

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

**Vivek Nigam<sup>1</sup> (IC), Cairo L. Nascimento Jr.<sup>1</sup> (PQ), Luiz Fernando C. Nascimento<sup>2</sup> (PQ)**

<sup>1</sup>Instituto Tecnológico de Aeronáutica (ITA)  
Divisão de Engenharia Eletrônica  
12228-900 - São José dos Campos - SP  
E-mails: vivek@h8.ita.br, cairo@ita.br

<sup>2</sup>Universidade de Taubaté (UNITAU)  
Departamento de Medicina  
12030-180 - Taubaté - SP  
E-mail: lfcnascimento@uol.com.br

### RESUMO

*Este artigo descreve a aplicação de duas técnicas de Inteligência Artificial, Sistemas Neuro-Fuzzy e Redes Neurais, 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 2 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%.*

### ABSTRACT

*This article 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. Two techniques from the artificial intelligence field are used: neuro-fuzzy systems and artificial neural 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 2 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 and between 63% and 66% for the artificial neural network models.*

## 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 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 efetiva 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 duas destas técnicas de modelagem, Redes Neurais e Sistemas Neuro-Fuzzy, 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.

Os modelos considerados neste trabalho utiliza as seguintes 8 entradas: 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; e 8) numeração desta gestação (p. ex., 3<sup>a</sup> gestação).

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; e 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 muito 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) até 2295 g, 2) entre 2295 g e 3370 g, e 3) acima de 3370 g. Uma vez feita tal separação dos dados, para cada técnica utilizada (Sistemas Neuro-Fuzzy ou 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. Quando 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.

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: a) a estrutura dos modelos baseados em sistemas neuro-fuzzy e redes neurais, b) os métodos de obtenção dos parâmetros dos modelos, c) os critérios de avaliação dos resultados obtidos pela aplicação dos modelos.

## 2. MODELO NEURO-FUZZY

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. Uma função de pertinência mede o grau com que um certo valor numérico da variável  $x_i$  satisfaz o conceito expresso pela função de pertinência. Diferentemente da lógica convencional, a saída de uma função de pertinência na lógica fuzzy pode assumir qualquer valor no intervalo [0,1]. Usualmente funções do tipo triangulares, trapezoidais ou gaussianas são usadas como funções de pertinência.
2. A abrangência de cada modelo local é definida por uma regra do tipo:

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

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:

$$y = \frac{\sum_{i=1}^N z_i y_i}{\sum_{i=1}^N z_i} \quad (2)$$

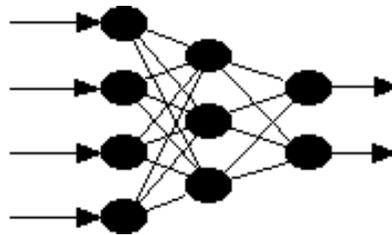
onde  $N$  denota o número de regras e  $z_i$  indica o coeficiente de ponderação da  $i$ -ésima regra. Tal

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* [1,2].

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 [2,3].

### 3. MODELO DE 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 1).



**Figura 1** – Estrutura do modelo baseado em Rede Neurais Artificiais (RNA) com 4 entradas, 3 unidades na camada escondida 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:

$$z_i = \tanh \left( \sum_{j=1}^8 w_j^{IH} x_j + bias_i^H \right), i = 1, \dots, NH \quad (3)$$

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

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. METODOLOGIA DE DESENVOLVIMENTO DOS MODELOS

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

- Conjunto dos dados de treinamento: é o conjunto de dados usado pelo algoritmo de treinamento para o ajuste dos parâmetros do modelo,
- 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

	Sub-conjunto I	Sub-conjunto II	Sub-conjunto III
<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 ser satisfatório, tanto no caso dos 3 modelos neuro-fuzzy como nos 3 modelos neurais, o erro medido sobre o conjunto de testes foi julgado como insatisfatório. 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:

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

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

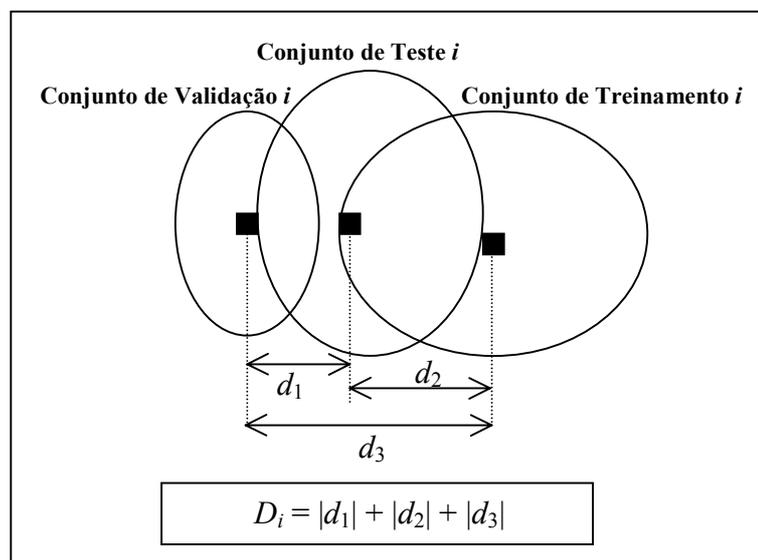
onde os índices  $d$  indicam as distâncias entre as médias vetoriais dos conjuntos de treinamento, validação e teste (conforme mostrado na figura 2) sendo que cada componente do vetor de entrada foi previamente normalizada linearmente entre  $-1$  e  $1$ , considerando todos os dados coletados. Com tal normalização o cálculo do módulo do vetor  $d$  independe do intervalo de cada componente do vetor de entrada  $x$ .

- 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.
- para a geração dos modelos neuro-fuzzy e neurais é então adotada a proposta de distribuição dos dados com o menor valor escalar  $S$ .

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]. 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 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).



**Figura 2** – Ilustração do cálculo dos índices de distribuição dos dados dos conjuntos de treinamento, teste e validação.

## 5. AVALIAÇÃO DOS RESULTADOS

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.

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 com os modelos *Neuro-Fuzzy*.**

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

**Tabela 3 – Resultados obtidos com os modelos neurais.**

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

## 6. 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, 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) investigar outras técnicas de modelagem como Redes de Kohonen [1], a técnica de “hints” [4] e a técnica de Composição de Especialistas Locais (MLM - *Multiple Local Models*) [5].

## AGRADECIMENTOS

Gostaríamos de agradecer a todas as mães que contribuíram com os dados de prontuário necessários para a realização deste estudo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] 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)).
- [2] Jang, J.-S. R.; Gulley, N.; *Fuzzy Logic Toolbox (for Use with MATLAB)*, The MathWorks, Inc., Natick, 2002 (<http://www.mathworks.com>).
- [3] 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.
- [4] Abu-Mostafa, Y. S.; “Learning from Hints in Neural Networks”, *Journal of Complexity*, vol. 6, pp. 192-198, Junho 1990.
- [5] 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.