

UM ESTUDO COMPARATIVO DE APLICAÇÃO DE REDE NEURAL ARTIFICIAL CONSTRUTIVA E REGRESSÃO MÚLTIPLA NA ESTIMAÇÃO DO ÍNDICE DE MASSA VENTRICULAR ESQUERDA

Domingos Vanderlei Filho, Fernando Menezes Campello de Souza

Universidade Federal de Pernambuco
Departamento de Engenharia Eletrônica e de Sistemas, Rua Acadêmico Hélio Ramos, s/n -
Cidade Universitária, Recife – PE 50740-530
dvanderlei@yahoo.com.br, fmcs@hotmail.com.br

Hilton de Castro Chaves Júnior

Universidade Federal de Pernambuco
Departamento de Cardiologia do Hospital das Clínicas da UFPE, Av. Prof. Moraes Rego, s/n -
Cidade Universitária, Recife – PE 50670-420
hchavesjr@uol.com.br

Mêuser Jorge Silva Valença

Universidade Salgado de Oliveira
Av. Mascarenhas de Moraes, 1919 – Imbiribeira, Recife – PE 51150-001
meuserv@yahoo.com.br

Resumo

Segundo a OMS as doenças coronarianas levam a morte mais de 7 milhões de pessoas por ano no mundo, na sua maioria em países em desenvolvimento como o Brasil. A hipertensão é apontada como uma das mais importantes causas de mortes prematuras no mundo que podem ser prevenidas, e contribui com cerca de 50% das doenças cardiovasculares. Considerando que as redes neurais artificiais (RNAs) podem realizar modelagem estatística não linear, este trabalho comparou experimentalmente o desempenho entre as abordagens com RNA auto construtiva e com regressão múltipla na estimação da variável resposta: índice de massa ventricular esquerda, a qual permite quantificar o grau de hipertrofia ventricular esquerda. Os resultados obtidos com as duas técnicas, comparados em termos de coeficiente de determinação, mostraram que a RNA pode ser usada como uma alternativa viável no apoio ao diagnóstico médico, além de encorajar investimentos em melhorias no sistema ora desenvolvido.

Palavras-Chaves: Redes Neurais Artificiais; Regressão Múltipla; Apoio ao Diagnóstico Médico; Cardiologia.

Abstract

According the World Health Organization the coronary heart disease kills more than 7 million people each year. Most of these deaths are in developing countries. High blood pressure contributes to around half of all cardiovascular disease. Artificial neural networks (ANN) apply non-linear statistics to classification problems. One such problem in the field of cardiology is to quantify the left ventricular hypertrophy (LVH). The objective of this work is to develop a model based on self constructed artificial neural networks to support a medical diagnosis. The results of the system were compared to those obtained with a classical approach used in medical area multiple regression. The results obtained with the two techniques in terms of coefficient of determination indicate that ANNs can be considered a

feasible approach to support a medical diagnosis and encourages future improvements.

Keywords: Artificial Neural Networks; Multiple Regression; Diagnosis Support System; Cardiology.

1. INTRODUÇÃO

O Ministério da Saúde (MS), no anuário estatístico de 2001 informa que as doenças do aparelho circulatório representam a principal causa de óbitos no país (32%) e em todas as regiões, seguidas pelas causas externas (15%), neoplasias (15%) e doenças do aparelho respiratório (11%) [1-2]. Outros trabalhos publicados também mostram e reforçam as conseqüências das doenças relacionadas ao Sistema Cardiovascular (SVC) na sociedade brasileira [3-6]

Analisando-se os dados apresentados no parágrafo anterior é possível perceber-se a importância social da área de cardiologia e abre-se a possibilidade de conceber e desenvolver modelos, técnicas e ferramentas para cooperar na busca de soluções para os problemas de diagnóstico e tratamento de tais doenças.

As RNA's e outras ferramentas de inteligência artificial têm sido aplicadas em diversas áreas como uma alternativa ao uso de técnicas clássicas de estatística, como exemplo a regressão múltipla. Em particular no caso da área de saúde é possível encontrar aplicações no apoio a decisão e diagnóstico médico [7-11].

Neste trabalho buscou-se principalmente mostrar de forma experimental que as RNA's em especial as redes neurais construtivas do tipo NSRBN podem representar uma alternativa bastante viável na estimação do IMVE em relação a regressão múltipla [12].

Este artigo está organizado em 7 seções. A seção 2 apresenta a regressão múltipla. Na seção 3 é apresentada a rede neural construtiva NSRBN. Na seção 4 apresenta-se a base de dados. A metodologia utilizada na realização dos experimentos é mostrada na seção 5. Os resultados obtidos com as duas técnicas são descritos na seção 6. A seção 7 apresenta as considerações finais e possibilidades de trabalhos futuros.

2. REGRESSÃO MÚLTIPLA

No modelo linear de regressão múltipla [13], a relação entre a variável dependente e as variáveis explicativas assume a forma:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

Onde a variável resposta Y está numa escala contínua (de razão). Embora variáveis explicativas categóricas possam ser usadas também na regressão múltipla, tipicamente, as variáveis independentes são também contínuas. O método de estimação dos parâmetros utilizado é o dos mínimos quadráticos. Podem ser usadas transformações nas variáveis explicativas, e o modelo apresenta a forma:

$$Y = \beta_0 + \beta_1 f_1(X_1) + \beta_2 f_2(X_2) + \dots + \beta_p f_p(X_p) + \varepsilon \quad (2)$$

As transformações introduzidas servem para garantir a linearidade. Para um modelo clássico, é imperioso que a variável Y seja gaussiana, isto cria a necessidade em alguns casos de também fazer-se uma transformação na variável resposta de forma a ser garantida esta condição. Um tipo de transformação realizada para este fim é através da aplicação do logaritmo neperiano.

Este modelo permite fazer a estimação do valor da variável resposta em função dos valores das variáveis explicativas, apresentando-se o resultado em termos de uma média e um intervalo de confiança ao redor deste valor. O modelo é considerado ajustado quando os testes estatísticos sobre os valores estimados dos parâmetros forem significativos, t de Student, e a análise de variância também for significativa, teste F. Este modelo é considerado tanto melhor quanto maior for o coeficiente de determinação, o qual expressa o percentual da variabilidade

observada na variável resposta que é explicado pela equação da regressão. Assim, uma vez atendida a condição de ser estatisticamente significativa, o modelo será tanto mais preciso na estimação do valor da variável resposta para um dado indivíduo, ou seja, mais estreito o intervalo de confiança ao redor do valor estimado, quanto maior for o seu coeficiente de determinação.

O que é desejado com a utilização de modelos de regressão múltipla é identificar o máximo de fatores explicativos, que possam esclarecer a variabilidade da resposta estimada empiricamente, procurando-se deixar o mínimo de variabilidade atribuída ao erro. Esta variabilidade imputada ao erro é chamada variabilidade não-explicada pela equação de regressão ou explicada pela variância.

3. REDE NEURAL CONSTRUTIVA

Neste trabalho foram utilizadas RNAs auto-organizáveis que caracterizam-se por utilizarem uma metodologia híbrida, redes neurais e conceitos estatísticos, com o objetivo de estimar RNAs de arquitetura ótima com uma estrutura que evolui durante o processo de treinamento. A classe de RNAs *feedforward* de alta ordem, chamadas de redes neurais por combinação de blocos de regressões sigmóides não-lineares (NSRBN), são capazes de aproximar qualquer função contínua definida sobre um conjunto compacto [12]. Essas RNAs têm como fundamento a auto-organização, além de utilizar um aprendizado incremental baseado em um algoritmo construtivo. Este algoritmo construtivo é responsável pela escolha do modelo de ótima complexidade, ou seja, pela definição da arquitetura da rede. As NSRBN podem ser vistas como aproximações polinomiais multivariadas, uma vez que elas têm como elemento de construção básico polinômios homogêneos de grau k , nos quais k varia de 1 a N (complexidade máxima, ou seja, máximo grau polinomial).

A arquitetura de uma RNA, de acordo com a possibilidade de mudanças no seu tamanho durante o processo de treinamento, pode ser classificada em estática ou dinâmica. Uma rede é dita de arquitetura estática quando o número de camadas e unidades de processamento permanece constante desde a concepção da RNA até a finalização do seu treinamento. Por outro lado, uma RNA de arquitetura dinâmica tem como característica principal um processo de mudanças no tamanho durante o treinamento. Nas NSRBN um algoritmo construtivo é utilizado para incrementar a estrutura da rede, partindo-se de uma pequena estrutura, até que se atinja uma estrutura de tamanho ótimo, segundo critérios estatísticos de seleção adequados.

O modelo NSRBN a rede é construída por blocos de polinômios homogêneos, utilizando uma função de ativação adequada (tangente hiperbólica, por exemplo), para os neurônios escondidos. Quanto aos neurônios de saída, estes têm uma ativação mais complexa, uma vez que realizam uma regressão logística não-linear com relação à saída dos neurônios escondidos. Desta forma, essas redes NSRBN podem aproximar uniformemente qualquer função contínua definida em um conjunto compacto limitado e também lidar com a realização de funções booleanas, o que não é possível com os algoritmos tradicionais do grupo GMDH – *Group Method Of Data Handling* (combinatorial e *multilayer*) [12].

O primeiro passo para a construção de um modelo combinatorial consiste em dividir a função $f(x)$ em blocos compostos de termos de igual grau, como apresentado na Figura 1.

$$f(x) = f_1(x) + f_2(x) + \dots + f_d(x) \quad (3)$$

O bloco de grau p ($p=1, 2, \dots, d$), de acordo com o polinômio de Kolmogorov-Gabor (teorema de Weierstrass), pode ser expresso como uma soma de produtos das entradas, isto é:

$$f_p(x) = \sum_{i_1=1}^N \sum_{i_2=i_1}^N \dots \sum_{i_N}^N a_p(i_1, i_2, i_3, \dots, i_N) \cdot x_1^{qp1(i_1)} \cdot x_2^{qp2(i_2)} \cdot \dots \cdot x_N^{qpN(i_N)} \quad (4)$$

em que, $q_{p1}(i_1) + q_{p2}(i_2) + q_{p3}(i_3) + \dots + q_{pN}(i_N) = p$
 e $p \geq q_{pm}(\cdot) \geq 0$ para $m=1,2,3,\dots,N$.

Tem-se como objetivo realizar o mapeamento de cada um dos polinômios homogêneos de grau p .

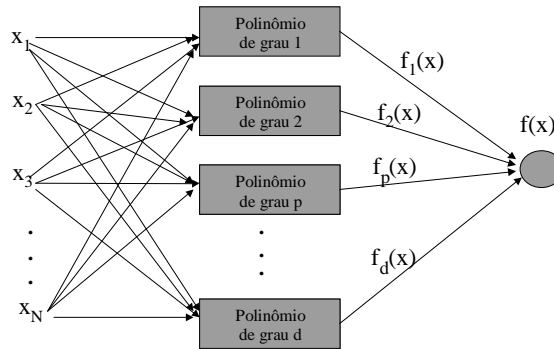


Figura 1: Arquitetura combinatorial

Uma forma simples seria realizar todos os produtos das variáveis de entrada procurando aproximar diretamente a hypersuperfície não-linear, entretanto isto não é prático, a menos que d e N sejam relativamente pequenos (modelo combinatorial). Logo, para se ter um modelo eficiente, faz-se necessário que ele possua funções de ativação que sejam capazes de realizar vários termos dos polinômios homogêneos de forma otimizada, isto é, sem que ocorra um crescimento explosivo no número de parâmetros necessários para ajustar a função desejada. Para esse objetivo, as NSRBN têm por base o algoritmo combinatorial (composta pela soma de blocos) com uma estrutura de blocos similar a uma rede *Multilayer Perceptron* (MLP), sendo que, diferentemente destas, os neurônios de saída realizam a técnica estatística de regressão logística não-linear [14], com relação aos neurônios escondidos.

Definição 1: uma rede NSRBN é uma rede combinatorial composta pela soma de p blocos ($p=1, \dots, d$) com estrutura similar a uma MLP, na qual as unidades escondidas desses blocos realizam uma regressão sigmóide linear das entradas (tal qual uma MLP), e as unidades de saída realizam uma regressão do tipo sigmóide não-linear das unidades escondidas (tal qual uma regressão logística não-linear). Essa nova estrutura formada é um aproximador universal com forte capacidade de mapeamento não-linear. Utiliza-se aqui o termo sigmóide, significando qualquer função sigmoidal, como uma generalização da regressão logística clássica.

A estrutura de cada bloco é composta de apenas uma camada escondida, tendo Nh unidades escondidas em cada bloco com função de ativação do tipo [14]:

$$f(net(h)) = (\sigma_{net(h)} + \theta_h) \tag{5}$$

em que $net(h)$ é a entrada líquida para cada neurônio escondido, h é um número inteiro que representa as unidades escondidas em cada bloco (variando de 1 até a ordem p do bloco), θ_h é a tendência e $\sigma_{net(h)}$ é, por exemplo, a função tangente hiperbólica dadas por:

$$\sigma_{net(h)} = \frac{e^{net(h)} - e^{-net(h)}}{e^{net(h)} + e^{-net(h)}} \tag{6}$$

Poder-se-ia utilizar outras funções, como é o caso da função recíproca representada na Equação 8, isto é,

$$f_r(x) = \frac{net(h)}{1 + |net(h)|} \tag{7}$$

O valor de p correspondente ao grau do bloco é quem vai definir o número de neurônios em cada bloco, isto é: o bloco de grau 1 ($p=1$) tem $h=1$, logo um neurônio escondido; o bloco de grau 2 ($p=2$) terá 2 neurônios escondidos ($h=1$ e $h=2$) e assim sucessivamente, de tal forma que o bloco de grau p terá $h=p$ e portanto p neurônios

escondidos.

A estrutura proposta para cada bloco é similar à da Figura 2, na qual apresenta-se o mapeamento para um bloco de grau p (a ordem dos neurônios escondidos h variará de 1 até p), para i variáveis de entrada (onde $i = 1, \dots, N$) e uma única unidade de saída.

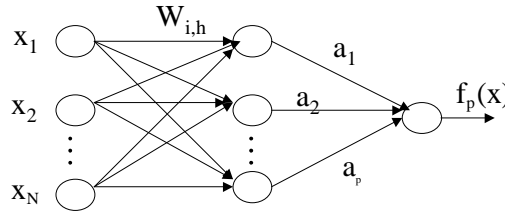


Figura 2: Arquitetura de um bloco polinomial

Esse bloco é uma rede *feedforward* completamente conectada com uma única camada escondida. A diferença básica desse bloco de grau p para uma rede MLP é que os neurônios de saída são expressos como uma regressão sigmóide não-linear das unidades escondidas.

Seja $x = [1, x_1, x_2, x_3, \dots, x_N]^T$ o vetor coluna das variáveis de entrada, no qual o elemento x_i denota um elemento qualquer dele. As entradas são ponderadas pelos vetores de pesos $\mathbf{w} = [w_{0h}, w_{1h}, w_{2h}, \dots, w_{Nh}]^T$, $h=1, \dots, p$ nos quais p é a ordem do bloco. Portanto, utilizando-se função de ativação estabelecida na Equação 5, a saída de cada unidade escondida, $f(\text{net}(h))$, desse bloco de grau p pode ser escrito como:

$$f(\text{net}(h)) = (\sigma_{\text{net}(h)} + \theta_h) \tag{8}$$

Em que: $\text{net}(h) = \mathbf{w}_h^T \mathbf{x}$, ou seja, $\text{net}(h) = \sum_{i=1}^N w_{ih}x_i + w_{0h}$, $h=1, 2, \dots, p$.

Logo, a unidade de saída (que realiza uma regressão logística não-linear) desse bloco pode ser representada pela seguinte equação:

$$f_p(x) = a_1(f(\text{net}(1)))^1 + a_2(f(\text{net}(2)))^2 + a_3(f(\text{net}(3)))^3 + \dots + a_p(f(\text{net}(p)))^p \tag{9}$$

sendo $a = [a_1, a_2, a_3, \dots, a_p]^T$ o vetor dos pesos das unidades escondidas para a unidade de saída.

Fazendo-se uso da Equação 8, a função $f_p(x)$ pode ser escrita da seguinte forma:

$$f_p(x) = a_1(\sigma_{\text{net}(1)} + \theta_1) + a_2(\sigma_{\text{net}(2)} + \theta_2)^2 + \dots + a_p(\sigma_{\text{net}(p)} + \theta_p)^p \tag{10}$$

ou

$$f_p(x) = \sum_{h=1}^p a_h(\sigma_{\text{net}(h)} + \theta_h)^h \tag{11}$$

com h representando a ordem dos neurônios na camada escondida, ($h=1, \dots, p$), do bloco de grau p . A arquitetura de uma rede NSRBN (Figura 1) é composta pela soma de blocos com grau variando de 1 a d . Aqui, está sendo utilizado o bloco p (Figura 2) para representar qualquer um desses blocos. Portanto, usando-se esta arquitetura, uma função desconhecida f em R^N pode ser aproximada por uma rede NSRBN, através da seguinte expressão:

$$f(x) = \sigma_{\text{net}(0)} \left(\sum_{p=1}^d f_p(x) \right) \tag{12}$$

em que: $\sigma_{\text{net}(0)}$ uma função de ativação adequada para a unidade de saída.

Por outro lado, a construção das redes NSRBN é feita de forma incremental, de maneira que pode-se representa-la em uma forma similar a de um bloco (Equação 10), conforme apresenta-se a seguir. Considere o passo do algoritmo para o qual se se adiciona à rede o bloco de ordem d , f_d , tendo-se anteriormente adicionado os blocos f_0, f_1, \dots, f_{d-1} . Fazendo-se $f_0 \equiv 0$, uma função f pode ser representada por:

$$f(x) = \sigma_{\text{net}(0)} \left(\sum_{p=0}^{d-1} f_p(x) + f_d(x) \right) \tag{13}$$

na qual os pesos dos blocos f_0, f_1, \dots, f_{d-1} são congelados ao adicionar-se o bloco de ordem d , f_d .

Portanto, pode-se escrever f_x da seguinte forma:

$$f(x) = \sigma_{net(o)} \{ c_0 + c_1(\sigma_{net(i)} + \theta_1) + c_2(\sigma_{net(2)} + \theta_2)^2 + \dots + c_d(\sigma_{net(d)} + \theta_d)^d \} \quad (14)$$

na qual o vetor $\mathbf{c} = [c_1, c_2, \dots, c_d]$ representa os pesos das unidades escondidas para as unidades de saída e a constante C_0 está representando todos os blocos congelados até a ordem $d-1$. A Equação 11 é, portanto, idêntica a uma regressão logística não-linear quando a função de ativação utilizada é a sigmóide logística tradicional.

A potencialidade dessas redes está fundamentada nas características impostas pela regressão sigmóide não-linear na unidade de saída. Esse algoritmo construtivo tem como base um método de aprendizagem construtiva (também chamado de auto-organização na teoria das redes GMDH). O princípio construtivo pode ser formulado como segue: quando um modelo incrementa gradualmente sua complexidade, certos critérios, que são chamados de critérios de seleção ou função objetivo e que têm a propriedade de complemento externo, passam através de um valor mínimo global. O encontro deste mínimo indicará a existência de um modelo de complexidade ótima.

Portanto, usando-se essa arquitetura proposta uma função qualquer $f(X)$ pode ser aproximada pelo uso direto de uma rede NSRBN por dividir $f(X)$ em blocos homogêneos de ordem 1 até o bloco de ordem d . Dessa forma, a função total $f(X)$ (polinômio completo) pode ser aproximada pela soma das funções parciais $f_p(X)$, $p=1, \dots, d$ (polinômios homogêneos), de maneira incremental podendo ser representada matematicamente por:

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + \hat{f}_p(x) + \dots + f_d(x)) \quad (15)$$

O modelo proposto tem como princípio manter congelados os pesos de todos os blocos anteriores ao atualmente otimizado, de tal forma que o número de pesos a ser otimizado a cada grau de complexidade não aumente explosivamente, como no algoritmo combinatorial tradicional. O algoritmo de abordagem construtiva procede como se descreve a seguir. Seja p um passo do algoritmo para qual o bloco é adicionado à rede, então uma função qualquer $f(X)$ é sucessivamente aproximada por:

$$f(x) = g(f_1(x)) \quad (16)$$

$$f(x) = g(\hat{f}_1(x) + f_2(x)) \quad (17)$$

$$f(x) = g(\hat{f}_1(x) + \hat{f}_2(x) + \dots + f_p(x)) \quad (18)$$

em que os pesos dos blocos $\hat{f}_{p-1}(x)$ são congelados quando o bloco de ordem p é adicionado à rede e g representa uma função de ativação adequada para a unidade de saída.

4. BASE DE DADOS

Os experimentos foram realizados utilizando-se a base de dados gerada por Chaves Júnior [15]. A base de dados final ficou composta por 101 indivíduos de ambos os sexos, com idades entre 21 e 81 anos, 27% de normotensos e 73% de hipertensos leves, moderados e graves, todos matriculados na Clínica de Hipertensão do Serviço de Cardiologia do Hospital das Clínicas da Universidade Federal de Pernambuco. Cerca de metade dos indivíduos (48) era portadora de hipertrofia ventricular esquerda concêntrica ao ecocardiograma.

O conjunto de dados incluiu variáveis antropométricas, biográficas, medidas de pressão e frequência cardíaca em consultório, série bioquímica, hormônios, resultados da monitoração ambulatorial da pressão arterial de 24h (MAPA-24h), dados do ecocardiograma e do eletrocardiograma. Os valores correspondentes a variável de saída IMVE foram obtidos pelo resultado do ecocardiograma (padrão ouro). O conjunto completo contém mais de 2.000 diferentes variáveis. Um estatística descritiva para uma pequena parte deste conjunto é

apresentada a título de exemplo na Tabela 1.

Tabela 1- Estatística Descritiva Parcial das Variáveis

VARIÁVEL	MÉDIA	MEDIANA	MÍNIMO	MÁXIMO	DESV. PADRÃO
Sexo (1-Masculino, 0-Feminino)	0,48	0	0	1	0,502
Idade (anos)	53,18	53	21	81	15,096
Peso (Kg)	73,24	73	48	115	14,79
Altura (m)	1,64	1,64	1,4	1,85	0,098
IMC (Kg/m)	27,27	26,99	18,07	40,31	4,43
Cintura (cm)	93,12	93	70	121	11,82
Quadril (cm)	101,01	99	84	149	10,873
Cintura/Quadril	0,92	0,92	0,77	1,07	0,072
Hipertenso	0,73	1	0	1	0,445
Pressão Sistól. Média Consultório (mmHg)	147,31	144	110	221,3	21,147
Pressão Diast. Média Consultório (mmHg)	89,75	88	58	131,3	13,875
Freq. Cardíaca Média Consultório (bpm)	77,4	77,33	49	113,7	14,367
Colesterol Total (mg / dl)	191,31	186	85	387	48,457
Pressão Sistólica Méd. Sono (mmHg)	126,7	124,5	95,7	177,5	15,76
Índice Eletrocardiográfico de Cornell (Contínuo)	19,25	18,5	4	53	8,239
Índice de Massa do Ventrículo Esquerdo (g/m ²)	121,08	106,9	62	251,5	40,14
Segmento QT (s)	0,334	0,36	0,12	0,6	0,091
Pressão Arterial Média (mmHg)	99,47	98,69	81,23	128,9	10,204
Pressão de Pulso (mmHg)	52,62	50,34	33,43	89,79	10,492

5. EXPERIMENTOS

Modelos usando redes NSRBN foram construídos, treinados e validados. Os resultados obtidos com estes modelos conexionistas foram comparados com outros gerados com a abordagem clássica estatística. As medidas conseguidas com o ecocardiograma foram utilizadas como padrão ouro.

Várias combinações das variáveis de entrada foram testadas e analisadas para permitir selecionar os modelos que forneceram melhores respostas em termos de coeficiente de determinação. Em todos os casos as redes tiveram como variável de saída os valores do índice de massa do ventrículo esquerdo. Outros critérios que foram considerados na seleção dos modelos de redes foram: escolha de variáveis de fácil obtenção ou de baixo custo financeiro; que apresentaram baixo custo computacional e complexidade em termos de menor número de blocos e entradas.

A análise de desempenho foi realizada utilizando-se um *software* simulador de redes NSRBN [16], o qual permitiu escolher e testar uma grande variedade de:

- Tamanhos dos conjuntos de treinamento, validação e teste;
- Valores dos parâmetros como taxa de aprendizado, momento e semente;
- Critério de parada para o treinamento.

Para cada simulação realizada estava disponível o mínimo erro médio quadrático (EMQ) no conjunto de validação, o erro percentual absoluto médio e o erro padrão no conjunto de teste. O coeficiente de determinação foi calculado através de um procedimento externo ao simulador.

Os experimentos envolvendo a regressão múltipla foram conduzidos como descrito em Chaves Júnior [15]. Ele estudou os seguintes aspectos: as variáveis que fortalecem o modelo; o comportamento das variáveis selecionadas; a importância relativa das variáveis

isoladas e combinadas e a estimativa do erro.

6. RESULTADOS OBTIDOS

Nesta seção estão apresentados os modelos de redes NSRBN e Regressão Múltipla que apresentaram o melhor desempenho conforme os critérios citados na seção 5.

6.1. MODELOS DE REDES NEURAIS

A arquitetura final da rede NSRBN para o melhor modelo usando 3 variáveis de entrada (chamado Modelo A) está mostrado na Figura 3, ele foi selecionado pelo conjunto de validação cruzada.

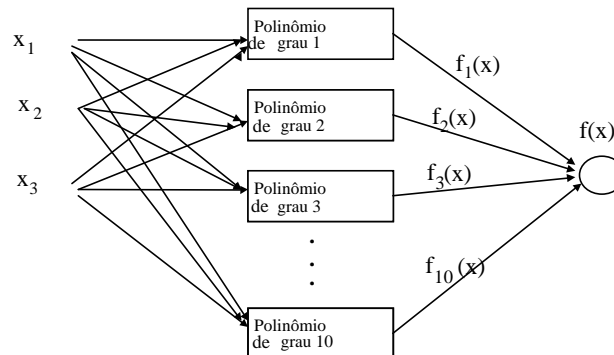


Figura 3 - Arquitetura NSRBN (Modelo A)

O Modelo A trabalhou com três variáveis: Idade, extensão do segmento QT do eletrocardiograma e a média do índice pulsátil da pressão arterial no período do sono. Esta última variável foi obtida a partir dos registros da MAPA-24h. A Figura 4 mostra a curva do EMQ para o conjunto de validação cruzada e a Figura 5 apresenta as curvas com os valores esperados e os obtidos com o Modelo A. O valor do Coeficiente de Determinação para este modelo foi de 66,37.



Figura 4 – Curva do EMQ no conjunto de validação (Modelo A)

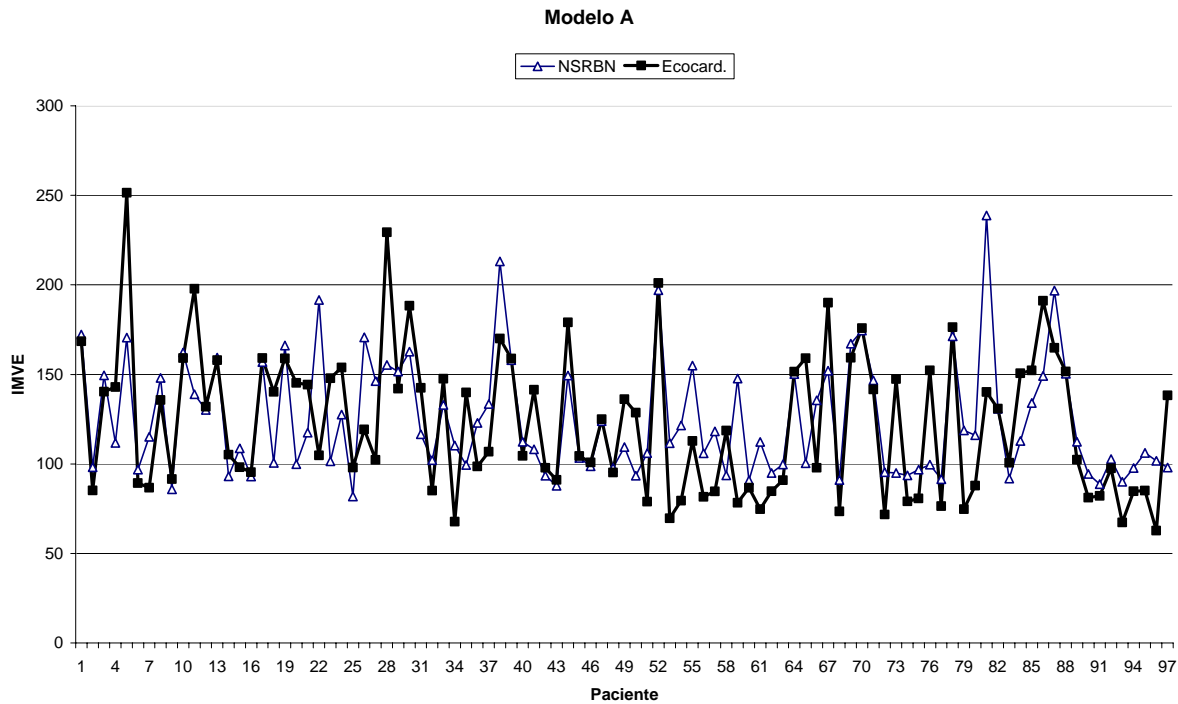


Figura 5 – Curvas dos valores esperados vs. obtidos (Modelo A)

Um segundo modelo foi obtido utilizando-se quatro variáveis de entrada. O melhor modelo foi obtido a partir dos erros no conjunto de validação cruzada e foi denominado Modelo B.

A arquitetura da rede para o Modelo B está apresentada na Figura 6.

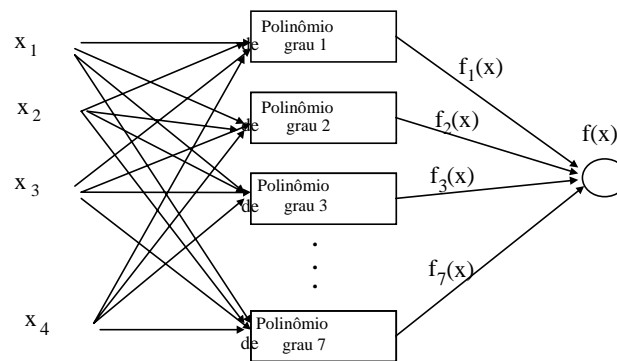


Figura 6 - Arquitetura NSRBN (Modelo B)

As variáveis que compuseram o modelo foram: Idade, a relação cintura/quadril, o segmento QT do eletrocardiograma e a média do índice pulsátil da pressão arterial no período do sono. A Figura 7 mostra a curva do EMQ para o conjunto de validação cruzada e a Figura 8 apresenta as curvas com os valores esperados e os obtidos com o Modelo B. O valor do Coeficiente de Determinação para este modelo foi de 85,82.



Figura 7 – Curva do EMQ no conjunto de validação (Modelo B)

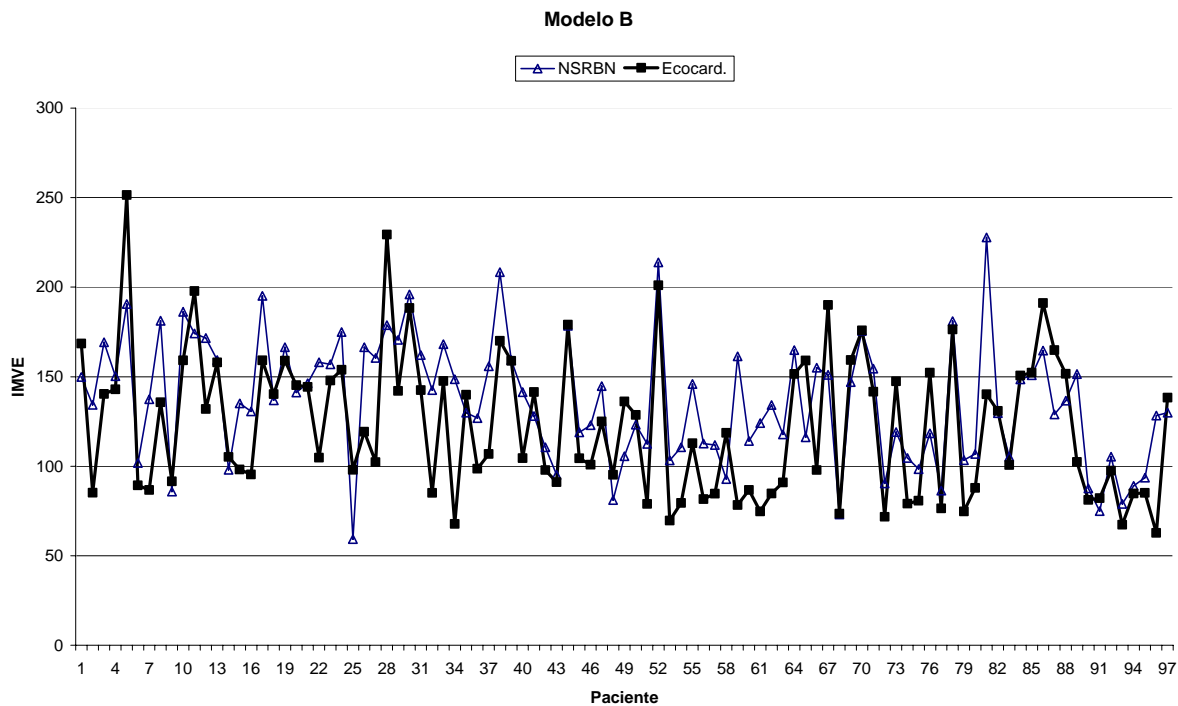


Figura 8 – Curvas dos valores esperados vs. obtidos (Modelo B)

6.2. MODELO REGRESSÃO MÚLTIPLA

As variáveis que compuseram o melhor modelo com a regressão múltipla (Modelo C) foram: Idade, Relação cintura/quadril (Cintquad), valores contínuos do índice eletro cardiográfico de Cornell (Cornellc) e a pressão sistólica média do sono (Pssmed). Elas explicaram 47% da variabilidade do logaritmo neperiano do índice de massa ventricular esquerda (coeficiente de determinação igual a 0,47). A expressão analítica do modelo está na Equação 19.

$$IMVE = \exp[2,6535 + 0,0065(Idade) + 0,8247(Cintquad) + 0,0080(Cornellc) + 0,0065(Pssmed)] \quad (19)$$

A Figura 9 apresenta as curvas com os valores esperados e os obtidos com o Modelo C.

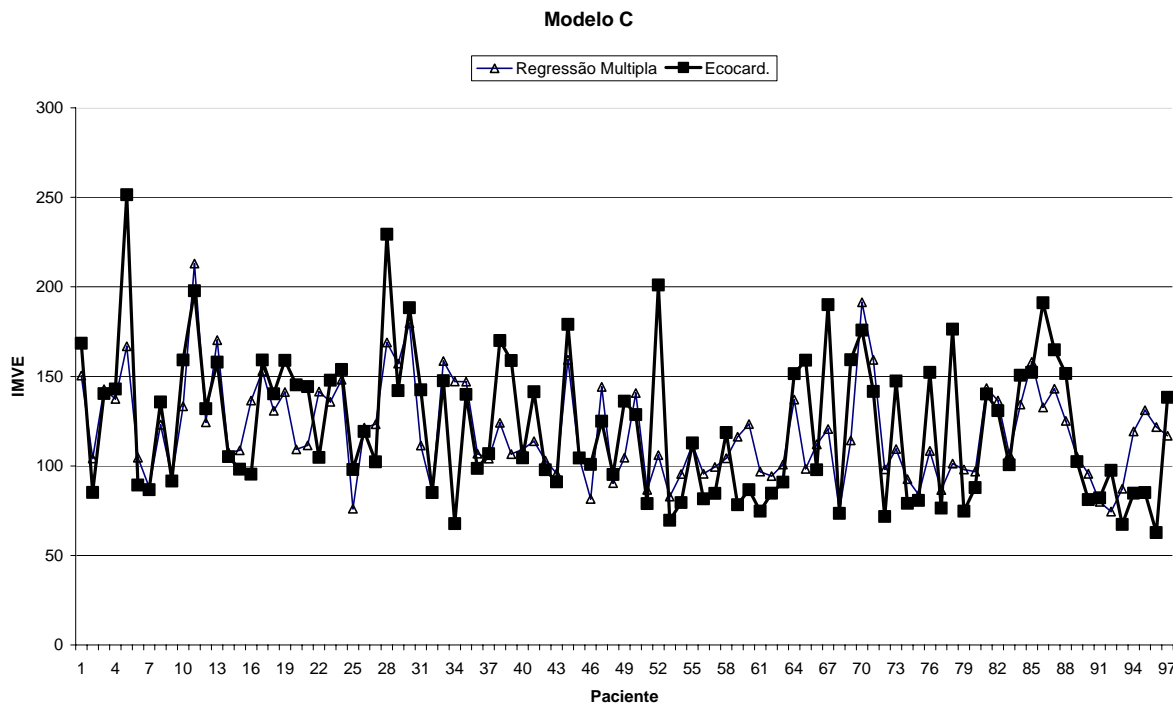


Figura 9 – Curvas dos valores esperados vs. obtidos (Modelo C)

7. CONCLUSÕES

Os modelos A e B, de três e quatro variáveis de entrada respectivamente, apresentaram resultados superiores do obtido a partir da regressão múltipla. A afirmação anterior considerou que em todos os três casos as variáveis utilizadas têm um custo aproximado, pois além de dados obtidos durante a consulta médica (idade e relação cintura/quadril) exigem mais dois exames complementares: eletrocardiograma e MAPA-24h.

Considerando a pequena amostra disponível para treinamento da rede neural, os resultados obtidos até agora indicam que o sistema proposto tem potencial para reduzir ou ainda justificar melhor a prescrição de exames e procedimentos de custo elevado na área de cardiologia. Isto é reforçado pelas características de das redes de detectar relações não lineares de alta ordem entre as variáveis além de sua robustez para tratar dados incompletos.

Contudo, mais experimentos são necessários antes que estes modelos de redes possam ser aceitos para suporte a decisão médica. Percebe-se a necessidade de continuação deste trabalho para validar o modelo neural em uma amostra maior de indivíduos.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Ministério da Saúde. Anuário estatístico de saúde do Brasil 2001. Disponível em <http://portal.saude.gov.br/saude/aplicacoes/anuario2001/index.cfm>. Capturado em março 2003.
- [2] Ministério da Saúde. Indicadores e Dados Básicos – Brasil – 2003. Disponível em <http://tabnet.datasus.gov.br/cgi/idb2003/matriz.htm>. Último acesso em 04/04/2005.Segunda referência.
- [3] A.P. Mansur, D. Favarato, M.F. De Souza, S.D. Avakian, J.M. Aldrighi, L.A.M. César, J.A.F. Ramires, “Tendência do Risco de Morte por Doenças Circulatórias no Brasil de 1979 a 1996”, Arq Bras Cardiol, vol. 76, No.6, p.497-503, 2001.

- [4] M.F. De Souza, A. Timerman, C.V. Serrano Júnior, R.D. Santos, A.P. Mansur, “Tendências do Risco de Morte por Doenças Circulatórias nas Cinco Regiões do Brasil de 1979 a 1996”. *Arq Bras Cardiol*, vol. 77, No.6, p.562-8, 2001.
- [5] A.P. Mansur, M.F. De Souza, A. Timerman, J.A.F. Ramires, “Tendência do Risco de Morte por Doenças Circulatórias, Cerebrovasculares e Isquêmicas do Coração em 11 Capitais do Brasil de 1980 a 1998”. *Arq Bras Cardiol*, vol. 79, No.3, p.269-76, 2002.
- [6] I. Lessa, “Assistência Médica e Óbitos por Doença Arterial Coronariana no Brasil, 1980-1999”. *Arq Bras Cardiol*, vol. 81, No.4, p.329-35, 2003.
- [7] W.G. Baxt, “Application of artificial neural networks to clinical medicine”, *Lancet*. 1995; 346(8983):1135-8.
- [8] D. Vanderlei Filho, M.A. Dos Santos, T.B. Ludermit, M.J.S. Valença, “Apoio ao Diagnóstico de LER/DORT Utilizando um Modelo Fuzzy”. VII Brazilian Symposium on Neural Networks. Ipojuca-PE, novembro
- [9] D. Vanderlei Filho, A.F.D. Silva, M.J.S. Valença, T.B. Ludermit, “Gerenciamento da Fila Única para Transplante de Órgãos com Abordagem *Fuzzy*.” LAPTEC'2002 - III Congress of Logic Applied to Technology. São Paulo-SP, novembro.
- [10] C.R.S. Lopes, T.B. Ludermit, M.P. De Souto, A.B. Ludermit, “Neural Networks for the analysis of Common Mental Disorders Factors”, Proceedings of the Brazilian Symposium on Neural Networks, SBRN-2002, Ipojuca-PE, novembro 2002.
- [11] A. Wolf, et al., “Localização das Ondas Componentes de Sinais Eletrocardiográficos Usando Redes Neurais Artificiais.” Anais do VI Brazilian Symposium on Intelligent Automation, SBAI-2003, Bauru, SP, Brasil.
- [12] M.J.S. Valença, “Análise e projeto de redes neurais com abordagem construtiva para modelagem de sistemas complexos”, Tese, UFPE, Brasil, 1999.
- [13] Draper, Smith, *Applied regression Analysis*, 1981, Ed. John Wiley and Sons.
- [14] M.J.S. Valença, T.B. Ludermit, “Selforganization neurons blocks networks”, ICCIMA, IEEE, pages 60-64, New Delhi, Índia, Sep. 1999.
- [15] H.D.C. Chaves Junior, Monitoração ambulatorial da pressão arterial (MAPA-24h): Sua reprodutibilidade e importância na construção de modelos para estimar a hipertrofia cardíaca, 2000. Tese de Doutorado, Universidade de São Paulo.
- [16] M.J.S. Valença, *Aplicando redes neurais: um guia completo*, Olinda-PE, 2005.