

***INSTITUTO TECNOLÓGICO DE AERONÁUTICA***



Vivek Nigam

**ESTUDO COMPARATIVO DA APLICAÇÃO DE  
TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL PARA A  
PREVISÃO DE PESO DE RECÉM-NASCIDOS**

*Trabalho de Graduação  
2003*

*Eletrônica*

VIVEK NIGAM

**ESTUDO COMPARATIVO DA APLICAÇÃO DE TÉCNICAS  
DE INTELIGÊNCIA ARTIFICIAL PARA A PREVISÃO DE  
PESO DE RECÉM-NASCIDOS**

Orientador

Prof. Dr. Cairo L. Nascimento Jr. (ITA)

Co-orientador

Prof. Dr. Luiz Fernando C. Nascimento (UNITAU)

**Divisão de Engenharia Eletrônica**

SÃO JOSÉ DOS CAMPOS

CENTRO TÉCNICO AEROESPACIAL

INSTITUTO TECNOLÓGICO DE AERONÁUTICA

2003

**Dados Internacionais de Catalogação-na-Publicação (CIP)**

**Divisão Biblioteca Central do ITA/CTA**

Nigam, Vivek  
Estudo Comparativo da Aplicação de Técnicas de Inteligência Artificial para a Previsão de Peso de Recém-Nascidos / Vivek Nigam  
São José dos Campos, 2003.  
41f.

Trabalho de Graduação – Divisão de Engenharia Eletrônica –  
Instituto Tecnológico de Aeronáutica, 2003. Orientadores: Prof. Dr. Cairo L. Nascimento Jr., Prof. Dr. Luiz Fernando C. Nascimento

1. Previsão. 2. Redes Neurais. 3. Sistemas nebulosos. 4. Mapas auto-organizáveis I. Centro Técnico Aeroespacial. Instituto Tecnológico de Aeronáutica. Divisão de Engenharia Eletrônica. III. Título

**REFERÊNCIA BIBLIOGRÁFICA**

NIGAM, Vivek. **Estudo comparativo da aplicação de técnicas de inteligência artificial para a previsão de peso de recém-nascidos**. 2003. 41f. Trabalho de Conclusão de Curso. (Graduação) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

**CESSÃO DE DIREITOS**

NOME DO AUTOR: Vivek Nigam

TÍTULO DO TRABALHO: Estudo comparativo da aplicação de técnicas de inteligência artificial para a previsão de peso de recém-nascidos

TIPO DO TRABALHO/ANO: Graduação / 2003

É concedida ao Instituto Tecnológico de Aeronáutica permissão para reproduzir cópias deste trabalho de graduação e para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta monografia de graduação pode ser reproduzida sem a autorização do autor.

---

Vivek Nigam

R. XXXXX

CEP XXXXXX – XXXXX

# ESTUDO COMPARATIVO DA APLICAÇÃO DE TÉCNICAS DE INTELIGÊNCIA ARTIFICIAL PARA A PREVISÃO DE PESO DE RECÉM- NASCIDOS

Essa publicação foi aceita como Relatório Final de Trabalho de Graduação

---

Vivek Nigam  
Autor

---

Prof. Dr. Cairo L. Nascimento Jr. (ITA)  
Orientador

---

Prof. Dr. Luiz Fernando C. Nascimento (UNITAU)  
Co-orientador

---

Prof. Dr. David Fernandes  
Coordenador do Curso de Engenharia Eletrônica

São José dos Campos, 20 de novembro de 2003

## Agradecimentos

*Primeiramente, gostaria de agradecer aos meus pais, Narendra Narain e Poonam Nigam, pelos anos dedicação e sacrifícios. Em especial, gostaria de agradecer ao meu irmão, Avishek Nigam, pelos anos de paciência e solidariedade que me ajudaram a ser uma pessoa melhor. Gostaria de agradecer também, ao meu orientador, Prof. Cairo L. Nascimento Jr., pelos ótimos conselhos não somente para a realização deste trabalho como para as difíceis decisões do meu futuro profissional. Agradeço também ao Prof. Luiz Fernando C. Nascimento e todas as mães que forneceram os dados necessários para realização deste trabalho. Finalmente, gostaria de agradecer à Deus, por ter me dado tantas oportunidades nesta vida.*

## *Pensamento*

*“Se os teus projetos forem para um ano, semeia o grão. Se forem para dez anos, planta uma árvore. Se forem para cem anos, educa o povo.”*

*(Provérbio chinês)*

## Resumo

Este trabalho descreve a aplicação de três técnicas de Inteligência Artificial, Sistemas *Neuro-Fuzzy*, Redes Neurais e Redes de *Kohonen*, para a criação de modelos matemáticos que possam ser usados algumas semanas antes do parto para a previsão da faixa de peso de recém-nascidos, a partir de certos dados da mãe. Estes modelos foram criados a partir dos dados coletados no ambulatório do Hospital Universitário da UNITAU durante os últimos três anos referentes a 550 nascimentos. As etapas para obtenção dos modelos de previsão foram: 1) seleção das variáveis de entrada para os modelos a partir dos dados coletados no ambulatório, 2) decisão sobre o número e os intervalos das faixas para o peso do recém-nascido, 3) separação dos dados coletados em dados de treinamento, validação e teste, 4) geração e testes dos modelos de previsão. Devido ao limitado número de dados disponíveis, o número de faixas para a classificação do peso do recém-nascido foi arbitrado em 3. Foi então desenvolvida uma metodologia para a definição dos intervalos destas 3 faixas. As 3 técnicas utilizadas então foram aplicadas para o desenvolvimento de modelos de previsão para cada uma das 3 faixas. As taxas de acerto medidas no conjunto de teste foram: 1) para a técnica *Neuro-Fuzzy*: de 56% a 70%, 2) para a técnica neural: de 63% a 66% e 3) Redes de *Kohonen*: de 58% a 64%.

## Abstract

This paper concerns the development of mathematical models to predict, a few weeks before birth, in which range the weight of a newborn will be, given some general physical information about the mother. Three techniques from the artificial intelligence field are used: neuro-fuzzy systems, artificial neural networks and *Kohonen* networks. These models were based on data which was collected at the UNITAU University Hospital in the last 3 years and contains information related to 550 births. The following stages were used in order to develop the models: 1) selection of the input variables, 2) decision about the number of bands and range of each band for the weight of the newborns, 3) partition of the data set in training, validation and test sets, 4) generation and testing of the models. Due to the small size of the data, only 3 bands were used to classify the weight of the newborn and the 3 artificial intelligence techniques were applied to develop a model for each band. Using the test set, the success rates were measured as between 56% and 70% for the neuro-fuzzy models, between 63% and 66% for the artificial neural network models and between 58% and 64% for *Kohonen* network models.



# Sumário

<b>1. INTRODUÇÃO .....</b>	<b>1</b>
<b>2. PROCESSO DE APRENDIZADO .....</b>	<b>4</b>
2.1. <i>Aprendizado Supervisionado.....</i>	<i>5</i>
2.2. <i>Aprendizado com Reforço .....</i>	<i>6</i>
2.3. <i>Aprendizado não Supervisionado.....</i>	<i>7</i>
<b>3. SISTEMAS NEURO-FUZZY.....</b>	<b>8</b>
3.1. <i>Lógica Fuzzy.....</i>	<i>8</i>
3.2. <i>Sistemas Neuro-Fuzzy tipo TSK .....</i>	<i>10</i>
<b>4. REDES NEURAIIS.....</b>	<b>12</b>
4.1. <i>Origem .....</i>	<i>12</i>
4.2. <i>Estrutura das Redes Neurais Artificiais (RNA).....</i>	<i>14</i>
4.3. <i>Algoritmo Back-Propagation .....</i>	<i>15</i>
<b>5. REDES DE KOHONEN .....</b>	<b>17</b>
5.1. <i>Algoritmo de Treinamento da Rede Neural de Kohonen .....</i>	<i>18</i>
<b>6. METODOLOGIA DE DESENVOLVIMENTO DOS MODELOS.....</b>	<b>20</b>
<b>7. AVALIAÇÃO DOS RESULTADOS .....</b>	<b>24</b>
7.1. <i>Critérios de Avaliação dos Modelos com Aprendizado Supervisionado</i>	<i>24</i>
7.2. <i>Critério de Avaliação dos Modelos de Redes de Kohonen .....</i>	<i>25</i>
7.3. <i>Resultados Obtidos .....</i>	<i>28</i>
<b>8. CONCLUSÃO.....</b>	<b>31</b>
<b>9. REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>32</b>

## 1. Introdução

A falta de uma infraestrutura adequada para a monitoração das mães no seu estado gestacional e o fato de não haver uma conscientização da população, especialmente de baixa renda, da importância da consulta pré-natal são os principais responsáveis pela existência de um elevado número de partos sem um profundo conhecimento do estado do feto e da mãe.

Visto que a faixa de peso do recém-nascido é o principal fator utilizado para decidir se o recém-nascido precisa ou não de tratamento especial, ou seja, de mais recursos humanos e de infra-estrutura, este trabalho se torna um estudo importante para auxiliar o médico nas decisões de atendimento da mãe em seu processo de parto. Com o resultado deste tipo de pesquisa, o médico ou administrador do centro médico poderá alocar de forma mais eficiente os recursos disponíveis para os casos mais críticos.

As técnicas de Inteligência Artificiais (IA) têm evoluído intensamente nas últimas décadas, surgindo diversas técnicas de construção de modelos não-lineares a partir de dados reais. Este trabalho visa comparar a aplicação de três destas técnicas de modelagem, Redes Neurais, Sistemas *Neuro-Fuzzy*, e Redes de *Kohonen* para a previsão da faixa do peso de recém-nascidos a partir de dados das mães coletados em ambulatório através do preenchimento de prontuários correspondentes a 550 diferentes gestações.

Além de comparar as diferentes técnicas de IA, este trabalho tem como objetivo construir um modelo quantitativo para realizar a previsão da faixa de peso do recém-nascido. Este objetivo se torna bastante desafiante, devido ao fato que atualmente, os especialistas conseguem somente desenvolver modelos de previsão qualitativos, de acordo com sua experiência. Não se conhecem nem mesmo quantas e quais são as variáveis de entrada para este tipo de modelo. Para este trabalho, devido a limitação dos tipos dados coletados, os modelos considerados utilizam as seguintes 8 entradas com informações da mãe:

- 1) Idade da mãe;
- 2) Número de filhos anteriores mortos no parto;
- 3) Número de filhos nascidos com menos de 2500 g;
- 4) Número médio de cigarros tragados por dia pela mãe durante a gestação;
- 5) Número de consultas ao longo da gestação;
- 6) Idade escolar da mãe (medida em anos);
- 7) Ganho de peso da mãe durante a gestação;
- 8) Numeração desta gestação (p. ex., 3<sup>a</sup> gestação);

Para este tipo de trabalho é fundamental que os dados utilizados sejam confiáveis, pois todo o conhecimento utilizado para realizar a construção dos modelos está contido nos dados. Este trabalho não teve como objetivo a validação das informações contidas nos prontuários, portanto os dados foram utilizados na forma que foram preenchidos pelas mães no ambulatório.

O peso do recém-nascido é categorizado pelos médicos em 3 faixas de peso:

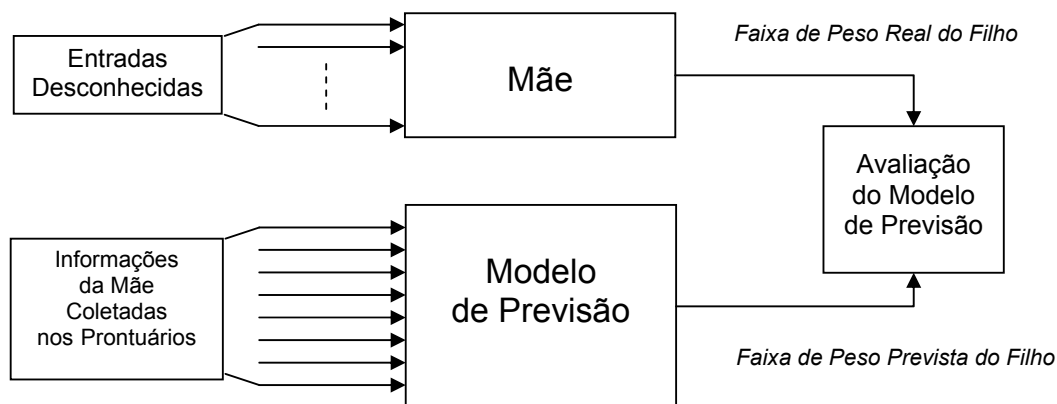
- 1) Crítico: peso menor que 2500 g;
- 2) Sub-crítico: peso entre 2500 g e 3500 g;
- 3) Normal: peso acima de 3500 g;

Este trabalho teve como objetivo inicial desenvolver modelos que indicassem na sua saída uma dessas 3 categorias de peso. Os testes iniciais indicaram que os dados coletados continham poucas amostras de situações onde o peso do recém-nascido estava na faixa de peso crítico (abaixo de 2500 g) o que tornava muito difícil a obtenção de modelos precisos para esta situação.

A solução adotada neste trabalho foi então a separação dos dados disponíveis em 3 faixas, adotando como critério de separação a variável peso do recém-nascido, de forma tal que cada uma das 3 faixas tivesse o mesmo número de dados. As novas 3 faixas de pesos então determinadas a partir dos dados foram:

- 1) Menor que 2995 g;
- 2) Entre 2995 g e 3370 g;
- 3) Maior que 3370 g;

Uma vez feita tal separação dos dados, para as técnicas baseadas em Sistemas *Neuro-Fuzzy* e Redes Neurais foram desenvolvidos 3 modelos independentes e correspondentes a cada uma das 3 regiões. Cada modelo deve indicar se os dados da mãe inseridos como entrada resultarão em um peso do recém-nascido dentro (saída 1) ou fora (saída -1) da faixa do modelo. Quanto mais próximo de 1 for a saída do modelo, maior será a possibilidade do peso do recém-nascido estar dentro da faixa associada ao modelo. O contrário se aplica quanto mais próximo a saída do modelo estiver de -1.



**Figura 1 - Ilustração da estrutura do problema**

Para a técnica de Redes de *Kohonen* também foram desenvolvidos três modelos (correspondentes as três faixas de peso), contudo ao inserirmos os dados da mãe, as saídas dos modelos resultam nas distâncias da região ativada pelos dados de entrada aos centros dos clusters correspondentes as regiões pertencentes à faixa de peso e não pertencentes a faixa de peso. Quanto menor a distância da região ativada ao cluster pertencente à faixa de peso, maior será a possibilidade do recém-nascido estar dentro da faixa associada ao modelo. O contrário se aplica quanto menor for a distância da região ativada ao cluster não pertencente à faixa de peso.

Para a determinação dos parâmetros do modelo (em Inteligência Artificial esta etapa é denominada treinamento) foi necessário o desenvolvimento de uma metodologia de seleção dos dados para o ajuste e teste de cada um dos modelos.

As próximas seções apresentam:

- 1) Tipos de aprendizados;
- 2) A estrutura dos modelos baseados em sistemas *neuro-fuzzy*, redes neurais e redes de *Kohonen* ;
- 3) Os métodos de obtenção dos parâmetros dos modelos;
- 4) Os critérios de avaliação dos resultados obtidos pela aplicação dos modelos;
- 5) Conclusões

## 2. Processo de Aprendizado

Antes de definirmos e explicarmos os modelos e técnicas utilizados neste trabalho (Redes Neurais, Sistemas *Neuro-Fuzzy* e Redes de *Kohonen*) é importante explicarmos os tipos de processos de aprendizados. Existem três tipos de aprendizados:

- Aprendizado Supervisionado(*supervised learning*)
- Aprendizado com Reforço (*reinforcement learning*)
- Aprendizado não Supervisionado (*unsupervised learning*)

As técnicas *Neuro-Fuzzy* e Redes Neurais utilizam o aprendizado supervisionado, enquanto que as Redes de *Kohonen* utilizam o aprendizado não supervisionado[9].

## 2.1. Aprendizado Supervisionado

Neste tipo de aprendizado é fundamental a existência de um “sistema” externo que possui o conhecimento do ambiente, representado por uma série de exemplos de entrada e saída. Neste trabalho o “sistema” seria os dados dos prontuários contendo as informações da mãe (dados de entrada) e os pesos dos recém-nascidos (dados de saída).

A saída do sistema é capaz de prover ao modelo a resposta necessária para dados de entrada. O ajuste dos parâmetros do modelo é feito utilizando o sinal erro obtido pela diferença da saída do modelo e da resposta do sistema, conforme a figura 2.

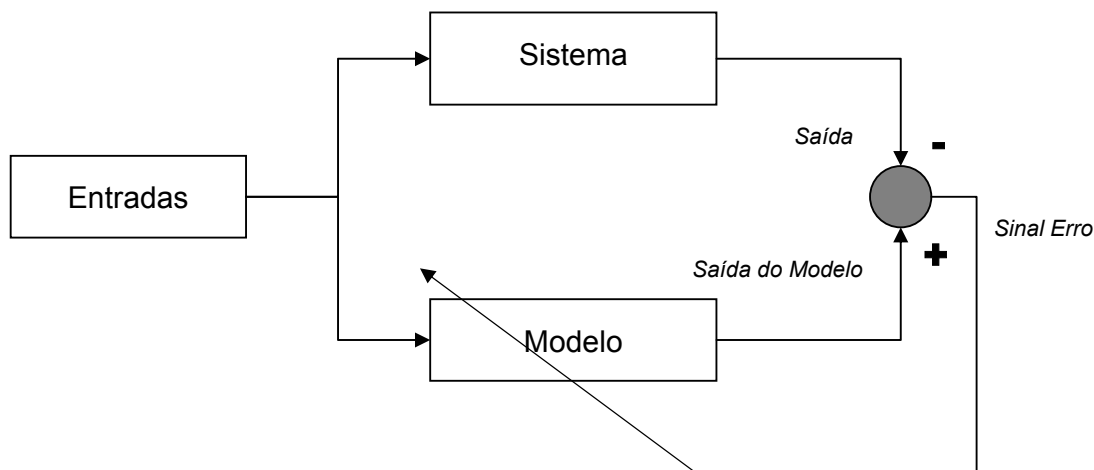


Figura 2 - Esquema do processo de aprendizado supervisionado

Finalmente quando os parâmetros do sistema é transmitido ao modelo, podemos dispensar o sistema e utilizar o modelo gerado para simular o conhecimento do ambiente.

## **2.2. Aprendizado com Reforço**

O aprendizado com reforço é feito através de um processo de tentativa e erro projetado para maximizar um índice de performance denominado sinal de reforço.

Este tipo de aprendizado é uma forma de aprendizado supervisionado, pois existe um retorno do ambiente. Contudo este retorno é um simples sinal sim/não, sendo, portanto apenas avaliativo e não instrutivo.

Um exemplo típico deste processo é a fase em que uma criança aprende a andar. Os parentes funcionam como uma espécie de crítico dizendo se a criança está no caminho correto ou não. Mas os pais não conseguem transferir o seu conhecimento de como andar para a criança, portanto a criança aprende por tentativa e erro até atingir o índice de performance, que neste caso seria, por exemplo, chegar ao pai.

### 2.3. Aprendizado não Supervisionado

O objetivo do aprendizado não supervisionado é descobrir regularidades ou similaridades nos dados de treinamento, de forma que as entradas semelhantes sejam agrupadas e produzam saídas semelhantes (no mesmo cluster). Portanto o modelo deve ser capaz de descobrir categorias (neste trabalho seriam as três faixas de peso do recém-nascido) para os dados de entradas.

Contudo neste tipo de aprendizado não existe um “sistema” externo ou algum “crítico” para vigiar o processo de aprendizado (como ocorria nos outros dois tipos de aprendizados). Portanto, é necessária uma ferramenta independente de medida da qualidade de representação da rede para o aprendizado e os livres parâmetros do modelo são otimizados com respeito a esta medição. Uma vez que a rede é ajustada para um número representativo de amostras de entrada, ela desenvolve a capacidade de descobrir as semelhanças nos dados de entrada e cria clusters automaticamente.

Este processo se divide em duas etapas:

- A formação de clusters de unidades neurais correspondendo aos clusters das entradas, para isso utiliza-se somente os dados de entrada não necessitando de um supervisor (representado pelo dado de saída).
- A rotulação (*labelling*) destes clusters, nesta fase o supervisor se faz necessário para fornecer os rótulos dos clusters. (No caso deste trabalho seria se os dados do cluster pertencem ou não a faixa de peso correspondente ao modelo).



### 3. Sistemas *Neuro-Fuzzy*

#### 3.1. Lógica *Fuzzy*

A lógica *Fuzzy* é uma forma simples de aproximar o raciocínio humano ao da lógica executada pela máquina. Diferente da lógica convencional, onde as transições de um estado membro a um estado não-membro é abrupta e repentina, na lógica fuzzy os estados membro e o não membro estão numa faixa gradual, sendo associado um grau ao elemento entre 0 (totalmente não membro) e 1 (totalmente membro).

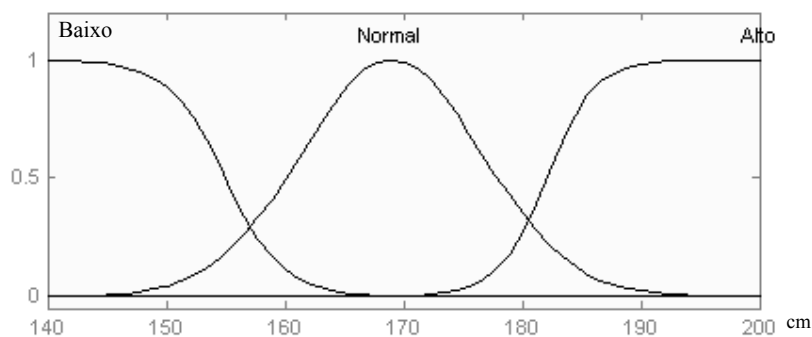
Por exemplo, para classificação de uma pessoa de acordo com a altura, na lógica convencional seriam definidas diversas faixas:

- Baixo – Menor que 1,60 m;
- Normal – Entre 1,60 e 1,80 m;
- Alto – Acima de 1,80 m;

Portanto, pessoas com 1,79 m de altura e 1,81 m de altura seriam classificadas de forma diferente, uma com altura normal e outra com altura alta. Porém na realidade o que diferencia estas duas pessoas são 2 cm, o que na vida cotidiana não é uma diferença muito grande.

Já na lógica *fuzzy* seria adotada uma função contínua de 0 a 1, denominada de função de pertinência, para cada uma das classificações de altura. Usualmente funções do tipo triangulares, trapezoidais ou gaussianas são usadas como funções de pertinência.

A figura 3 apresenta três funções de pertinência para cada classificação da altura de uma pessoa. Portanto nesta lógica uma pessoa teria graus pertinência para cada característica, por exemplo, uma pessoa com 1,79 m de altura seria 0 baixo, 0,35 normal e 0,25 alto. Enquanto que uma pessoa de 1,81 m seria 0 baixo, 0,25 normal e 0,40 alto.



**Figura 3 - Demonstração das Funções de Pertinência da Lógica *Fuzzy***

Este tipo de lógica é útil, pois as transições de um estado para outro são bastante suaves e intuitivas.

### 3.2. Sistemas *Neuro-Fuzzy* tipo TSK

Neste trabalho foram usados Sistemas *Neuro-Fuzzy* do tipo TSK (Takagi-Sugeno-Kang). Nestes modelos a saída é obtida pela combinação de vários modelos usualmente lineares de abrangência local. Um modelo *neuro-fuzzy* é construído da seguinte forma:

1. A cada variável de entrada  $x_i$  são associadas várias funções de pertinência.
2. A abrangência de cada modelo local é definida por uma regra do tipo:

$$\text{Regra } i: \text{Se } (x_1 \in MF_{1j}) \text{ e } \dots (x_8 \in MF_{8k}) \text{ então } y_i = w_{1i} * x_1 + \dots + w_{8i} * x_8 + bias_i. \quad (1)$$

onde  $MF_{1j}, \dots, MF_{8k}$  são as funções de pertinência associadas às variáveis de entrada usadas pela  $i$ -ésima regra.

3. A saída escalar do modelo  $y$  é calculada pela ponderação das saídas dos modelos associados a cada uma das regras, ou seja:

$$y = \left[ \sum_{i=1}^N z_i y_i \right] / \left[ \sum_{i=1}^N z_i \right] \quad (2)$$

onde  $N$  denota o número de regras e  $z_i$  indica o coeficiente de ponderação da  $i$ -ésima regra. Tal coeficiente é calculado pelo grau de satisfação da regra (grau de veracidade do antecedente da regra) para um certo valor numérico das entradas de acordo com a lógica de inferência *fuzzy* [4,5,10].

- 3.3. A determinação do número de regras e das funções de pertinência e dos parâmetros dos modelos ( $w_{1i}, \dots, w_{8i}, bias_i$ ) (respectivamente antecedente e consequente) em cada regra é feita pelo método ANFIS (*Adaptive Neuro-Fuzzy Inference System*). Tal método usa a técnica de clusterização de dados denominada *Sub-Clustering* para, a partir dados de treinamento disponíveis e com pouco esforço numérico, criar um modelo grosseiro do tipo entrada-saída. Posteriormente o algoritmo de treinamento

*Back-Propagation* é então usado para um ajuste final das funções de pertinência e dos parâmetros dos modelos lineares de cada regra [5,6,10].

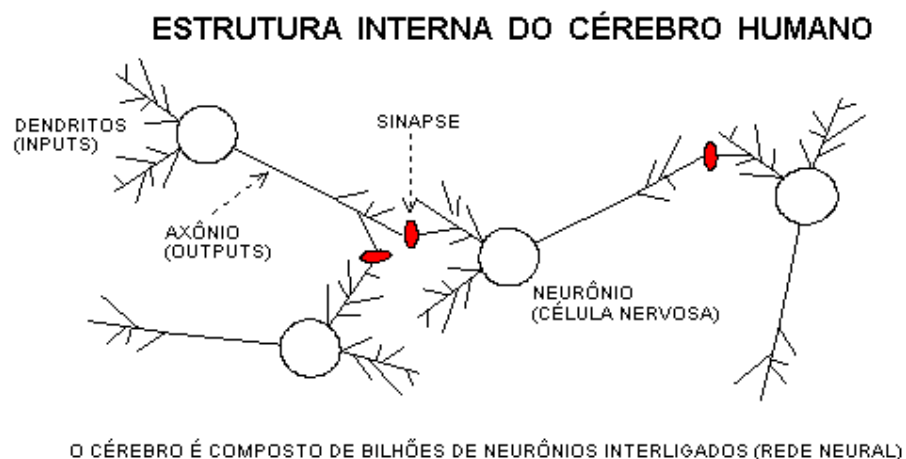
## 4. Redes Neurais

### 4.1. Origem

A idéia das redes neurais começou ao tentar-se entender melhor o cérebro humano. Alguns cientistas acreditavam que o cérebro era uma massa contínua e homogênea, como outros órgãos humanos, contudo existiam outros cientistas que acreditavam que o cérebro era discreto, formado por algum tipo de “elemento”.

Até que os cientistas ganhadores do Prêmio Nobel, Golgi e Cajal, desenvolveram um método para “colorir” o cérebro usando um corante que fez com que os axônios tomassem uma coloração diferente da massa cerebral, evidenciando efetivamente que o cérebro não é uma massa homogênea.

O cérebro humano é formado por bilhões desses “elementos”, denominados de neurônios, formando uma rede neural extremamente complexa. Os neurônios são formados por dois principais constituintes: o axônio e o dendrito. Os axônios são caracterizados por serem longos terminais de saída e os dendritos um conjunto de terminais de entrada.



**Figura 4 - Estrutura Interna do Cérebro Humano**

Os neurônios se comunicam através de impulsos elétricos transmitidos na região onde o axônio e o dendrito entram em contato, denominados de sinapse. Os impulsos recebidos por um neurônio são disparados somente quando este impulso atinge um limiar de ação, fazendo

com que o neurônio dispare um neurotransmissor que aumenta a polaridade da membrana pós-sináptica, inibindo ou excitando a geração dos pulsos para um outro neurônio.

Concebendo procedimentos matemáticos análogos ao funcionamento dos neurônios biológicos, Warren McCulloch e Walter Pitts estabeleceram, em 1943, as bases da neurocomputação. Esta contribuição foi puramente conceitual já que os autores não sugeriram aplicações práticas para o seu trabalho, mesmo porque os sistemas propostos por eles não tinham capacidade de aprender.

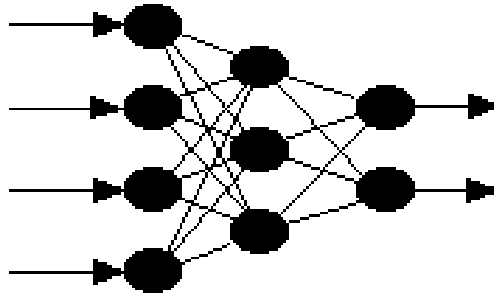
Em 1949 Donald Hebb deu um passo muito importante na história das redes neurais. Ele propôs um modo de proporcionar capacidade de aprendizado às redes neurais artificiais. Sua proposta, traduzida matematicamente, viabilizou o desenvolvimento de redes neurais eficazes.

Nos anos 50 já existiam várias implementações de Redes Neurais através de circuitos analógicos e, naquela época, acreditou-se que o caminho para o entendimento da inteligência humana havia sido descoberto [1]. Para reproduzir o comportamento do cérebro humano pensava-se que bastaria construir uma rede neural suficientemente grande. Uma rigorosa análise matemática, no entanto, comprovou o pouco poder computacional dos modelos de rede neural utilizados na época, levando as pesquisas neste campo a ficarem esquecidas de meados dos anos 60 até o início da década de 80.

Nos anos 80 o interesse pela área retornou, devido, em grande parte, ao surgimento de novos modelos neurais, como o proposto por *Hopfield* [2] e *Kohonen* [3] e de algoritmos de aprendizado mais poderosos. Além disso, nesta mesma época, ocorreu o surgimento de computadores mais rápidos e poderosos, facilitando a implementação das redes neurais.

## 4.2. Estrutura das Redes Neurais Artificiais (RNA)

Na técnica de modelagem baseada em Redes Neurais Artificiais (RNA) a estrutura deste modelo é constituída de unidades computacionais (chamadas de “células neurais”) distribuídas em camadas. Tipicamente são usadas 3 camadas que são então denominadas de entrada, escondida e de saída (figura 5).



**Figura 5 - Estrutura do Modelo Baseado em Redes Neurais Artificiais com 4 Entradas, 3 Unidades Escondidas e 2 Saídas**

As unidades enviam as suas saídas para as camadas superiores, sendo que nenhum processamento é feito nas unidades situadas nas camadas de entrada. Cada modelo neural desenvolvido neste trabalho possui 8 entradas e 1 saída calculada pela seguinte expressão:

$$z_i = \tanh \left( \sum_{j=1}^8 w_j^{IH} x_j + bias_i^H \right), i = 1, \dots, NH$$

$$y = \tanh \left( \sum_{j=1}^{NH} w_j^{HO} z_j + \sum_{j=1}^8 w_j^{IO} x_j + bias \right) \quad (3)$$

onde NH indica o número de unidades usadas na camada escondida.

Neste trabalho todos os modelos neurais desenvolvidos usaram 10 unidades na camada escondida com ligação direta entre as unidades na camada de entrada e a unidade de saída.

Os pesos das conexões ( $w^{IH}$ ,  $w^{HO}$ ,  $w^{IO}$ ,  $bias^H$ ,  $bias$ ) que ligam as diversas unidades são os parâmetros do modelo que devem ser ajustados pelo algoritmo de treinamento da rede neural. Neste trabalho o algoritmo *Back-Propagation* foi utilizado para o treinamento dos modelos baseados em RNA [1].

### 4.3. Algoritmo *Back-Propagation*

O aprendizado por *back-propagation* baseia-se na propagação retrógrada do erro para os níveis anteriores da rede, de acordo com o grau de participação que cada neurônio teve no nível posterior. Portanto se trata de um aprendizado supervisionado, pois a rede produz sua própria saída e compara-se com a saída desejada (obtida com o “professor” externo).

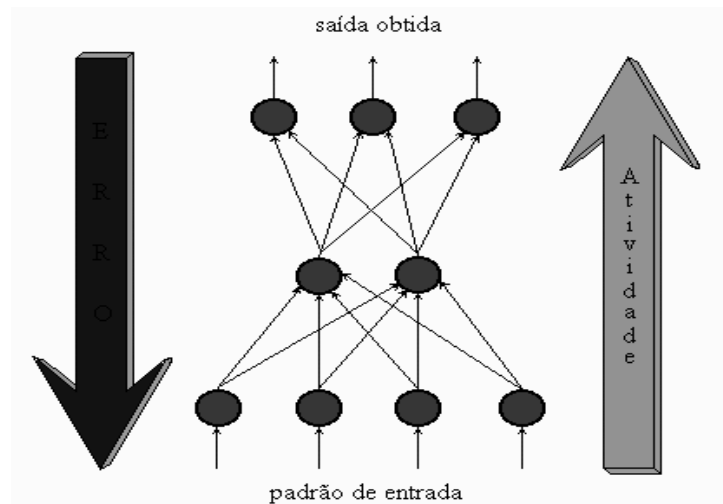


Figura 6 - Ilustração do algoritmo *Back-Propagation*

O algoritmo de aprendizado *back-propagation* apresenta os seguintes passos:

Passo 1: Inicialização dos pesos com valores aleatórios dentro de uma determinada faixa;

Passo 2: Apresentação do vetor entrada com as respectivas saídas desejadas:

$$X_i = [x_1, x_2, \dots, x_n, saída]$$

Passo 3: Cálculo das saídas de todas as unidades da RNA:

$$[y_1^H(1), \dots, y_k^H(N_{hid}), y_k^S(1), \dots, y_k^S(N_{out})]$$

onde:

$N_{hid}$  é o número de unidades na célula escondida;

$N_{out}$  é o número de saídas do modelo;



Passo 4: Cálculo do erro de cada unidade de saída:

$$e_k^S = d_k^S - y_k^S$$

onde:

$d_k^S$  é o valor de saída desejado (obtida com “professor” externo)

Passo 5: Cálculo do erro de cada unidade das camadas escondidas:

$$e_k^H = - \sum_{i=1}^{N_{out}} e_k^S(i) \bar{f}_k^S(i) w_k^S(i, j)$$

onde:

$\bar{f}_k^S(i)$  - Derivada da função de ativação da unidade de saída;

$w_k^S(i, j)$  - Peso que parte da unidade i para a unidade de saída j;

Passo 6: Atualizar os pesos das unidades:

- Unidades de Saída:

$$\Delta w_k^S = \eta e_k^S(i) \bar{f}_k^S(i) y_k^H(j)$$

- Unidade da Camada Escondida

$$\Delta w_k^H = \eta \left[ \sum_{l=1}^{N_{out}} e_k^S(l) \bar{f}_k^S(l) w_k^S(l, i) \right] \bar{f}_k^H(i) u_k(j)$$

onde:

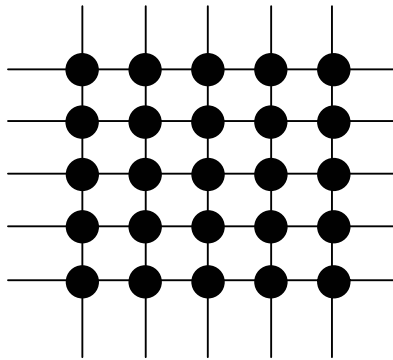
$\eta$  - Índice de aprendizado;

$u_k$  - Dados de entrada;

Passo 7: Voltar ao passo 2 até o término do conjunto de padrões de treinamento, ou até o número máximo de iterações ou até atingir um erro de saída aceitável;

## 5. Redes de *Kohonen*

Uma rede de *Kohonen* é formada por uma rede neural de uma camada com M unidades e N entradas dispostas fisicamente em um arranjo geométrico uniforme (figura 7) em um espaço de 1,2 ou 3 dimensões. Se duas unidades estão topologicamente próximas neste espaço, estas unidades são vizinhas e a distância entre elas é a distância euclidiana.



**Figura 7 - Ilustração de uma Rede de *Kohonen* com Dimensão 5x5**

Nesta técnica todas as M unidades recebem a mesma entrada e de acordo com a saída de cada unidade é selecionada uma unidade vencedora. Esta seleção é feita de acordo com o seguinte critério:

$$S^v = \min(\|x_k - w^i\|)$$

onde:

$x_k$  - vetor com os dados de entrada;

$w^i$  - vetor com os pesos da unidade  $i$ ;

$v$  - a unidade vencedora;

No processo de aprendizado da rede de *Kohonen* (não supervisionado) são definidos na fase de rotulação os clusters desta rede. Determinado a unidade vencedora  $v$ , determina-se o cluster mais perto desta unidade vencedora e n conseqüentemente através do rótulo deste cluster obtém-se a classificação dos dados de entrada (Neste trabalho seria se os dados de entrada pertencem ou não à faixa de peso do modelo).

### 5.1. Algoritmo de Treinamento da Rede Neural de *Kohonen*

O algoritmo de treinamento desta técnica é composto pelos seguintes passos:

- Passo 1: Iniciar os pesos da rede neural com valores aleatórios;
- Passo 2: Fazer  $k = 1$ ;
- Passo 3: Escolher uma entrada do conjunto de treinamento  $x_k$ ;
- Passo 4: Calcular a similaridade entre os pesos da unidade  $i$  da rede neural e a entrada  $x_k$ ;
- Passo 5: Identificar a unidade vencedora  $v$

$$S_k^v = \min(\|x_k - w_k^i\|)$$

- Passo 6: Atualizar os pesos da unidade vencedora e de sua vizinhança seguindo:

$$w_{k+1}^i = w_k^i + \eta_k (x_k - w_k^i), \text{ se } i \in \text{Viz}(v);$$

$$w_{k+1}^i = w_k^i, \text{ se } i \notin \text{Viz}(v);$$

onde:

$\eta_k$  - taxa de aprendizado;

$\text{Viz}(v)$  – representa a vizinhança da unidade vencedora  $v$ ;

**Observação:** A vizinhança de uma unidade central  $i$  é determinada por um parâmetro  $d$ , todas as unidades que tiverem uma distância euclidiana à unidade central  $i$  menor ou igual à  $d$  serão pertencentes à vizinhança da unidade central  $i$ .

Este passo é fundamental para estabelecer uma região de agrupamento com características semelhantes e conseqüentemente formar um cluster, pois os pesos de todas as unidades da vizinhança da unidade vencedores  $v$  são ajustados na mesma proporção.

- Passo 7: Reduzir o tamanho da vizinhança  $\text{Viz}(v)$ , ou seja diminuir o parâmetro  $d$ ;
- Passo 8: Reduzir a taxa de aprendizado  $\eta_k$ ;
- Passo 9: Fazer  $k = k + 1$ ;
- Passo 10: Retornar ao passo 3

- Passo 11: Após termino das iterações de treinamento - Utilizar amostras representativas de cada classificação para rotular os cluster obtidos até o momento

## 6. Metodologia de Desenvolvimento dos Modelos

Para o desenvolvimento dos modelos, os dados de entrada-saída disponíveis são divididos em 3 conjuntos:

- a) Conjunto dos dados de treinamento: é o conjunto de dados usado pelo algoritmo de treinamento para o ajuste dos parâmetros do modelo,
- b) Conjunto dos dados de validação: é o conjunto de dados usado durante o treinamento para medir o grau de “aderência” excessiva do modelo aos dados de treinamento. Isto ocorre quando o erro medido sobre o conjunto de validação aumenta enquanto o erro medido sobre o conjunto de treinamento diminui, uma indicação do chamado “overfitting”, que é causado pela excessiva flexibilidade da estrutura do modelo em relação aos dados de treinamento.
- c) Conjunto dos dados de teste: é o conjunto de dados usado após o treinamento para uma avaliação quantitativa e efetiva da qualidade do modelo.

Para a geração dos modelos de previsão do peso do recém-nascido em cada uma das 3 faixas foi adotado que os dados devem possuir as seguintes propriedades:

- a) nos dados de treinamento: 50% dos dados devem corresponder a dados de entrada que produziram saídas dentro da faixa-alvo, 50% dos dados devem corresponder a dados de entrada que produziram saídas fora da faixa-alvo;
- b) o mesmo princípio deve ser adotado para os dados de validação e para os dados de teste. Denominando de sub-conjunto I os dados na faixa-alvo e sub-conjuntos II e III os dados nas outras 2 faixas, foi arbitrada a proporção entre os conjuntos de treinamento, validação e teste mostrada na tabela 1.

**Tabela 1** – Proporção entre os conjuntos de Treinamento, Validação e Teste

	<b>Sub-conjunto I</b>	<b>Sub-conjunto II</b>	<b>Sub-conjunto III</b>
<b>Treinamento</b>	80 %	40 %	40 %
<b>Validação</b>	10 %	5 %	5 %
<b>Teste</b>	10 %	5 %	5 %

Inicialmente os dados disponíveis foram separados nestes 3 conjuntos de forma aleatória. Apesar do erro medido sobre o conjunto de dados de treinamento ser satisfatório, no

caso dos 3 modelos *neuro-fuzzy* e nos 3 modelos neurais, o erro medido sobre o conjunto de testes foi julgado como insatisfatório (nas Redes de *Kohonen* não foi utilizado uma separação aleatória, foi utilizada a metodologia sugerida abaixo). Dada a natureza real dos dados, foi feita a hipótese que esta diferença na medida dos erros sobre os conjuntos estaria sendo causada pela não distribuição uniforme dos dados coletados nos conjuntos nos seus devidos espaços de entrada.

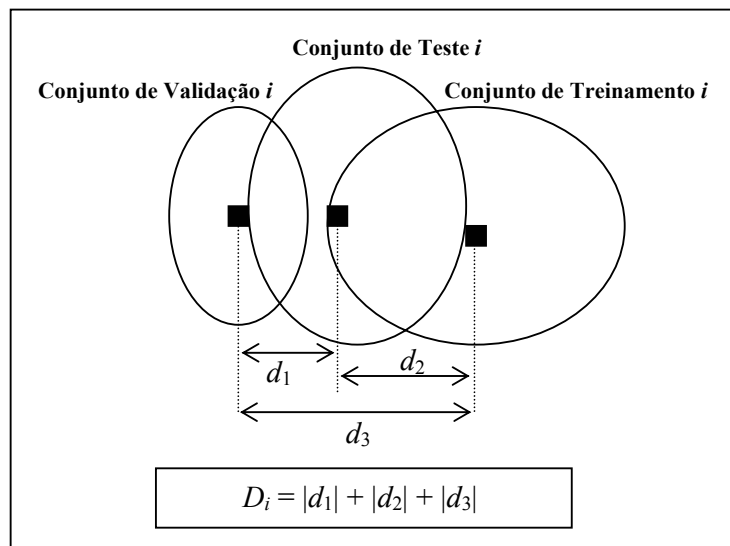
Foi decidido então que possivelmente uma maneira para melhorar o desempenho dos modelos sobre o conjunto de testes seria a implementação de uma melhor distribuição dos dados nos conjuntos de treinamento, validação e testes. Uma nova metodologia de formação dos conjuntos de dados foi então proposta baseada no princípio que os conjuntos de treinamento, validação e teste devem estar na mesma região do espaço de entrada. Os passos desta metodologia são:

- a) gerar aleatoriamente de acordo com a tabela 1 um grande número de propostas de distribuição dos dados nos conjuntos de treinamento, validação e teste,
- b) para cada proposta de distribuição de dados medir o seguinte índice em cada sub-conjunto:

$$D_i = |d_1| + |d_2| + |d_3| \quad (4)$$

onde os índices  $d$  indicam as distâncias entre as médias vetoriais dos conjuntos de treinamento, validação e teste (conforme mostrado na figura 8) sendo que cada componente do vetor de entrada foi previamente normalizada linearmente entre  $-1$  e  $1$ , considerando todos os dados coletados. Com tal normalização o cálculo do módulo do vetor  $d$  independe do intervalo de cada componente do vetor de entrada  $x$ .

- c) o índice final da proposta de solução é calculado pela soma  $S = D_I + D_{II} + D_{III}$ , onde  $D_I, D_{II}$  e  $D_{III}$  significam respectivamente o índice  $D$  para os sub-conjuntos I, II e III da tabela 1.
- d) para a geração dos modelos *neuro-fuzzy*, neurais e de redes de *Kohonen* é então adotada a proposta de distribuição dos dados com o menor valor escalar  $S$ .



**Figura 9 - Esquema com o Cálculo dos Índices de Distribuição dos Dados dos Conjuntos de Treinamento, Validação e Teste**

Para cada uma das 3 faixas de peso foram geradas e analisadas neste trabalho 100 propostas de distribuição dos dados.

O pacote computacional MATLAB foi usado tanto na separação dos dados em conjuntos como no desenvolvimento dos modelos *neuro-fuzzy* e neurais. Para o desenvolvimento dos 3 modelos *neuro-fuzzy* foram usadas as funções disponíveis no *Fuzzy Logic Toolbox* do MATLAB® [2] (O código fonte está em arquivos em anexo). Os modelos desenvolvidos apresentam as seguintes estruturas:

- a) Modelo I (abaixo de 2995 g): 16 regras e 16 funções de pertinências associadas a cada componente de entrada;
- b) Modelo II (entre 2995 g e 3370 g): 8 regras e 8 funções de pertinência associadas a cada componente de entrada;
- c) Modelo III (acima de 3370 g): 5 regras e 5 funções de pertinências associadas a cada componente de entrada.

Para o desenvolvimento dos 3 modelos neurais foi definido um novo conjunto de treinamento pela união dos conjuntos de treinamento e validação mencionados na tabela I. Uma implementação própria dos autores do algoritmo *Back-Propagation* (Código fonte está em arquivos em anexo) em MATLAB foi então utilizada por 300 iterações sobre o novo conjunto de treinamento para a obtenção dos 3 modelos neurais. A estrutura dos modelos neurais está expressa na eq. (3).

Para o desenvolvimento dos 3 modelos de Redes de *Kohonen* foi definido um novo conjunto de teste pela união dos conjuntos de validação e de teste mencionados na tabela I. Uma implementação da estrutura e do processo de aprendizado definido na seção 5.1 foi realizada pelos autores no MATLAB (Código fonte está nos arquivos em anexo). Contudo, para efetuarmos a rotulagem dos dois clusters (dos dados pertencentes e não pertencentes à faixa de peso) obtidos nos modelos utilizamos a média dos dados de treinamento daquela faixa como amostra de rotulagem. Porém para determinarmos o cluster dos dados não pertencentes à faixa de peso tivemos que escolher uma das duas médias obtidas das outras faixas de peso. As características de cada modelo estão descritas abaixo:

- a) Modelo I (abaixo de 2995 g):- Estrutura bidimensional com 20x20 unidades, 400 iterações com uma taxa de aprendizado inicial de 0,5 e ajuste feito pela função  $\exp(-10 \cdot \text{iter} / 400)$ . A média dos dados de treinamento da faixa III foi utilizada como amostra de dados fora da faixa I
- b) Modelo II (entre 2995 g e 3370 g): Estrutura bidimensional com 15x15 unidades, 400 iterações, com uma taxa de aprendizado inicial de 0,4 e ajuste feito pela função  $\exp(-5 \cdot \text{iter} / 400)$ . A média dos dados de treinamento da faixa I foi utilizada como amostra de dados fora da faixa II
- c) Modelo III (acima de 3370 g): Estrutura bidimensional com 22x22 unidades, 400 iterações com uma taxa de aprendizado inicial de 0,8 e ajuste feito pela função  $\exp(-10 \cdot \text{iter} / 400)$ . A média dos dados de treinamento da faixa I foi utilizada como amostra de dados fora da faixa III



## 7. Avaliação dos Resultados

### 7.1. Critérios de Avaliação dos Modelos com Aprendizado Supervisionado

Lembrando que um modelo com saída  $-1$  ou  $1$  indica que o peso do recém-nascido está respectivamente fora ou dentro da faixa prevista pelo modelo, foram utilizados os seguintes dois critérios para avaliação dos resultados considerando o conjunto de dados de teste:

**Critério I:** Os valores de saída do modelo acima de  $0$  foram considerados como uma previsão afirmando que o peso estará na região do modelo e os casos abaixo de  $0$  foram considerados como uma afirmação do peso não estar na região do modelo. Com este critério foram contabilizados os números de acertos e erros ao comparar a saída do modelo com o resultado real.

**Critério II:** Os valores de saída acima de  $0,5$  foram considerados como uma previsão afirmando que o peso estará na região do modelo e os casos abaixo de  $-0,5$  foram considerados como uma afirmação do peso não estar na região do modelo. Os casos entre  $-0,5$  e  $0,5$  foram considerados como o modelo está incerto em relação à previsão. Foram contabilizados os números de acertos, erros e incertezas ao comparar a saída do modelo com o resultado real.

## 7.2. Critério de Avaliação dos Modelos de Redes de *Kohonen*

Lembrando que em cada um dos 3 modelos são definidos dois cluster na fase de aprendizado:

- a) Cluster dos dados pertencentes à faixa de peso do modelo;
- b) Cluster dos dados não pertencentes à faixa de peso do modelo;

Para definirmos a que cluster um conjunto de entrada pertence, devemos calcular a distância euclidiana da unidade vencedora deste conjunto de entrada aos dois clusters. O cluster que estiver mais perto da unidade vencedora foi considerado o cluster a que pertence o conjunto de entrada. Porém, caso a unidade vencedora esteja a uma distância equivalente aos dois clusters, é considerado um estado de incerteza do modelo.

O número de acertos, erros e incertezas são contabilizados de acordo com o valor real do peso do recém-nascido. Por exemplo, um caso em que o recém-nascido tenha 3000 g (portanto da faixa II de peso) e a unidade vencedora para este recém-nascido no modelo I (modelo correspondente à faixa I - para pesos abaixo de 2995 g) esteja mais perto do cluster não pertencente à faixa é contabilizado um acerto, caso contrário um erro.

Por exemplo, seja um caso de um recém-nascido pertencente à faixa III (peso maior que 3370 g) for inserido em um modelo de previsão para a faixa I (peso abaixo 2995 g). A célula vencedora  $v$  para esta entrada está disposta na figura 9, neste caso a distância para o cluster pertencente à faixa I (na figura representado pelo símbolo P) é maior que a distância para o cluster não pertencente a faixa I (na figura representado pelo símbolo N), portanto neste caso é contabilizado um acerto.

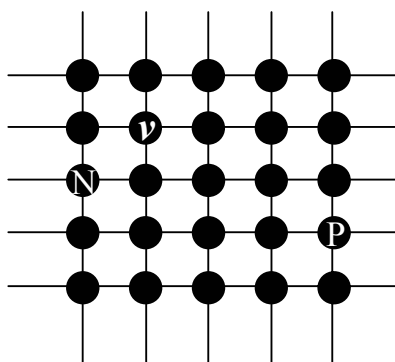


Figura 10 - Demonstração do Critério de Avaliação das Redes de Kohonen

Investigamos também outro critério de avaliação com um processo de rotulagem mais complexo. Ao invés de usar somente a média dos conjuntos de treinamento como representante dos clusters, utilizamos todos os dados de treinamento das três faixas de peso. Utilizamos os seguintes passos para efetuar a rotulagem do modelo:

Passo 1: Iniciar um vetor  $M$  com  $N$  posições, e para cada posição 3 índices, correspondentes a cada faixa com o valor zero;

onde  $N$  corresponde ao número de unidades no modelo;

Passo 2: Inserir um dado de um dos conjuntos de treinamento no modelo;

Passo 3: Verificar qual a unidade vencedora utilizando o mesmo critério definido na seção 5;

Passo 4: Para esta unidade vencedora (associada a uma posição do vetor  $M$ ) somamos 1 ao índice do vetor  $M$  correspondente à faixa do dado inserido;

Passo 5: Repetir o passo 2 até que todos os dados forem inseridos no modelo;

Realizado o processo de rotulagem, iniciamos o processo de teste do modelo utilizando os seguintes passos como critério de avaliação:

Passo 1: Inserir um dado do conjunto de teste no modelo;

Passo 2: Verificar qual a unidade vencedora, utilizando o mesmo critério definido na seção 5;

Passo 3: Para cada posição do vetor  $M$ , verificar qual o maior dos 3 índices. Esta posição será considerada um representante da faixa com o maior valor. Caso tenha um empate no maior valor em dois índices a posição é considerada uma incerteza e, portanto não é mais utilizada para os próximos passos;

Passo 4: Determinar a distância da unidade vencedora para cada unidade representante da faixa 1, faixa 2 e faixa 3;

Passo 5: Determinar qual o representante com a menor distância. Caso tenha-se duas unidades representantes de duas faixas distintas com a menor distância, é considerado uma incerteza e encerra-se o teste para este dado de entrada;

Passo 6: Determinado qual o representante, contabilizar um acerto ou um erro de acordo com o dado de entrada e do representante. Caso o dado de entrada e o representante sejam ambos pertencentes ou não à faixa do modelo é contabilizado um acerto, caso contrário é contabilizado um erro;

Mesmo sendo um processo de rotulagem mais complexo, este critério não foi muito eficaz na previsão da faixa de peso do recém-nascido, portanto continuamos a adotar o critério de avaliação e de rotulagem inicialmente propostos.

### 7.3. Resultados Obtidos

Utilizando os 3 modelos *neuro-fuzzy*, os 3 modelos neurais, a metodologia de seleção de dados explicada anteriormente e os critérios citados acima, os resultados mostrados nas tabelas 2 e 3 foram obtidos. Note que as taxas foram calculadas usando os conjuntos de testes e que o número de amostras é indicado entre parênteses.

**Tabela 2** – Resultados obtidos utilizando os dados de teste com os modelos *Neuro-Fuzzy*.

Modelos <i>Neuro-Fuzzy</i>	Critério I		Critério II		
	Acertos	Erros	Acertos	Erros	Incertezas
<b>Modelo I</b> ( <b>&lt; 2995 g</b> )	56,4% (22)	43,6% (17)	38,5% (15)	17,9% (7)	43,6% (17)
<b>Modelo II</b> ( <b>2995 g – 3370 g</b> )	70,0% (27)	30,0% (12)	5,1% (2)	17,9% (7)	76,9%(30)
<b>Modelo III</b> ( <b>&gt;3370 g</b> )	70,0% (27)	30,0% (12)	22,5% (9)	10,0% (4)	67,5% (26)

**Tabela 3** – Resultados obtidos utilizando os dados de teste com os modelos neurais.

Modelos Neurais	Critério I		Critério II		
	Acertos	Erros	Acertos	Erros	Incertezas
<b>Modelo I</b> ( <b>&lt; 2995 g</b> )	63,2% (25)	36,8% (14)	60,5% (24)	26,3% (10)	13,2% (5)
<b>Modelo II</b> ( <b>2995 g – 3370 g</b> )	57,9% (23)	42,1% (16)	50,0% (20)	23,7% (9)	26,3%(10)
<b>Modelo III</b> ( <b>&gt;3370 g</b> )	65,8% (26)	34,2% (13)	44,8% (17)	15,8% (6)	39,5% (15)

Pelo critério I os resultados obtidos com os modelos *neuro-fuzzy* foram um pouco superiores aos resultados dos modelos neurais. Porém pelo critério II as taxas de acerto dos modelos neurais foram bem superiores às taxas dos modelos *neuro-fuzzy* visto que as taxas de incerteza exibiram comportamento contrário e as taxas de erro não foram muito diferentes para as duas técnicas.

Trabalhos futuros deverão explicar porque pelo critério II as taxas de incertezas dos modelos *neuro-fuzzy* foram bem superiores às dos modelos neurais.

Comparando os resultados obtidos com os dados de teste com os dados de treinamento, mostrados nas tabelas 4 e 5, percebemos que as taxas de acertos foram bastante parecidas, confirmando que a metodologia de seleção dos dados da seção 6 foi bastante eficaz.

**Tabela 4** – Resultados obtidos utilizando os dados de treinamento com os modelos *Neuro-Fuzzy*.

<b>Modelos <i>Neuro-Fuzzy</i></b>	<b>Critério I</b>		<b>Critério II</b>		
	<b>Acertos</b>	<b>Erros</b>	<b>Acertos</b>	<b>Erros</b>	<b>Incertezas</b>
<b>Modelo I</b> <b>(&lt; 2995 g)</b>	81,2% (189)	18,8% (55)	49,3% (144)	2,0% (6)	48,6 (142)
<b>Modelo II</b> <b>(2995 g – 3370 g)</b>	72,3% (211)	27,7% (81)	19,5% (57)	2,4% (7)	78,1% (228)
<b>Modelo III</b> <b>(&gt;3370 g)</b>	76,3% (223)	23,7% (69)	30,5% (89)	2,0% (6)	67,5% (197)

**Tabela 5** – Resultados obtidos utilizando os dados de treinamento com os modelos neurais.

<b>Modelos Neurais</b>	<b>Critério I</b>		<b>Critério II</b>		
	<b>Acertos</b>	<b>Erros</b>	<b>Acertos</b>	<b>Erros</b>	<b>Incertezas</b>
<b>Modelo I</b> <b>(&lt; 2995 g)</b>	64,6% (189)	35,4% (103)	58,3% (170)	29,7% (87)	12,0% (35)
<b>Modelo II</b> <b>(2995 g – 3370 g)</b>	61,7% (180)	38,3% (112)	49,6% (145)	25,9% (75)	24,5%(72)
<b>Modelo III</b> <b>(&gt;3370 g)</b>	61,2% (179)	38,8% (113)	44,9% (131)	25,4% (74)	29,7% (87)

Os resultados obtidos com os 3 modelos de Redes de *Kohonen*, explicados na seção 5, estão mostrados nas tabelas 4,5 e 6. O conjunto de teste foi utilizado para efeito de contagem do número de acertos, erros e incertezas de cada modelo.

**Tabela 4** – Resultados obtidos com o modelo I de Redes de *Kohonen*

<b>Modelo I</b> <b>(&lt; 2995 g)</b>	<b>Acertos</b>	<b>Erros</b>	<b>Incertezas</b>
<b>Dados da Faixa I</b> <b>(&lt; 2995 g)</b>	21	17	0
<b>Dados da Faixa II</b> <b>(2995 g – 3370 g)</b>	13	6	0
<b>Dados da Faixa III</b> <b>(&gt;3370 g)</b>	13	7	0
<b>TOTAL</b>	61% (47)	39% (30)	0% (0)

**Tabela 5** – Resultados obtidos com o modelo II de Redes de Kohonen

<b>Modelo II (2995 g – 3370 g)</b>	<b>Acertos</b>	<b>Erros</b>	<b>Incertezas</b>
<b>Dados da Faixa I (&lt; 2995 g)</b>	11	8	0
<b>Dados da Faixa II (2995 g – 3370 g)</b>	26	12	0
<b>Dados da Faixa III (&gt;3370 g)</b>	12	8	0
<b>TOTAL</b>	64% (49)	36% (28)	0% (0)

**Tabela 6** – Resultados obtidos com o modelo III de Redes de Kohonen

<b>Modelo III (&gt;3370 g)</b>	<b>Acertos</b>	<b>Erros</b>	<b>Incertezas</b>
<b>Dados da Faixa I (&lt; 2995 g)</b>	13	6	0
<b>Dados da Faixa II (2995 g – 3370 g)</b>	11	8	0
<b>Dados da Faixa III (&gt;3370 g)</b>	21	18	0
<b>TOTAL</b>	58% (45)	42% (32)	0% (0)

Os modelos de Redes de *Kohonen* não obtiveram os mesmos índices de acerto das outras duas técnicas, contudo pode-se perceber que em todos os três modelos os acertos superaram os erros na previsão de todos os tipos de dados, mostrando que esta técnica também foi capaz de relacionar os dados de entrada da mãe com o peso do recém-nascido.

Uma possível causa do desempenho inferior das Redes de *Kohonen* seria a ausência de amostras confiáveis de cada faixa, necessárias na fase de rotulagem dos clusters. Seria necessária uma quantidade de dados muito grande para obtermos através do uso da média dos parâmetros uma amostra representativa de cada faixa de peso.

## 8. Conclusão

Este estudo mostra que a aplicação de técnicas de desenvolvimento de modelos baseadas em Inteligência Artificial a partir de dados reais podem ser uma ferramenta importante na tomada de decisões. Para o caso específico da previsão do peso do recém-nascido, os resultados obtidos foram bastante promissores com taxas médias de 65% de acerto para as técnicas com aprendizado supervisionado e de 60% para as Redes de *Kohonen*, apesar da grande dificuldade do problema abordado visto que a relação entrada-saída dos dados não está quantitativamente comprovada.

Futuros trabalhos deverão:

- 1) aumentar a base de dados coletados,
- 2) apresentar novas propostas de entradas para os modelos e realizar uma nova coleta de dados,
- 3) investigar outras técnicas de modelagem como a técnica de “hints” [7] e a técnica de Composição de Especialistas Locais (MLM - *Multiple Local Models*) [8],
- 4) investigar técnicas de extração do conhecimento adquirido pelos modelos.



## 9. Referências Bibliográficas

- [1] HECHT-NIELSEN, R. Neurocomputing. Addison-Wesley. Massachusetts, EUA. 1991.
- [2] HOPFIELD, J. J. Neural Networks end Physical Systems with Emergent Collective Computational Abilities. *Proceedings of National Academy of Sciences*, Washington, v.79, p.2554-2558, Apr. 1982.
- [3] KOHONEN, T. Clustering Taxonomy and Topological Maps of Patterns. In: *International Conference on Pattern Recognition*, 6., 1982. Proceedings...[S.l:s.n.], 1982. P.114-128.
- [4] Nascimento Jr., C. L.; Yoneyama, T.; *Inteligência Artificial em Controle e Automação*, Edgard Blücher, São Paulo, 2000 ([http://www.ele.ita.br/ia\\_contaut](http://www.ele.ita.br/ia_contaut)).
- [5] Jang, J.-S. R.; Gulley, N.; *Fuzzy Logic Toolbox (for Use with MATLAB)*, The MathWorks, Inc., Natick, 2002 (<http://www.mathworks.com>).
- [6] Jang, J.-S. R.; “ANFIS: Adaptive-Network-Based Fuzzy Inference System”, IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665-685, Maio/Junho 1993.
- [7] Abu-Mostafa, Y. S.; “*Learning from Hints in Neural Networks*”, Journal of Complexity, vol. 6, pp. 192-198, Junho 1990.
- [8] de Melo, B.; *Previsão de Séries Temporais usando Modelos de Composição de Especialistas Locais*, Tese de Mestrado, PG-EAM/P, Instituto Tecnológico de Aeronáutica, 2003.
- [9] Guimarães, U. K. O.; *Coordenação Visual-Motora de um Braço Robótico em 3D Usando Redes de Kohonen*, Trabalho de Graduação, Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, 1998.
- [10] Nigam, V.; Nascimento Jr., C.L., Nascimento; L.F.C.; *Estudo Comparativo da Aplicação de Técnicas de Inteligência Artificial para a Previsão da Faixa de Peso de Recém-Nascidos*, IX Encita, Instituto Tecnológico de Aeronáutica, 2003.

## FOLHA DE REGISTRO DO DOCUMENTO

1. CLASSIFICAÇÃO/TIPO  TC	2. DATA  10 de novembro de 2003	3. DOCUMENTO Nº  CTA/ITA-IEE/TC-004/2003	4. Nº DE PÁGINAS  41
5. TÍTULO E SUBTÍTULO: Estudo comparativo da aplicação de técnicas de inteligência artificial para a previsão da faixa de peso de recém-nascidos			
6. AUTOR(ES): <b>Vivek Nigam</b>			
7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES): Instituto Tecnológico de Aeronáutica. Divisão de Engenharia Eletrônica – ITA/IEE			
8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR: Redes Neurais; Neuro-Fuzzy; Redes de Kohonen; Previsões; Recém-Nascidos;			
9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO: Redes neurais; Sistemas nebulosos; Mapas auto-organizáveis (redes neurais); Modelos matemáticos; Análise numérica; Previsão; Inteligência artificial; Medicina; Controle			
10. APRESENTAÇÃO: <div style="text-align: right; margin-right: 50px;"><b>X Nacional</b></div> <div style="text-align: right; margin-right: 50px;"><b>Internacional</b></div> Trabalho de Graduação, ITA, São José dos Campos, 2003. 41 páginas.			
11. RESUMO: <p>Este trabalho descreve a aplicação de três técnicas de Inteligência Artificial, Sistemas Neuro-Fuzzy, Redes Neurais e Redes de Kohonen, para a criação de modelos matemáticos que possam ser usados algumas semanas antes do parto para a previsão da faixa de peso de recém-nascidos, a partir de certos dados da mãe. Estes modelos foram criados a partir dos dados coletados no ambulatório do Hospital Universitário da UNITAU durante os últimos três anos referentes a 550 nascimentos. As etapas para obtenção dos modelos de previsão foram: 1) seleção das variáveis de entrada para os modelos a partir dos dados coletados no ambulatório, 2) decisão sobre o número e os intervalos das faixas para o peso do recém-nascido, 3) separação dos dados coletados em dados de treinamento, validação e teste, 4) geração e testes dos modelos de previsão. Devido ao limitado número de dados disponíveis, o número de faixas para a classificação do peso do recém-nascido foi arbitrado em 3. Foi então desenvolvida uma metodologia para a definição dos intervalos destas 3 faixas. As 3 técnicas utilizadas então foram aplicadas para o desenvolvimento de modelos de previsão para cada uma das 3 faixas. As taxas de acerto medidas no conjunto de teste foram: 1) para a técnica Neuro-Fuzzy: de 56% a 70%, 2) para a técnica neural: de 63% a 66% e 3) Redes de Kohonen: de 58% a 64%.</p>			
12. GRAU DE SIGILO: <div style="display: flex; justify-content: space-between; margin-top: 10px;"> <span><b>(X) OSTENSIVO</b></span> <span><b>( ) RESERVADO</b></span> <span><b>( ) CONFIDENCIAL</b></span> <span><b>( ) SECRETO</b></span> </div>			