta da especificidade, ou seja, quando o teste é muito sensível a positivos, ele tende a gerar um número grande de falsos positivos. Consequentemente, a especificidade diminui, porque o número de falsos negativos cai. O oposto acontece quando o teste tem uma baixa sensibilidade: a especificidade aumenta devido ao grande número de falsos negativos. Em um contexto de UTI, é preferível ter falso positivos do que falso negativos (ou seja, exagerar no lado da prevenção é considerado eticamente mais aceitável), desta forma, a melhor rede seria a de 20 nodos, que tem 79% de acurácia no grupo de teste, 68% de sensibilidade, 82,7% de especificidade, 56,7% de preditividade positiva e uma curva ROC com uma AUC de 0,795. Isto significa que a rede errou apenas 6 em 282 casos no grupo de treinamento e 21 em 100 casos no grupo de teste. Entretanto, como o número excessivo de falsos positivos poderia aumentar os custos no tratamento intensivo, consideramos que rede com 5 neurônios intermediários poderia ser outra alternativa, uma vez que apresenta um índice de 81% de acurácia no grupo de teste, 64% de sensibilidade, 86,7% de especificidade, uma preditividade positiva de 61,5% e uma Curva ROC com uma AUC de 0,787. Porém, considerando os resultados obtidos no teste de classificação de todos os 382 pacientes dos grupos de treinamento e teste

(Tabela 4.2) para as redes com 20 e 5 nodos intermediários, podemos concluir que a melhor rede para a tarefa de classificação proposta foi a Rede (30-5-1). Este desempenho pode ser considerado muito bom, porque:

1. tanto o número de variáveis quanto a variabilidade eram muito grandes;
2. a associação com o prognóstico não era muito grande para a maioria das variáveis;
3. as intervenções terapêuticas não eram consideradas;
4. a casuística utilizada apresentava enorme variedade de etiologias no grupo de dados.

Os resultados apresentados mostram que as redes com grande número de elementos na camada intermediária não são, necessariamente, as melhores. A camada intermediária parece funcionar como uma banda de passagem no mapeamento das variáveis de entrada para o nodo de saída permitindo que a rede neural aprenda uma função matemática não linear capaz de separar, tão precisamente quanto possível, as duas classes de prognóstico sobrevida/óbito. Podemos dizer que uma rede “memorizou” quando aprendeu a representar um modelo particular de dados, e “generalizou” quando ela abstraiu um modelo genérico. Normalmente, quanto maior o número de nodos na camada intermediária maior é a chance da rede memorizar a amostra, ou seja, “decorar” os casos que lhe são apresentados. Quando isto acontece, durante a fase de teste onde uma série nova de dados são apresentados à rede, o desempenho classificatório obtido é ruim. Por outro lado, quanto menor o número de nodos na camada intermediária, maior a chance da rede neural aprender a generalizar, ou seja, ser capaz de “aprender” uma função matemática que lhe permita separar, especialmente, os grupos prognósticos. Porém, um número muito pequeno de nodos na camada intermediária impede tanto a memorização quanto a generalização. Não

existe um formalismo científico para obter uma rede com uma boa generalização para uma determinada tarefa. Até o momento, isto só é obtido após uma série de experimentações. Neste trabalho, especificamente, foram realizados 39 experimentos. Foram assim identificados alguns fatores importantes para a obtenção deste resultado:

- a) o tamanho da camada intermediária, já discutido anteriormente;
- b) o número e tipo de variáveis - muitas variáveis podem dificultar a rede em proceder uma separação da amostra; por outro lado, um volume pequeno de variáveis pode não ser suficiente para o estabelecimento de uma relação causal entre as mesmas;
- c) o número e a variabilidade dos exemplares de treinamento - quanto maior a variabilidade dos exemplos, maior deverá ser o número de casos utilizados para o treinamento da rede.

Pelo exposto, podemos concluir que a convergência da curva de aprendizado não representa definitivamente que uma rede aprendeu uma tarefa classificatória, ela apenas nos assinala que, pelo menos, a rede foi capaz de memorizar a amostra. Uma conclusão definitiva que uma rede aprendeu a generalizar só é possível após uma avaliação dos resultados de classificação do grupo de teste.

Um aspecto interessante deste trabalho é que, normalmente, os valores obtidos na saída das redes neurais no grupo de teste é do tipo contínuo, ou seja, ocorre uma dispersão dos resultados ao longo do intervalo determinado para os resultados de classificação. Neste caso, observamos que apenas 7 a 10% dos resultados obtidos nas saídas das redes, durante a fase de classificação do grupo de teste, apresentaram valores intermediários entre os extremos do intervalo  $[0,1]$ . Tal comportamento demonstra que as redes foram capazes de criar um modelo discriminatório para um prognóstico binário óbito/sobrevida. Além disso, o formato em U das curvas obtidas



(Figura 4.6) demonstra que a função implementada pelas redes para executar esta tarefa classificatória é de natureza não linear.

Analisando-se comparativamente as curvas de aprendizado para a Rede (30-5-1) obtidas pelo treinamento da mesma com diferentes valores de taxas de aprendizado, pode-se comprovar que as maiores taxas (principalmente aquelas acima de 1), produzem um comportamento instável na rede, caracterizado por bruscas alterações em certos pontos da curva. A mudança na taxa de aprendizado da rede não afeta significativamente a capacidade de se atingir a convergência, mas apenas a rapidez com que ela era conseguida (Figura 4.4).

Outro aspecto interessante observado no treinamento das redes neurais foi a obtenção de um desempenho muito melhor em termos de rapidez de aprendizado, bem como nos resultados de classificação, para a rede que teve a apresentação dos casos na camada de entrada, realizado de forma aleatória. Embora a rede, cuja amostra de entrada foi sempre apresentada na mesma sequência (apresentação determinística), tenha sido treinada até 45.000 ciclos e tenha apresentado um erro final menor, seu desempenho global foi inferior. Provavelmente, a apresentação determinística dos casos, durante o treinamento da rede, não impede que a mesma fique “viciada” com a amostra e, desta forma, não consiga escapar de algum erro local mínimo que tenha encontrado durante a fase de aprendizado. O algoritmo de retropropagação não garante que o erro global mínimo seja achado durante o processo de treinamento, mas apenas algum erro local mínimo.

Avaliando a contribuição individual das variáveis de entrada, constatamos que 22 das 30 variáveis de entrada apresentaram o mesmo comportamento para os quatro experimentos onde alteramos o número de nodos da camada intermediária das redes neurais. Dentro destas 22 variáveis, apenas 6 apresentaram uma influência inversa.

Avaliando estas variáveis, verificamos que quatro se referiam ao tipo de internação/admissão do paciente na UTI, ou seja, internação clínica, cirurgia de emergência, cirurgia eletiva ou pós-cirúrgico imediato. Quanto às duas restantes, uma se referia ao Ritmo Cardíaco Sinusal e, a outra, ao Índice de Glasgow.

Como as quatro variáveis que identificam o tipo de internação/admissão do paciente na UTI possuem efeito inverso, independente do prognóstico do paciente, provavelmente elas não estão desempenhando um papel discriminante que influencie no resultado prognóstico da rede, portanto, a sua exclusão do grupo de variáveis de entrada deve ser considerado em novos experimentos. O comportamento inverso da variável Ritmo Cardíaco Sinusal se justifica pela própria característica fisiológica desta variável, ou seja, um ritmo cardíaco normal tem um efeito prognóstico positivo. Chama a nossa atenção o fato da variável Índice de Glasgow apresentar uma correlação altamente inversa. Isto decorre do fato desta ser uma variável de forte valor preditivo com relação a evolução do paciente, sendo muito importante para o cálculo do Índice APACHE. Este comportamento deve-se ao fato de que quanto maior for o Índice de Glasgow maior é o nível de consciência do paciente, ou seja, melhor o prognóstico (Tabela 3.6). Quando esta variável foi mapeada para o intervalo [1,0] os piores índices ficaram com os valores mais próximos a zero. Porém, como o método utilizado para averiguar a influência de cada variável consistia em atribuir o valor 0 para todas as variáveis, com a exceção da que estaria sendo estudada, que receberia o valor 1, apresentamos para as redes sempre um ótimo valor clínico de Glasgow. Consequentemente, as redes fizeram uma associação entre um alto valor atribuído a esta variável e sua influência inversa sobre a evolução a óbito.

Dentre as variáveis que apresentaram uma influência direta, chamamos atenção para as variáveis Sexo e Coeficiente Diagnóstico. O sexo masculino (variável

com valor igual a 1) apresentou uma alta correlação com o óbito. Avaliando-se o universo dos 382 casos utilizados para o treinamento e teste das redes neurais observamos que 266 (69,7 %) eram homens e destes, 73 pacientes faleceram, o que corresponde a 67,1% de todos os óbitos do universo estudado. Por sua vez, a variável Coeficiente Diagnóstico corresponde a um valor de severidade relacionado com a patologia registrada na admissão do paciente. Quanto maior o valor atribuído a esta variável, maior a severidade da patologia. Consequentemente, pior o prognóstico.

Observando as 8 variáveis que apresentaram comportamentos não padronizados, constatamos que as influências diretas e inversas possuíam 2 tipos de padrões:

1. As variáveis apresentavam influência direta e inversa de baixa amplitude, variáveis Pressão Arterial Média, Fibrilação Atrial, pH Sanguíneo e Creatinina Sérica;
2. As variáveis apresentavam influência direta e inversa de grande amplitude, variáveis Peritonite Purulenta, Frequência Cardíaca, FIO<sub>2</sub> e Hematócrito;

No caso do padrão de baixa amplitude, provavelmente, para as redes neurais, estas variáveis não contribuíam de forma clara para a definição de um prognóstico, sendo a sua direção irrelevante ou aleatória. Já para o caso do padrão de grande amplitude, observamos que as redes de 5 e 20 possuem uma influência direta, enquanto que as redes 10 e 15 possuem uma influência inversa. Provavelmente, para um mesmo valor de acurácia existiriam várias soluções possíveis determinando, assim, arranjos diferentes para estas variáveis. Isto é um sinal que estas variáveis não contribuem de forma consistente para a tarefa de discriminação sendo, provavelmente, utilizadas pelas redes para ajustarem seus “modelos” de predição prognóstica.

Os resultados relativos ao valor preditivo do modelo de rede neural artificial desenvolvido no presente trabalho para mortalidade de pacientes críticos em UTI são comparáveis aos publicados na literatura. No estudo mais extenso (e mais semelhante ao nosso), Doig *et al.* (1993), usando apenas 15 variáveis, obtiveram áreas sob a curva ROC de 0,9993 e 0,82 para os conjuntos total e de teste respectivamente. No presente trabalho, a melhor rede (30-5-1) obteve áreas sob a curva ROC de 0,97365 e de 0,7871, respectivamente, e de 0,9732 para o conjunto total, valores esses que estão um pouco abaixo daqueles observados por Doig *et al.* (1993). Entretanto, é necessário notar que existem algumas diferenças metodológicas entre o trabalho desses autores e o aqui relatado. Primeiro, o número de variáveis de entrada é diferente (15 e 30, respectivamente). O número de pacientes usados por Doig é superior ao nosso, e, aparentemente, a qualidade dos dados foi mais bem controlada. Esses fatores, principalmente o conjunto e a escolha das variáveis utilizadas, podem ser ponderáveis na determinação do desempenho da rede.

Segundo Doig *et al.* (1993), a preditividade observada poderia ser maior se os conjuntos de dados fossem maiores e tivessem mais casos de óbitos. Nesse caso, o desempenho teórico da rede neural poderia se aproximar de 95 %, o que seria um importante resultado clínico. Acreditamos que também esse seria o caso com os nossos dados, mas o principal fator seria ainda um melhor controle da qualidade dos dados, e um maior empenho em conseguir dados mais homogêneos (ou estratificá-los através de alguma variável, como o código de diagnóstico).

Embora usando o mesmo conjunto de dados do APACHE II, o trabalho de Chen *et al.* (1992) preferiu usar a rede neural para prever severidade de doença, que é uma tarefa de baixa confiabilidade e repetibilidade, uma vez que não existe um "padrão-ouro" (ou seja, um prognóstico fechado, e de fácil determinação, como é o

caso da mortalidade). Este trabalho é de difícil comparação com o nosso, pois não há detalhes suficientes publicados sobre quando e que variáveis eram colhidas, quais eram os critérios de inclusão de pacientes, e qual era o grau de mortalidade. Além disso, o número de pacientes era muito pequeno (35 para o treinamento e apenas 10 para o teste), para se chegar a qualquer conclusão válida. Apesar dessas deficiências, os autores observaram uma correlação interessante a respeito do tamanho da rede neural e do número de ciclos de treinamento, que coincide com as observadas em nosso trabalho. Eles notaram que não havia influência significativa do número de nodos da camada intermediária sobre o desempenho da rede, a partir de duas unidades (havia um aumento significativo de desempenho entre as redes de um e dois nodos intermediários). Ao contrário do nosso trabalho, entretanto, Chen *et al.* (1992) observaram uma degradação da porcentagem de acertos da rede em função do número de ciclos de treinamento, mas apenas para as redes com maior número de nodos intermediários. Não notamos tal efeito com nossos dados.

Com relação aos trabalhos que usaram redes neurais para prever duração da internação na UTI (Tu e Guerriere, 1993; Buchman *et al.*, 1994), nenhum deles se preocupou em avaliar o prognóstico de severidade ou de mortalidade. Como nossos dados não incluíam estas variáveis, fica difícil fazer qualquer comparação. Entretanto, vale a pena comentar que o escore do APACHE (e portanto a mortalidade) se correlaciona bem com a duração da internação, embora não se saiba se existe alguma variável que possa estar criando correlações espúrias. A partir desse ponto de vista, observa-se que o desempenho das redes neurais nessa tarefa não são muito diferentes, quantitativamente, das observadas nos trabalhos de prognóstico de mortalidade.

---

## *Conclusão*

## 6. Conclusão

---

As Redes Neurais Artificiais do tipo *backpropagation* são, tradicionalmente, reconhecidas por sua capacidade de reconhecimento de padrões em situações onde a relação entre a entrada e a saída não é linear e o grupo de dados são abundantes. A boa performance das redes obtida neste trabalho de classificação do prognóstico evolutivo dos pacientes em Unidades de Terapia Intensiva, revela que o problema é classificável.

Podemos concluir que esta abordagem utilizando redes neurais artificiais, foi bem sucedida para a predição de um prognóstico em um grupo de dados extremamente variável. Em relação ao APACHE II, as redes neurais apresentaram um desempenho superior (Tabela 4.2), com a vantagem de obter seus parâmetros de dados regionais, ao invés de modelos universais.

Concluimos também que a utilização apenas da acurácia como um medidor do desempenho das redes neurais não é um bom método. Isto decorre do fato que nem sempre um bom índice de acurácia é acompanhado de um bom resultado de sensibilidade e especificidade. A utilização das Curvas ROC deve ser considerada como uma solução mais completa para avaliar o desempenho global das redes.

A experiência obtida neste trabalho nos mostra que ainda não existe um método ideal para a identificação da melhor rede dentro de um problema de classificação proposto. A utilização dos métodos citados, em conjunto, é que levaram à identificação da melhor topologia de rede.

Acreditamos que o aperfeiçoamento e a melhoria da *performance* dos meios de predição garantirá um aumento na qualidade e segurança ao atendimento do paciente, na alocação de recursos e na avaliação de novas terapias. Quando as redes neurais conseguirem apresentar uma alta preditividade, elas se tornarão um dos mais importantes auxiliares para o processo de tomada de decisão médica em nível individual de pacientes.

## 6.1 Sugestões para Pesquisas Adicionais

Ao final deste primeiro trabalho sobre a utilização de Redes Neurais Artificiais para predição de óbitos em pacientes internados em Unidades de Terapia Intensiva, percebemos que uma série de novas linhas de pesquisa poderiam ser desenvolvidas. A seguir, comentamos algumas:

- Experimentação do modelo desenvolvido em uma outra Unidade de Terapia Intensiva. Desta forma, poderemos avaliar se o modelo pode ser replicado para outros serviços “as is”, ou seja, da forma estabelecida por nós neste trabalho, ou alguns ajustes no modelo deveriam ser realizados;
- Utilização de Redes Neurais Artificiais de aprendizado não supervisionado para realizar uma pré-classificação dos registros dos pacientes. Este tipo de RNA atribui automaticamente categorias ou classes distintas a cada grupo significantemente homogêneo de exemplares, executando uma operação semelhante à análise de aglomerados (*cluster analysis*), uma técnica de classificação numérica multivariada. Por conseguinte, esta RNA poderia agrupar os casos mais semelhantes, à revelia de qualquer padrão classificatório imposto e,



a seguir, a partir de cada agrupamento determinado, uma RNA semelhante à desenvolvida nesta pesquisa poderia ser utilizada para realizar a classificação do prognóstico. Estaríamos assim trabalhando com redes neurais em castata onde a primeira teria a função de classificação do grupo a que pertence o paciente, e a segunda teria a função de avaliar o prognóstico.

---

- Avaliação temporal das variáveis de entrada. Neste trabalho, utilizamos os valores disponíveis na base de dados da UTI do HC-UNICAMP que registrava apenas os piores valores registrados dentro das primeiras 24 horas a partir da admissão do paciente na unidade. Propomos a realização de um estudo temporal das variáveis identificadas, onde os seus valores diários fossem apresentados às redes, de forma que estas pudessem avaliar a tendência global do conjunto para procederem à classificação do prognóstico do paciente. Assim, teríamos um método baseado na análise do comportamento dinâmico da patologia vigente e não apenas baseado nos valores determinados em um quadro (*frame*) de tempo.

---

## ***Bibliografia***

## 7. Bibliografia

---

- Ambroso, C.; Bowes, C.; Chambrin, M.C.; Gilhooly, K.; Green, C.; Kari, A.; Logie, R.; Marraro, G.; Mereu, M.; Rembold, P.; *et al.* (1992) - INFORM: European survey of computers in intensive care units. *Int. J. Clin. Monit. Comput.*, 9(1): 53-61.
- Arruda-Botelho, A.G.; Sabbatini, R.M.E. - A high-level language and microcomputer program for the description and simulation of neural architectures. *Resumos da VI Reunião Anual da Federação de Sociedades de Biologia*, Caxambú, MG, p. 628, agosto 1991.
- Ashutosh, K.; Lee, H.; Mohan, C.K.; Ranka, S.; Mehrotra, K.; Alexander, C. - Prediction criteria for successful weaning from respiratory support: statistical and connectionist analyses. *Crit. Care Med.*, 20(9): 1295-1301, 1992.
- Astion, M.L.; Wilding, P. - Application of neural networks to the interpretation of laboratory data in cancer diagnosis. *Clin.Chem.*, 38(1):34-8, 1992a.
- Avanzolini, G.; Barbini, P.; Gnudi, G. - Unsupervised learning and discriminant analysis applied to identification of high risk postoperative cardiac patients. *Int. J. Bio-Med. Comput.*, 25(2-3):207-221, 1990.
- Baxt, W.G. - Use of an artificial neural network for data analysis in clinical decision making.: the diagnosis of acute coronary occlusion. *Neural Computation*, 2: 480-9, 1991a.
- Baxt, W.G. - Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann.Intern. Med.*, 115: 843-848, 1991b.
- Beck, J.R.; Schultz, E.K. - The use of relative operating characteristic (ROC) curves in test performance evaluation. *Arch. Path. Lab. Med.* 110: 13-20, 1986.
- Bentz, H.J.; Kotter, R. - Neural networks in Medicine: present day and future applications. *Softw. Kur. Med. Psychol.* 3(5): 129-37, 1990.
- Bergeron, B.P.; Morse, A.N.; Greenes, R.A. - A generic neural network-based tutorial supervisor for computer aided instruction. *Proceed. 14th. Annual Symp. Comput. Applic. Med. Care*. New York: IEEE Computer Society Press, p. 435-439, 1991.
- Blumenfeld, B. - A connectionist approach to the recognition of trend in time ordered medical parameters. *Proc. 13th. Ann. Symp. Comput. Appl. Med. Care*. New York: IEEE Press, p. 288-294, 1989.

- Blumenfeld, B. - A connectionist approach to the recognition of trend in time ordered medical parameters. *Comput. Methods Progr. Biomed.*, 32 (1): 53-62, 1990.
- Boone, J.M.; Gross, G.; Shaber, G. - Computer-aided radiologic diagnosis using neural networks. *Proceed. IJCNN-90 - International Joint Conference on Neural Networks*, Washington, DC., p. 98-101, Vol. 2, 1990.
- Boone, J.M.; Gross, G.W.; Greco-Hunt, V. - Neural networks in radiologic diagnosis. I. Introduction and illustration. *Invest. Radiol.*, 25(9): 1012-1016, 1990a.
- Boone, J.M.; Sigillito, V.G.; Shaber, G.S. - Neural networks in Radiology - An introduction and evaluation in a signal detection task. *Med. Phys.*, 17(2):234-241, 1990b.
- Bounds, D.; Lloyd, P.; Matthew, B.; Wadell, G. - A multilayer perceptron network for the diagnosis of back pain. *Proc. IEEE Int. Conf. Neural Networks*, Vol.II, p. 481-9, 1988.
- Bounds, D.G.; Lloyd, P.J.; Matthew, B.G. - A comparison of neural network and other pattern recognition approaches to the diagnosis of low back disorders. *Neural Networks*, 3(3): 583-592, 1990.
- Brown, T.H.; Kairiss, E.W.; Keenan, C.L. - Hebbian synapses - biophysical mechanisms and algorithms. *Ann. Rev. Neurol.*, 13:475-511, 1990.
- Buchman, T.G.; Kubos, K.L.; Seidler, A.J.; Siegforth, M.J. (1994) - A comparison of statistical and connectionist models for the prediction of chronicity in a surgical intensive care unit. *Crit. Care Med.*, 22(5): 750-62.
- Chen, Y.; Chen, L.; Chi, M.; Cai, X - The use of neural network in decision making for nursing diagnosis. *Proceed. 7th World Congress Medical Informatics (MEDINFO 92)* - K.C. Lun; Degoulet, P.; Piemme, T.E.; Rienhoff, O. (Eds.). Amsterdam: North-Holland, p. 1063, 1992.
- Chen, S.; Evens, M.; Trace, D.A.; Naeymi-Rad, F. (1992) - Severity measurements using neural networks.
- Clemmer, T.P.; Gardner, R.M. (1992) - Medical informatics in the intensive care unit: state of the art 1991. *Int. J. Clin. Monit. Comput.*, 8(4): 237-50.
- Cliquet Jr., A.; Mendeck, A.; Quesnel, D.R.F.; Sovi, F.X.; Felipe Jr., P.; Oberg, T.D.; Lemos, G.J.P.; Guimaraes, E.A.; Quevedo, A.A.F. - Restoration of upper limb function to tetraplegics using neural networks and a voice processing system. In: *Proceed. VI Mediterrean Conf. Med. Biol. Engineer.* International Federation for Medical and Biological Engineering, Capri, Italia, 1992a.
- Cliquet Jr., A.; Mendeck, A.; Quesnel, D.R.F.; Sovi, F.X.; Felipe Jr., P.; Oberg, T.D.; Lemos, G.J.P.; Guimaraes, E.A.; Quevedo, A.A.F. - A neural network-voice controlled neuromuscular electrical stimulation system for tetraplegics.

In: *Rehabilitation Engineering Society of North America - RESNA International'92*. Toronto, Canada, 1992b.

Cohen, M.E.; Hudson, D.L.; Anderson, M.D. - A neural network learning algorithm with medical applications. *Proc. 13th. Ann. Symp. Comput. Appl. Med. Care*. New York: IEEE Press, p. 307-311, 1989a.

Collura, T.F.; Jacobs, E.C.; Braun, D.S.; Burgess, R.C. (1993) - EView--a workstation-based viewer for intensive clinical electroencephalography. *IEEE Trans Biomed Eng.*, 40(8): 736-44.

Cullen, J. D. - Reassessing critical care: Illness, outcome, and cost. *Critical Care Medicine*, 17(12)Supplement: S172 - S173, 1989.

DaPonte, J.S.; Parikh, J.; Katz, D.A. - Detection of liver metastasis using the backpropagation algorithm and linear discriminant analysis. *Proceed. SPIE Conf. Applications of Artificial Intelligence and Neural Networks*, Bellingham, WA: SPIE, p. , 1991.

Dodd, N. (1990) - Intensive care ward monitoring and "default" training *Proceed. Int. Neural Network Conf. (INNC'90)*. Dordrecht: Kluwer Acad.Publ., p.326-329.

Doig, G.S.; Inman, K.J.; Sibbald, W.J.; Martin, C.M.; Robertson, J.M. (1993) - Modeling mortality in the intensive care unit: comparing the performance of a back-propagation, associative-learning neural network with multivariate logistic regression. *Proc. Annu. Symp. Comput. Appl. Med. Care*, p. 361-5.

Dojat, M.; Brochard, L.; Lemaire, F.; Harf, A. (1992) - A knowledge-based system for assisted ventilation of patients in intensive care units. *Int. J. Clin. Monit. Comput.*, 9(4): 239-50.

Eberhart, R.C.; Dobbins, R.W. - Early neural network development history: the age of Camelot. *IEEE Engineer. Med. Biol. Magaz.*, 9(3):, 15-18, 1990a.

Eberhart, R.C.; Dobbins, R.W. - Neural networks versus bayesian diagnosis of appendicitis. *Proceed. 12th. Ann. Conf. IEEE Engineer. Med. Biol. Soc.*, Philadelphia, PA, USA, 1990b.

Eberhart, R.C.; Dobbins, R.W.; Hutton, L.V. - Neural network paradigm comparisons for appendicitis diagnosis. *Proc. 4th Ann. IEEE Symp. Computer-Based Med. Syst.*, p. 298-304, 1991.

Fahlman, S.E. - Faster learning variations on back-propagation: an empirical study. In: *Proceed. 1988 Connectionist Models Summer School* (Touretzky, R., Ed.). Morgan Kauffman, 1989.

Felipe Jr., P.; Sabbatini, R.M.E.; Carvalho-Júnior, P.; Beseggio, R.E.; Terzi, R.G.G. - Outcome prediction for critical patients under intensive care, using

- backpropagation neural networks. *Anais I Fórum Nacional de Ciência e Tecnologia em Saúde*, Caxambu, MG, pp. 344-47, novembro de 1992.
- Ferraris VA, Propp ME. - Outcome in critical care patients: a multivariate study. *Crit Care Med.* Jul 20(7):967-76, 1992.
- Ferrer Salvans, P.; Alonso Valles, L.; Osorio Gullon, A.; Suarez Menendez, M.; Vilaplana Birba, J.; Rubio Garcia, R.; Valino Blanco, J. - An empirical comparison of backpropagation and multivariate discriminant analysis to the prediction of colon cancer recurrence after surgical treatment. *International Neural Network Conference. Proceed. INNC 90 Paris*, July 1990. Dordrecht, Netherlands: Kluwer. Vol.1, p. 348, 1990.
- Forrest, D.V.; Flory, M.J.; Anderson, S. - Neural network programming. *NY State J. Med.*, 91(12):553, 1991.
- Frenster, J.H. - Neural networks for pattern recognition in medical diagnosis. *Proceed. 12th. Ann. Conf. IEEE Engineer. Med. Biol. Soc.*, Philadelphia, PA, USA, 1990.
- Gillespy, T.; Staab, E.V.; Lawrence, E. (1990) - Electronic imaging in a teaching hospital intensive care unit: evaluation of the clinical review system. *J. Digit. Imaging*, 3(2): 124-8.
- Gustafson, D.H.; Fryback, D.; Rose, J. *et al* - *Theoretical Methodology for Severity Index Development*. Madison, WI: Center for Health Systems Research and Analysis. University of Wisconsin, 1981.
- Halpern, N.A.; Thompson, R.E.; Greenstein, R.J. (1992) - A computerized intensive care unit order-writing protocol. *Ann. Pharmacother.*, 26(2): 251-4.
- Hanley, J.A.; McNeil, R.J. - The meaning and use of the area under receiver operating characteristic (ROC) curve. *Radiology*, 143: 29-36, 1982.
- Harrison, R.F.; Marshall, S.J.; Kennedy, R.L. - A connectionist aid to the early diagnosis of myocardial infarction. *Proceed. Third European Conf. Artificial Intelligence in Medicine*, Maastricht, The Netherlands, June 1991a.
- Harrison, R.F.; Marshall, S.J.; Kennedy, R.L. - The early diagnosis of heart attacks: a neurocomputational approach. *Proceed. Internat. Joint Conf. Neural Networks*, Vol.1: p. 1-5, 1991b.
- Hart, A.; Wyatt J. - Connectionist models in medicine: an investigation of their potential. In: Hunter, J.; Cookson, J.; Wyatt, J. (Eds.) *Proceedings of Artificial Intelligence in Medicine Europe*. Lecture Notes in Medical Informatics, 38: 115-124, 1989.
- Hart, A.; Wyatt, J. - Evaluating black-boxes as medical decision aids: issues arising from a study of neural networks. *Medical Informatics*, 15(3): 229-36, 1990.

- Higgins, S.B.; Jiang, K.; Swindell, B.B.; Bernard, G.R. (1991) - A graphical ICU workstation. *Proc. Annu. Symp. Comput. Appl. Med. Care*, p. 783-7.
- Hopfield, J.J. - Neural networks and physical systems with emergent collective computational abilities. *Proc. Nat. Acad. Sci. USA*, 79: 2554-2558, 1982.
- Howell, R.M.; Fowler, D. - A neural network as an instrument of prediction. *Proceed. 14th Ann. Symp. Computer Applications in Medical Care*. New York: IEEE Computer Society Press, p. 299-302, 1990.
- Hripcsak, G. - Problem solving using neural networks. *MD Computing*, 5(3): 25-37, 1988a.
- Huang, W.Y. & Lippman, R.P. - Comparisons between neural nets and conventional classifiers. *Proc. 1st. International Neural Networks Conference*. IEEE, New York, p. 485, 1987.
- Hudson, D.L.; Cohen, M.E.; Lammers, R.K. - Use of a hybrid expert system to predict wound infections. *Proceed. 7th World Congress Medical Informatics (MEDINFO 92)* - K.C. Lun; Degoulet, P.; Piemme, T.E.; Rienhoff, O. (Eds.). Amsterdam: North-Holland, p. 546-551, 1992.
- Hulley, S.B.; Cummings, S.R. - *Designing Clinical Research. An Epidemiological Approach*. Baltimore: William & Wilkins, 1988.
- Imhoff, M. (1992) - Acquisition of ICU data: concepts and demands. *Int. J. Clin. Monit. Comput.*, 9(4):229-37.
- Jansen, B.H. - Artificial neural nets for K-complex detection. *IEEE Engineer. Med. Biol. Magaz.*, 9(3): 50-52, 1990.
- Jayaweers, A.; Drake, K.C.; Abbott, R.; Kaul, S. - Determination of long-term outcome in patients with coronary artery disease using an artificial neural network. *J. Amer. Coll. Cardiol.*, 21(2): 7A, 1993.
- Kampmann, J.; Lau, G.; Kropp, S.; Schwarzer, E.; Sande, C.H. (1991) - Connection of electronic medical devices in ICU according to the standard 'MIB'. *Int. J. Clin. Monit. Comput.*, 8(3): 163-6.
- Kari, A.; Ruokonen, E.; Takala, J. (1990) - Comparison of acceptance and performance of automated and manual data management systems in intensive care. *Int. J. Clin. Monit. Comput.*, 7(3): 157-62.
- Kinser, W.; Indrayanto, A.; Langi, A. - A study of backpropagation, counterpropagation and adaptive resonance theory neural network models. *Proceed. 12th. Ann. Conf. IEEE Engineer. Med. Biol. Soc.*, Philadelphia, PA, USA, 1990.

- Knaus, W. A.; Draper, E. A.; Wagner, D.; Zimmerman, J. E. - APACHE II: A severity of disease classification system. *Critical Care Medicine*, 13 (10): 818 - 829, 1985.
- Knaus, W.A.; Wagner, D.; Draper, E. - Development of APACHE. *Critical Care Medicine*, 17(12)Supplement: S176 - S185, 1989.
- 
- Kohonen, T. - *Self Organization and Associative Memory*. Berlin: Springer-Verlag, 1984.
- Kolmogorov, A.N. - On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *Dokl. Akad. Nauk USSR*, 114, 953-956, 1957.
- Kosko, B. - Bidirectional associative memories. *IEEE Trans. Syst. Man, Cybern.*, SMC-18: 42-60, 1988.
- Lau, F.; Vincent, D.D. (1993) - Formalized decision support for cardiovascular intensive care. *Comput. Biomed. Res.*, 26(3): 294-309.
- Lemeshow, S.; Teres, D.; Pastides, H.; Avrunin, J.S.; Steingrub, J.S. - A method for predicting survival and mortality of ICU patients using objectively devised weights. *Critical Care Medicine*, 13: 519, 1985.
- Lippman, R.P. - An introduction to computing with neural nets. *IEEE Trans. Acoust. Speech & Sign. Process.*, 36: 4-22, 1987.
- McIntyre, J.W.; Nelson, T.M. (1989) - Application of automated human voice delivery to warning devices in an intensive care unit: a laboratory study. *Int. J. Clin. Monit. Comput.*, 6(4): 255-62.
- Meistrell, M.L.; Spackman, K.A. - Evaluation of neural network performance by receiver operating characteristic (ROC) analysis: examples from the biotechnology domain. *Proc. 13th. Ann. Symp. Comput. Appl. Med. Care*. New York: IEEE Press, p. 295-301, 1989 and *Comput. Methods Programs. Biomed.* 32(1):73-80, 1990.
- Moura Júnior, J.M.A.; Danna Neto, L.; Carvalho, A.C.M.; Ohl, A.; Anção, M.S.; Sigulem, D. - Clinical history simulation model. *Proceed. 7th World Congress Medical Informatics (MEDINFO 92)* - K.C. Lun; Degoulet, P.; Piemme, T.E.; Rienhoff, O. (Eds.). Amsterdam: North-Holland, p. 1607, 1992.
- Mulsant, B. - A neural network as an approach to clinical diagnosis. *MD Computing*, 7(1): 25-36, 1990.
- O'Neill, M.C. - Training back-propagation neural networks to define and detect DNA-binding sites. *Nucleic Acids Res.*, 19(2):313-8, 1991.



- Overall, J.E.; Klett, C.J. - *Applied Multivariate Analysis*. New York: McGraw-Hill, 1972.
- Owens D, Sox HJ. - Medical decision making: probabilistic medical reasoning. *Ann Intern Med.*, 18(2):70-116, 1990.
- Ravdin, P.M.; Clark, G.M.; Hilsenbeck, S.G.; Owens, M.A.; Vendely, P.; Pandian, M.R.; McGuire, W.L. - A demonstration that breast cancer recurrence can be predicted by neural network analysis. *Breast Cancer Res. Treat.*, 21(1): 47-53, 1992.
- Reggia, J.A. - Artificial neural systems in medical science and practice. *MD Computing*, 5(3): 4-6, 1988.
- Reggia, J.A. - Neural computation in Medicine. *Artificial Intelligence in Medicine* (in press), 1992.
- Reggia, J.A.; Sutton III, G.G. - Self-processing networks and their biomedical implications. *Proc. IEEE*, 76(6): 580-92, 1988.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. - Learning representations by back-propagating errors. *Nature*, 323, 533-536, (1986).
- Rutledge R, Fakhry SM, Rutherford EJ, Muakkassa F, Baker CC, Koruda M, Meyer AA. - Acute Physiology and Chronic Health Evaluation (APACHE II) score and outcome in the surgical intensive care unit: an analysis of multiple intervention and outcome variables in 1,238 patients. *Crit Care Med.* 1991 Aug;19(8):1048-53.
- Sabbatini, R.M.E. - A multilayered neural network for processing 2D tomographic images in neurosurgery. *Proceed. IEEE Nuclear Science and Medical Imaging Symposium*, Orlando FL, USA, 1992c (abstract).
- Sabbatini, R.M.E. - A practical course on knowledge engineering for Health professionals. In: Barber, B.; Cao, D.; Qin, D.; Wagner, G. (Eds.) - *Proceedings Sixth World Congress of Medical Informatics (MEDINFO'89, Beijing, China and Singapore)*: International Medical Informatics Association, North-Holland Press, Amsterdam, Netherlands, p. 949-952, 1989.
- Sabbatini, R.M.E. - Applications of connectionist systems in Biomedicine. *Proceed. 7th World Congress on Medical Informatics (MEDINFO 92)*. International Federation of Medical Informatics. North Holland, Amsterdam, p. 418-426, 1992a.
- Sabbatini, R.M.E. - Microcomputers Applications in Medicine: a Review. Geneva: World Health Organization, Information Systems Support Division, mimeo, 1985.

- Sabbatini, R.M.E.; Carvalho-Júnior, P.M.; Freitas, N.C.; Felipe Jr., P. - NEURONET: a connexionist shell system for automatic pattern classification in Biomedicine. *Anais I Forum Nacional de Ciência e Tecnologia em Saúde*, Caxambu, MG, Brazil, 1992.
- Sabbatini, R.M.E.; Ortiz, J.; Ghefter, C.G.M. - Neural network-based approach to outcome prognosis for patients with diastolic dysfunction. **In:** Hudson, D. & Cohen, M. (Eds.) - *Comparative Approaches in Medical Decision Making*. Amsterdam: Elsevier, 1993a (*in press*).
- Schwaiger, J.; Haller, M.; Finsterer, U. (1992) - A framework for the knowledge-based interpretation of laboratory data in intensive care units using deductive database technology. *Proc. Annu. Symp. Comput. Appl. Med. Care.*, p. 13-7.
- Seneff, M.; Knaus, W.A. - Predicting patient outcome from intensive care: A guide to APACHE, MPM, SAPS, PRISM, and other prognostic scoring systems. *J Intensive Care Med* 5:33-52, 1990.
- Shoemaker, W.C. - Methodologic assessment of outcome. *Critical Care Medicine*, 17(12)Supplement: S169, 1989.
- Shortliffe, E.H. - Clinical decision support systems. **In:** Shortliffe, E.H.; Perreault, L.E.; Wiederhold, G.; Fagan, L.M. - *Medical Informatics. Computer Applications in Health Care*. New York: Addison & Wesley, p. 466-502, 1990.
- Simpson, P.K. - *Artificial Neural Systems: Foundations, Paradigms, Applications and Implementations*. Pergamon Press, London, 1990.
- Smith, J.W.; Everhart, J.E.; Dickson, W.C.; Knowler, W.C.; Johannes, R.S. - Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. *Proc. 12th. Ann. Symp. Comput. Appl. Med. Care*. New York: IEEE Press, p. 261-265, 1988.
- Stubbs, D.F. - Neurocomputers. *MD Computing*, 5(3): 14-24, 1988.
- Teres, D. - Peer review, publication policy, and APACHE. *Critical Care Medicine*, 17(12)Supplement: S169 - S172, 1989.
- Terzi, R.G.G.; Joffily, M.; Scandiucci, M.; Meira, S.; Beseggio, R.E.; Gomez, M.I.; Araujo, S.; Mühlen, S.S. - Comparison of three predictive models of hospital death in ICU admissions for head trauma (*submitted for publication*), 1992.
- Tu J.V.; Guerriere, M.R. (1993) - Use of a neural network as a predictive instrument for length of stay in the intensive care unit following cardiac surgery. *Comput. Biomed. Res.*, 26(3): 220-9.
- Uckun, S.; Dawant, B.M.; Lindstrom, D.P. (1993) - Model-based diagnosis in intensive care monitoring: the YAQ approach. *Artif. Intell. Med.*, 5(1):31-48.

van den Heuvel, J.; Stemerding, J.D.; Bogers, A.J.; Bree, D.S. (1990) - GUUS: an expert system in the intensive care unit. *Int. J. Clin. Monit. Comput.*; 171-5.

Wenkebach, U.; Pollwein, B.; Finsterer, U. (1992) - Visualization of large datasets in intensive care. *Proc. Annu. Symp. Comput. Appl. Med. Care.* p. 18-22.

---

Wiederhold, G. & Perreault, L.E. - Clinical research systems. **In:** Shortliffe, E.H.; Perreault, L.E.; Wiederhold, G.; Fagan, L.M. - *Medical Informatics. Computer Applications in Health Care*. New York: Addison & Wesley, p. 503-532, 1990.

Yoshida, K.; Hayashi, Y.; Imura, A.; Shimada, N. - TI:Fuzzy neural expert system for diagnosing hepatobiliary disorders. *Proc. Int. Conf. Fuzzy Logics & Neural Networks (IIZUKA'90)*. Iizuka, Japan, p. 539-543, 1990.

---

## *Apêndices*

## 8. Apêndice A

---

### 8.1 Descrição do programa Neuronet

O programa Neuronet, atualmente em sua versão 2.09, é um aplicativo genérico para implementação de três tipos de redes neurais organizadas em camadas múltiplas e é operado através de uma interface interativa de usuário baseada em menus de fácil utilização. As seguintes funções estão disponíveis:

- **Definição:** permite criar e modificar todos os parâmetros e arquitetura de uma rede;
- **Exemplos:** permite adicionar, remover e visualizar exemplos do conjunto de treinamento;
- **Arquivamento:** permite gravar no disco a definição da rede, juntamente com o conjunto de exemplos de treinamento. É gravado o último estado de treinamento da rede. Uma rede previamente gravada pode ser recuperada a partir do disco com um comando de carregamento.
- **Treinamento:** permite efetuar um número fixo de ciclos de treinamento utilizando o conjunto de exemplos. O programa exibe continuamente na tela, durante o treinamento, o número de ciclos completados, o tempo gasto em segundos, o erro global da rede, e exibe um gráfico da curva do aprendizado, em tempo real;
- **Parâmetros da rede:** o usuário pode modificar quaisquer dos parâmetros predefinidos para uma rede (vide adiante);

- **Classificação:** este módulo permite a entrada de padrões pelo teclado, ou a partir de um arquivo de teste, e realizar a sua classificação pela rede treinada carregada em memória. Os resultados dessa classificação podem simplesmente ser exibidos em tela, ou o programa pode, opcionalmente, gravar um arquivo contendo os valores de saída desejados e obtidos, o qual posteriormente poderá ser analisado pelos módulos seguintes;
- **Análise da rede:** como o NEURONET é voltado para a experimentação, foram implementados diversas ferramentas de análise de uma rede. O programa pode mostrar todos os dados registrados da rede (inclusive todas as matrizes de pesos de conexão); bem como analisar o grau de contribuição das variáveis de entrada sobre as de saída. Existe também um módulo para analisar e comparar diversas curvas de aprendizado, na forma de gráficos de evolução do erro global ou da derivada da curva de aprendizado.
- **Análise de desempenho:** o desempenho de uma rede treinada pode ser avaliado de forma bastante completa, submetendo ao programa o arquivo de resultados gerado pelo módulo de classificação:
  1. Cálculo de acurácia, sensibilidade, especificidade e preditividade, erro médio, teste estatístico de significância por qui-quadrado, distribuição dos valores de saída, lista de exemplos discrepantes, etc., a partir de um valor de corte determinado;
  2. Cálculo e exibição da curva ROC (Receiver Operating Characteristic) para um conjunto de pontos de corte definidos previamente pelo usuário em um arquivo de disco. Os valores das curvas ROC podem ser gerados em um arquivo ASCII em disco, para posterior exame por outro módulo do

programa, ou por outros programas de análise estatística e elaboração de gráficos.

## 8.2 Parâmetros de Definição e Treinamento

Os diversos parâmetros de definição de uma rede neural podem ser escolhidos através de um menu próprio de acesso:

- **Arquitetura:** O usuário pode escolher ou modificar o tamanho da rede (número de nodos das camadas de entrada, intermediária e saída), bem como o tipo da rede a ser utilizada. Existem três arquiteturas disponíveis:
  - perceptron de duas camadas com regra de aprendizado supervisionado do tipo delta generalizado, com uma camada de nodos de entrada (com um máximo de 50 elementos) e uma camada de nodos de saída (com um máximo de 50 elementos);
  - perceptron de três camadas com regra de aprendizado supervisionado do tipo retropropagação de erros, com uma camada de nodos de entrada (com um máximo de 50 elementos); uma camada de nodos intermediários (com um máximo de 20 elementos) e uma camada de nodos de saída (com um máximo de 50 elementos);
  - quantizador linear vetorial (LVQ) de duas camadas, com regra de aprendizado supervisionado ou não supervisionado de Kohonen, com uma camada de nodos de entrada (com um máximo de 50 elementos) e uma camada de nodos de saída (com um máximo de 50 elementos);
- **Tipo de associação:** rede auto-associativa ou hetero-associativa;

- **Tipos de entradas e saídas:** podem ser analógicas (a ativação do nodo é um número real no intervalo  $[0,1]$ ), binárias (0 ou 1) ou gráficas (neste caso, cada nodo é um pixel de uma imagem retangular);
- **Inicialização das conexões:** determinística ou aleatória (no intervalo  $[-1,+1]$ , com a magnitude do intervalo podendo ser escolhida pelo usuário). Este recurso permite alocar valores às matrizes de conexões, antes de ser iniciado o treinamento. Conexões determinísticas permitem reproduzir determinadas condições de treinamento, para fins de comparação experimental;
- **Taxa de aprendizado:** um parâmetro escolhido pelo usuário, e que pode ser mudado manualmente ao longo do aprendizado, no intervalo real  $[0,2]$ . Representa a fração do erro (delta) que é transferida na modificação iterativa e retropropagada das conexões;
- **Momento:** um parâmetro escolhido pelo usuário, e que pode ser mudado manualmente ao longo do aprendizado, no intervalo real  $[0,1]$ . Representa a fração da variação de peso no ciclo de aprendizado imediatamente anterior, a ser transferida para as conexões;
- **Tipo de variação dos pesos:** pode ser do tipo determinística ou aleatória. No último caso, a cada cálculo de peso, o programa adiciona um pequeno valor aleatório, cujo valor médio pode ser escolhido pelo usuário.
- **Apresentação dos exemplos:** existem três possibilidades:
  - os exemplos de treinamento são apresentados sempre na mesma ordem em que se encontram no conjunto (arquivo) de dados
  - a ordem de apresentação dos exemplos é aleatorizada logo ao início do treinamento (primeiro ciclo), e depois permanece constante



— a ordem de apresentação dos exemplos é aleatorizada a cada ciclo de treinamento.

- **Métrica de erro:** é a expressão matemática usada para comparar os valores desejados com os obtidos, nas formas de aprendizado supervisionado. Três formas são disponíveis: erro quadrático (derivada da função de transferência), função de similaridade e função de máxima verossimilhança;
- **Função de transferência do nodo:** pode ser escolhida entre função sigmoideal e função dupla exponencial;

Os valores de entrada e saída são automaticamente normalizados no intervalo  $[0,1]$  antes de poderem ser utilizados pela rede.

A estrutura da rede, uma vez definida ou modificada de forma interativa pelo usuário, é armazenada em um arquivo seqüencial de disco, em ASCII, contendo também um título para a rede, nomes para cada camada e nodo de entrada e de saída, os parâmetros de inicialização, tipo, arquitetura e inicialização, e as matrizes de pesos entrada-saída (para perceptrons de duas camadas), entrada-intermediária, e intermediária-saída (para perceptrons de três camadas), além de todos os exemplos registrados. Os exemplos de teste são armazenados em outro arquivo ASCII separado. Tanto o conjunto de treinamento quanto o de teste podem conter até 300 exemplos, que são carregados na memória central, antes de serem utilizados.

Ao se realizar um treinamento de uma rede neural, o programa permite que o usuário escolha outros parâmetros, que só funcionam durante aquele treinamento, e não são armazenados no arquivo de definição da rede:

- **Número de ciclos de treinamento com o qual a rede será treinada:** o treinamento pode ser feito em etapas, por exemplo, realizar 1.000 ciclos, depois parar e alterar parâmetros, realizar mais 2.000 ciclos, etc.

matemático opcional do tipo Intel 80287 ou 80387, ou ainda microprocessadores com coprocessamento matemático embutido, como o Intel 80486.

---

## 8.4 Desempenho

Devido às limitações da linguagem e da plataforma de hardware escolhidas, o programa NEURONET tem baixa velocidade de execução (medida em torno de 10.000 cálculos de conexão por segundo em um microcomputador do tipo PC 486, com coprocessador numérico) e uma capacidade relativamente pequena (redes máximas de 50 x 20 x 50, com até 300 exemplos).

## 9. Apêndice B

---

**Tabela 9.1:** ANÁLISE DE DETEÇÃO: r1nn20tr.OUT - CURVA ROC  
Arquivo de saída: r1nn20tr.ROC  
Rede (30-20-1) - Classificação do Grupo de Treino

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	70	111	99	2	64,2	97,2	52,9	75,0	41,4	98,2	53,93
0,010	69	154	56	3	79,1	95,8	73,3	84,6	55,2	98,1	101,16
0,050	68	206	4	4	97,2	94,4	98,1	96,3	94,4	98,1	236,65
0,100	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,200	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,300	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,400	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,500	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,600	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,700	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,800	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,900	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,950	64	208	2	8	96,5	88,9	99,0	94,0	97,0	96,3	226,40
0,990	37	208	2	35	86,9	51,4	99,0	75,2	94,9	85,6	110,26
0,999	0	210	0	72	74,5	0,0	100,0	50,0	0,0	74,5	0,00

**Tabela 9.2:** ANÁLISE DE DETEÇÃO: r1nn20ts.OUT - CURVA ROC  
Arquivo de saída: r1nn20ts.ROC  
Rede (30-20-1) - Classificação do Grupo de Teste

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	22	38	37	3	60,0	88,0	50,7	69,3	37,3	92,7	10,05
0,010	19	52	23	6	71,0	76,0	69,3	72,7	45,2	89,7	14,01
0,050	18	56	19	7	74,0	72,0	74,7	73,3	48,6	88,9	15,57
0,100	17	58	17	8	75,0	68,0	77,3	72,7	50,0	87,9	15,21
0,200	17	61	14	8	78,0	68,0	81,3	74,7	54,8	88,4	19,09
0,300	17	62	13	8	79,0	68,0	82,7	75,3	56,7	88,6	20,57
0,400	16	62	13	9	78,0	64,0	82,7	73,3	55,2	87,3	17,63
0,500	16	62	13	9	78,0	64,0	82,7	73,3	55,2	87,3	17,63
0,600	16	62	13	9	78,0	64,0	82,7	73,3	55,2	87,3	17,63
0,700	15	63	12	10	78,0	60,0	84,0	72,0	55,6	86,3	16,25
0,800	14	63	12	11	77,0	56,0	84,0	70,0	53,8	85,1	13,58
0,900	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,950	13	69	6	12	82,0	52,0	92,0	72,0	68,4	85,2	20,81
0,990	11	70	5	14	81,0	44,0	93,3	68,7	68,8	83,3	16,77
0,999	0	75	0	25	75,0	0,0	100,0	50,0	0,0	75,0	0,00

**Tabela 9.3:** ANÁLISE DE DETEÇÃO: r1nn15tr.OUT - CURVA ROC  
Arquivo de saída: r1nn15tr.ROC  
Rede (30-15-1) - Classificação do Grupo de Treino

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	68	130	80	4	70,2	94,4	61,9	78,2	45,9	97,0	66,03
0,010	68	166	44	4	83,0	94,4	79,0	86,7	60,7	97,6	117,90
0,050	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,100	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,200	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,300	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,400	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,500	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,600	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,700	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,800	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,900	68	209	1	4	98,2	94,4	99,5	97,0	98,6	98,1	251,11
0,950	65	209	1	7	97,2	90,3	99,5	94,9	98,5	96,8	236,21
0,990	33	209	1	39	85,8	45,8	99,5	72,7	97,1	84,3	99,80
0,999	0	210	0	72	74,5	0,0	100,0	50,0	0,0	74,5	0,00

**Tabela 9.4:** ANÁLISE DE DETEÇÃO: rlnn15ts.OUT - CURVA ROC  
Arquivo de saída: rlnn15ts.ROC  
Rede (30-15-1) - Classificação do Grupo de Teste

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	19	38	37	6	57,0	76,0	50,7	63,3	33,9	86,4	4,38
0,010	17	51	24	8	68,0	68,0	68,0	68,0	41,5	86,4	8,61
0,050	15	56	19	10	71,0	60,0	74,7	67,3	44,1	84,8	8,56
0,100	15	57	18	10	72,0	60,0	76,0	68,0	45,5	85,1	9,42
0,200	14	59	16	11	73,0	56,0	78,7	67,3	46,7	84,3	9,14
0,300	14	62	13	11	76,0	56,0	82,7	69,3	51,9	84,9	12,33
0,400	14	62	13	11	76,0	56,0	82,7	69,3	51,9	84,9	12,33
0,500	13	63	12	12	76,0	52,0	84,0	68,0	52,0	84,0	11,11
0,600	12	63	12	13	75,0	48,0	84,0	66,0	50,0	82,9	8,85
0,700	11	64	11	14	75,0	44,0	85,3	64,7	50,0	82,1	7,77
0,800	11	64	11	14	75,0	44,0	85,3	64,7	50,0	82,1	7,77
0,900	10	66	9	15	76,0	40,0	88,0	64,0	52,6	81,5	7,82
0,950	10	69	6	15	79,0	40,0	92,0	66,0	62,5	82,1	12,00
0,990	9	71	4	16	80,0	36,0	94,7	65,3	69,2	81,6	13,00
0,999	0	75	0	25	75,0	0,0	100,0	50,0	0,0	75,0	0,00

**Tabela 9.5:** ANÁLISE DE DETEÇÃO: r1nn10tr.OUT - CURVA ROC  
Arquivo de saída: r1nn10tr.ROC  
Rede (30-10-1) - Classificação do Grupo de Treino

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	70	95	115	2	58,5	97,2	45,2	71,2	37,8	97,9	40,98
0,010	68	132	78	4	70,9	94,4	62,9	78,7	46,6	97,1	68,23
0,050	68	168	42	4	83,7	94,4	80,0	87,2	61,8	97,7	121,79
0,100	68	191	19	4	91,8	94,4	91,0	92,7	78,2	97,9	179,31
0,200	68	205	5	4	96,8	94,4	97,6	96,0	93,2	98,1	232,09
0,300	67	209	1	5	97,9	93,1	99,5	96,3	98,5	97,7	246,10
0,400	67	209	1	5	97,9	93,1	99,5	96,3	98,5	97,7	246,10
0,500	67	209	1	5	97,9	93,1	99,5	96,3	98,5	97,7	246,10
0,600	67	209	1	5	97,9	93,1	99,5	96,3	98,5	97,7	246,10
0,700	67	209	1	5	97,9	93,1	99,5	96,3	98,5	97,7	246,10
0,800	67	209	1	5	97,9	93,1	99,5	96,3	98,5	97,7	246,10
0,900	65	209	1	7	97,2	90,3	99,5	94,9	98,5	96,8	236,21
0,950	60	209	1	12	95,4	83,3	99,5	91,4	98,4	94,6	212,28
0,990	34	209	1	38	86,2	47,2	99,5	73,4	97,1	84,6	103,52
0,999	0	210	0	72	74,5	0,0	100,0	50,0	0,0	74,5	0,00

**Tabela 9.6:** ANÁLISE DE DETEÇÃO: r1nn10ts.OUT - CURVA ROC  
Arquivo de saída: r1nn10ts.ROC  
Rede (30-10-1) - Classificação do Grupo de Teste

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	20	33	42	5	53,0	80,0	44,0	62,0	32,3	86,8	3,62
0,010	19	40	35	6	59,0	76,0	53,3	64,7	35,2	87,0	5,37
0,050	18	52	23	7	70,0	72,0	69,3	70,7	43,9	88,1	11,59
0,100	16	56	19	9	72,0	64,0	74,7	69,3	45,7	86,2	10,68
0,200	16	62	13	9	78,0	64,0	82,7	73,3	55,2	87,3	17,63
0,300	15	63	12	10	78,0	60,0	84,0	72,0	55,6	86,3	16,25
0,400	14	63	12	11	77,0	56,0	84,0	70,0	53,8	85,1	13,58
0,500	14	63	12	11	77,0	56,0	84,0	70,0	53,8	85,1	13,58
0,600	13	63	12	12	76,0	52,0	84,0	68,0	52,0	84,0	11,11
0,700	12	64	11	13	76,0	48,0	85,3	66,7	52,2	83,1	9,96
0,800	11	68	7	14	79,0	44,0	90,7	67,3	61,1	82,9	13,01
0,900	11	69	6	14	80,0	44,0	92,0	68,0	64,7	83,1	14,76
0,950	10	69	6	15	79,0	40,0	92,0	66,0	62,5	82,1	12,00
0,990	8	71	4	17	79,0	32,0	94,7	63,3	66,7	80,7	10,23
0,999	0	75	0	25	75,0	0,0	100,0	50,0	0,0	75,0	0,00



**Tabela 9.7:** ANÁLISE DE DETEÇÃO: r1nn5tr.OUT - CURVA ROC  
Arquivo de saída: r1nn5tr.ROC  
Rede (30-5-1) - Classificação do Grupo de Treino

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	70	131	79	2	71,3	97,2	62,4	79,8	47,0	98,5	74,06
0,010	70	168	42	2	84,4	97,2	80,0	88,6	62,5	98,8	130,34
0,050	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,100	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,200	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,300	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,400	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,500	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,600	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,700	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,800	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,900	70	208	2	2	98,6	97,2	99,0	98,1	97,2	99,0	256,32
0,950	68	208	2	4	97,9	94,4	99,0	96,7	97,1	98,1	246,16
0,990	38	208	2	34	87,2	52,8	99,0	75,9	95,0	86,0	114,09
0,999	0	210	0	72	74,5	0,0	100,0	50,0	0,0	74,5	0,00

**Tabela 9.8:** ANÁLISE DE DETEÇÃO: r1nn5ts.OUT - CURVA ROC  
Arquivo de saída: r1nn5ts.ROC  
Rede (30-5-1) - Classificação do Grupo de Teste

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	21	42	33	4	63,0	84,0	56,0	70,0	38,9	91,3	10,52
0,010	18	52	23	7	70,0	72,0	69,3	70,7	43,9	88,1	11,59
0,050	17	57	18	8	74,0	68,0	76,0	72,0	48,6	87,7	14,08
0,100	17	59	16	8	76,0	68,0	78,7	73,3	51,5	88,1	16,42
0,200	16	62	13	9	78,0	64,0	82,7	73,3	55,2	87,3	17,63
0,300	16	63	12	9	79,0	64,0	84,0	74,0	57,1	87,5	19,11
0,400	16	64	11	9	80,0	64,0	85,3	74,7	59,3	87,7	20,72
0,500	16	65	10	9	81,0	64,0	86,7	75,3	61,5	87,8	22,45
0,600	16	65	10	9	81,0	64,0	86,7	75,3	61,5	87,8	22,45
0,700	15	66	9	10	81,0	60,0	88,0	74,0	62,5	86,8	21,13
0,800	14	67	8	11	81,0	56,0	89,3	72,7	63,6	85,9	19,89
0,900	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,950	12	67	8	13	79,0	48,0	89,3	68,7	60,0	83,8	14,08
0,990	9	72	3	16	81,0	36,0	96,0	66,0	75,0	81,8	15,28
0,999	0	75	0	25	75,0	0,0	100,0	50,0	0,0	75,0	0,00

**Tabela 9.9:** ANÁLISE DE DETEÇÃO: r1nn5tr.OUT - CURVA ROC  
 Arquivo de saída: r1nn5tr.ROC  
 Rede (30-5-1) - Classificação do Grupo de Treino (apresentação de  
 exemplos do tipo determinístico)

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	70	149	61	2	77,7	97,2	71,0	84,1	53,4	98,7	97,46
0,010	70	201	9	2	96,1	97,2	95,7	96,5	88,6	99,0	225,06
0,050	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,100	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,200	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,300	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,400	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,500	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,600	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,700	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,800	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,900	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,950	70	210	0	2	99,3	97,2	100,0	98,6	100,0	99,1	266,40
0,990	66	210	0	6	97,9	91,7	100,0	95,8	100,0	97,2	246,23
0,999	0	210	0	72	74,5	0,0	100,0	50,0	0,0	74,5	0,00

**Tabela 9.10:** ANÁLISE DE DETEÇÃO: rlnn5ts.OUT - CURVA ROC  
Arquivo de saída: rlnn5ts.  
Rede (30-5-1) - Classificação do Grupo de Teste (apresentação de  
exemplos do tipo determinístico)

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0,001	19	49	26	6	68,0	76,0	65,3	70,7	42,2	89,1	11,33
0,010	16	63	12	9	79,0	64,0	84,0	74,0	57,1	87,5	19,11
0,050	15	66	9	10	81,0	60,0	88,0	74,0	62,5	86,8	21,13
0,100	15	66	9	10	81,0	60,0	88,0	74,0	62,5	86,8	21,13
0,200	14	67	8	11	81,0	56,0	89,3	72,7	63,6	85,9	19,89
0,300	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,400	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,500	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,600	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,700	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,800	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,900	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,950	13	67	8	12	80,0	52,0	89,3	70,7	61,9	84,8	16,90
0,990	12	68	7	13	80,0	48,0	90,7	69,3	63,2	84,0	15,79
0,999	0	75	0	25	75,0	0,0	100,0	50,0	0,0	75,0	0,00

**Tabela 9.11:** ANÁLISE DE DETEÇÃO: r1nn5tt.OUT - CURVA ROC  
 Arquivo de saída: r1nn5tt.  
 Rede (30-5-1) - Classificação de todos os registros (Grupo de Treino + Grupo de Teste).

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0.001	70	228	82	2	78.0	97.2	73.5	85.4	46.1	99.1	119.21
0.010	70	264	46	2	87.4	97.2	85.2	91.2	60.3	99.2	183.67
0.050	70	302	8	2	97.4	97.2	97.4	97.3	89.7	99.3	316.27
0.100	70	305	5	2	98.2	97.2	98.4	97.8	93.3	99.3	332.47
0.200	70	305	5	2	98.2	97.2	98.4	97.8	93.3	99.3	332.47
0.300	70	305	5	2	98.2	97.2	98.4	97.8	93.3	99.3	332.47
0.400	70	305	5	2	98.2	97.2	98.4	97.8	93.3	99.3	332.47
0.500	70	305	5	2	98.2	97.2	98.4	97.8	93.3	99.3	332.47
0.600	70	306	4	2	98.4	97.2	98.7	98.0	94.6	99.4	338.16
0.700	70	306	4	2	98.4	97.2	98.7	98.0	94.6	99.4	338.16
0.800	70	306	4	2	98.4	97.2	98.7	98.0	94.6	99.4	338.16
0.900	70	306	4	2	98.4	97.2	98.7	98.0	94.6	99.4	338.16
0.950	63	306	4	9	96.6	87.5	98.7	93.1	94.0	97.1	294.32
0.990	40	307	3	32	90.8	55.6	99.0	77.3	93.0	90.6	168.87
0.999	0	310	0	72	81.2	0.0	100.0	50.0	0.0	81.2	0.00

**Tabela 9.12:** ANÁLISE DE DETEÇÃO: r1nn10tt.OUT - CURVA ROC  
 Arquivo de saída: r1nn10tt.  
 Rede (30-10-1) - Classificação de todos os registros (Grupo de Treino  
 + Grupo de Teste).

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0.001	68	130	180	4	51.8	94.4	41.9	68.2	27.4	97.0	32.38
0.010	68	167	143	4	61.5	94.4	53.9	74.2	32.2	97.7	53.23
0.050	68	209	101	4	72.5	94.4	67.4	80.9	40.2	98.1	88.16
0.100	68	209	101	4	72.5	94.4	67.4	80.9	40.2	98.1	88.16
0.200	68	209	101	4	72.5	94.4	67.4	80.9	40.2	98.1	88.16
0.300	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.400	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.500	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.600	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.700	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.800	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.900	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.950	66	244	66	6	81.2	91.7	78.7	85.2	50.0	97.6	124.87
0.990	34	269	41	38	79.3	47.2	86.8	67.0	45.3	87.6	40.67
0.999	0	310	0	72	81.2	0.0	100.0	50.0	0.0	81.2	0.00

**Tabela 9.13:** ANÁLISE DE DETEÇÃO: rlnn15tt.OUT - CURVA ROC  
 Arquivo de saída: rlnn15tt.  
 Rede (30-15-1) - Classificação de todos os registros (Grupo de Treino  
 + Grupo de Teste).

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0.001	68	130	180	4	51.8	94.4	41.9	68.2	27.4	97.0	32.38
0.010	68	166	144	4	61.3	94.4	53.5	74.0	32.1	97.6	52.57
0.050	68	209	101	4	72.5	94.4	67.4	80.9	40.2	98.1	88.16
0.100	68	209	101	4	72.5	94.4	67.4	80.9	40.2	98.1	88.16
0.200	68	209	101	4	72.5	94.4	67.4	80.9	40.2	98.1	88.16
0.300	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.400	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.500	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.600	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.700	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.800	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.900	68	210	100	4	72.8	94.4	67.7	81.1	40.5	98.1	89.21
0.950	65	238	72	7	79.3	90.3	76.8	83.5	47.4	97.1	111.31
0.990	33	269	41	39	79.1	45.8	86.8	66.3	44.6	87.3	37.71
0.999	0	310	0	72	81.2	0.0	100.0	50.0	0.0	81.2	0.00

**Tabela 9.14:** ANÁLISE DE DETEÇÃO: rlnn20tt.OUT - CURVA ROC  
 Arquivo de saída: rlnn20tt.  
 Rede (30-20-1) - Classificação de todos os registros (Grupo de Treino  
 + Grupo de Teste).

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
0.001	70	213	97	2	74.1	97.2	68.7	83.0	41.9	99.1	100.56
0.010	70	266	44	2	88.0	97.2	85.8	91.5	61.4	99.3	188.44
0.050	69	287	23	3	93.2	95.8	92.6	94.2	75.0	99.0	245.00
0.100	69	296	14	3	95.5	95.8	95.5	95.7	83.1	99.0	281.15
0.200	69	300	10	3	96.6	95.8	96.8	96.3	87.3	99.0	299.86
0.300	69	301	9	3	96.9	95.8	97.1	96.5	88.5	99.0	304.84
0.400	69	301	9	3	96.9	95.8	97.1	96.5	88.5	99.0	304.84
0.500	69	301	9	3	96.9	95.8	97.1	96.5	88.5	99.0	304.84
0.600	69	301	9	3	96.9	95.8	97.1	96.5	88.5	99.0	304.84
0.700	69	301	9	3	96.9	95.8	97.1	96.5	88.5	99.0	304.84
0.800	69	301	9	3	96.9	95.8	97.1	96.5	88.5	99.0	304.84
0.900	67	302	8	5	96.6	93.1	97.4	95.2	89.3	98.4	297.41
0.950	64	302	8	8	95.8	88.9	97.4	93.2	88.9	97.4	278.94
0.990	46	305	5	26	91.9	63.9	98.4	81.1	90.2	92.1	190.54
0.999	0	310	0	72	81.2	0.0	100.0	50.0	0.0	81.2	0.00



**Tabela 9.15:** ANÁLISE DE DETEÇÃO: APACHE II - CURVA ROC - (Grupo de Treino + Grupo de Teste)

Classificação de todos os registros (Grupo de Treino + Grupo de Teste).

Nivel	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
10%	84	81	204	13	43,19	86,59	28,42	57,50	70,83	86,17	8,79
12%	76	131	154	21	54,18	78,35	45,96	62,15	66,95	86,18	17,86
14%	74	179	106	23	66,23	76,28	62,80	69,54	58,88	88,61	44,39
16%	65	215	70	32	73,29	67,01	75,43	71,22	51,85	87,04	57,06
18%	51	237	48	46	75,39	52,57	83,15	67,86	48,48	83,74	48,13
20%	45	257	28	52	79,05	46,39	90,17	68,28	38,35	83,17	62,60
22%	33	267	18	64	78,53	34,02	93,68	63,85	35,29	80,66	48,01
24%	25	270	15	72	77,22	25,77	94,73	60,25	37,50	78,94	32,47
26%	19	274	11	78	76,70	19,58	96,14	57,86	36,66	77,84	24,73
28%	11	276	9	86	75,13	11,34	96,84	54,09	45,00	76,24	9,76
30%	8	282	3	89	75,91	8,24	98,94	53,59	27,27	76,01	13,39
32%	5	283	2	92	75,39	5,15	99,29	52,22	28,57	75,46	7,97
34%	2	285	0	95	75,13	2,06	100,00	51,03	0,00	75,00	5,90
36%	1	285	0	96	74,86	1,03	100,00	50,51	0,00	74,80	2,94

**Tabela 9.16:** ANÁLISE DE DETEÇÃO: APACHE II - CURVA ROC - Classificação do Grupo de Teste.

Nível	PV	NV	PF	NF	Acur	Sens	Espec	S+E	Pred+	Pred-	X2
10%	25	26	46	3	51,00	89,28	36,11	62,69	64,78	89,65	6,31
12%	22	40	32	6	62,00	78,57	55,55	67,06	59,25	86,95	9,45
14%	22	52	20	6	74,00	78,57	72,22	75,39	47,61	89,65	21,35
16%	20	59	13	8	79,00	71,42	81,94	76,68	39,39	88,06	25,97
18%	18	63	9	10	81,00	64,28	87,50	75,89	33,33	86,30	27,43
20%	15	67	5	13	82,00	53,57	93,05	73,31	25,00	83,75	27,39
22%	10	69	3	18	79,00	35,71	95,83	65,77	23,07	79,31	17,74
24%	6	70	2	22	76,00	21,42	97,22	59,32	25,00	76,08	9,52
26%	3	71	1	25	74,00	10,71	98,61	54,66	25,00	73,95	4,56
28%	1	72	0	27	73,00	3,57	100,00	51,78	0,00	72,72	2,59
30%	1	72	0	27	73,00	3,57	100,00	51,78	0,00	72,72	2,59
32%	1	72	0	27	73,00	3,57	100,00	51,78	0,00	72,72	2,59
34%	1	72	0	27	73,00	3,57	100,00	51,78	0,00	72,72	2,59
36%	1	72	0	27	73,00	3,57	100,00	51,78	0,00	72,72	2,59