Redes neurais e modelos de regressão para previsão de uso tecnológico em unidades de cuidados intensivos neonatais

NEURAL NETWORKS AND REGRESSION MODELS FOR PREDICTING TECHNOLOGY UTILIZATION IN NEONATE INTENSIVE CARE UNITS

João Domingos Scalon
Departamento de Ciências Exatas, Universidade Federal de Lavras

RESUMO

Modelos que conseguem prever corretamente o uso de tecnologias de saúde podem ser úteis para o planejamento de unidades de cuidados de saúde. No momento, os modelos de regressão são as únicas técnicas estatísticas usadas para esse fim. Nos últimos anos, observa-se um crescente uso de redes neurais para modelagem de problemas nos quais, tradicionalmente, são usados modelos de regressão. Este artigo compara o uso de redes neurais e modelos de regressão para previsão de uso de tecnologias de saúde em uma Unidade de Cuidados Intensivos Neonatais. Os resultados obtidos indicam que redes neurais podem não ser superiores aos modelos de regressão para prever o uso tecnológico, quando não existe uma explícita relação não linear entre as co-variáveis e o uso tecnológico.

PALAVRAS-CHAVES

Cuidados intensivos neonatais. Redes neurais. Regressão. Tecnologias de saúde.

INTRODUÇÃO

Clinical and managerial decision-making requires some degree of resource forecasting in order to guarantee the availability of essential technologies and resource optimization. Hence, statistical models for predicting the use of health technologies can be useful as an adjunct to management of a health unit in order to purchase new pieces of equipment, to plan a new unit or to restructure an existing one to match the epidemiological transitions (JENNET, 1986).

The statistical techniques most often used for predicting the relationship of technology utilization to multiple predictor variables are both linear and logistic regressions (ALMEIDA et al., 1991; SCALON; FREIRE; CUNHA, 1998; SCALON; VIDAL MELO; PANERAI, 1996). In regression models, a functional form is imposed on the data. In the case of linear and logistic regressions, this assumption is that the outcome is related to a linear combination of the predictor variables. Thus, linear regression models have limitations in some clinical situations, as the relationship between an outcome and predictor variable may be non-linear (DRAPER; SMITH, 1998).

Although, the presence of non-linear relationships can be modeled by various modifications of the regression models, the nature of the relationship are not always known a priori. An alternate approach to overcome these limitations of the linear regression models is not to assume any functional relationship and let the data speak for themselves. This is the basis of the power of artificial neural networks once they provide a universal approximation of function (HORNIK, STINCHCOMBE; WHITE, 1989). This fact explains why this approach is a promising tool to estimate health technology utilization.

Neural networks have been frequently applied to several different situations such as to diagnose acute illness (ABALAVANAN; CARLO, 2001; BAXT, 1994) and to predict clinical outcomes (mortality) as a function of technological use (CHACON; MAUREIRA, 2004). A body of theoretical work suggests that neural networks have the ability to consistently match or exceed the performance of other traditional regression methods (HORNIK; STINCHCOMBE; WHITE, 1989) while other papers suggest that neural networks may not be superior to regression analysis (ABALAVANAN; CARLO, 2001; TU, 1996). The purpose of this paper is to compare the predictive performance of neural networks and regression models for predicting the use of health technologies in a Neonate Intensive Care Unit.

METHODS

The data set used to build the models came from a cohort of sequential admissions to the NICU of Fernandes Figueira Institute, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil, over a period of 20 months. Patients were rejected if they were transferred into the hospital from elsewhere (69 cases), stayed for less than 24 hours (27 cases), or had incomplete medical records (223 cases). These criteria resulted in 193 records with complete data for analysis.

Six neonatal technologies, based on their clinical importance for neonatal care, were selected and their use was measured. The total number of utilization episodes in each neonate measured the use for gasometry (GASO). The duration of utilization, in number of days, measured the use for continuous positive airway pressure (CPAP) and mechanical ventilation (VENTIL). The use of echocardiography (ECHO), bicarbonate (BICAR) and adrenaline (ADRE) was approximated by a binary variable reflecting either the use or non-use in each neonate of these technologies.

Sixteen prognostic variables for the technology utilization were extracted, retrospectively, from each patient chart, as follows: AGE - age of the mother (in years); GA - gestational age (weeks); WEIGHT - birth weight (grams); APG1 - Apgar score at 1 minute post delivery; APG5 - Apgar score at 5 minutes post delivery; PRE - prenatal consultations; ABOR - abortion in previous gestations; HYPER - hypertensive disorder; ABP - abruptio placentae; CEPHA - non-cephalic delivery, RESP - respiratory complications, diagnosed at admission (d.a.a.); BLOOD - hematological complications, d.a.a; INFEC - infectious disease of the newborn, d.a.a; CARDIO - cardiovascular complications, d.a.a; CYANO - cyanotic newborn, d.a.a and NEURO - neurological complications, d.a.a. The first six variables (AGE to APG5) are continuous while the other variables are binary reflecting either the presence or absence of specific risk states. All the quantitative variables were normalized on the interval [0, 1].

The database with 193 neonates was split into two groups: the adjustment group, composed of 148 neonates who were sequentially admitted during the first 15 months of the study and the test group, composed of 45 neonates admitted in the last five months. The adjustment group was used to fit the models and the test group was used to evaluate the differences, in terms of prediction performance, among the models.

Multiple linear regression (MLR) was applied to assess the linear relationship between the utilization of GASO, CPAP and VENTIL and the prognostic variables. The MLR model has the form $Y = \beta X$, where Y is the vector of the technology utilization, X is the matrix of prognostic variables, and β is the vector of regression coefficients. The regression coefficients were estimated by the ordinary least squares (DRAPER; SMITH, 1998). A forward stepwise process was adopted to choose the variables to be included in the models. The limits for inclusion and exclusion of variables in the models were selected as the levels of probability $p = 0.10$ and $p = 0.15$, respectively (BENDEL; AFIFI, 1977). The goodness-of-fit of the MLR model was assessed by using the adjusted coefficient of determination (R^2), residual analysis, t and F-tests (DRAPER; SMITH, 1998).

Multiple logistic regression (LOG) is a popular statistical modeling technique in which the probability of use of a dichotomous outcome (Y) of each technology (ECHO, ADRE and BICAR) is related to a set of potentially prognostic variables (X) in the form $\log [p/(1-p)]= \beta X$ where p is the probability of the outcome, β are the coefficients associated which each prognostic variable. The logarithm of the odds $\{\log[p/(1 - p)]\}$ is related in a linear manner to the potentially prognostic variables. The coefficients β and the significance of the model prognostic variables were estimated by the maximum likelihood method. Goodness-of-fit of the models was assessed by using the total precision of classification, t and Hosmer-Lemeshow tests (HOSMER; LEMESHOW, 2000).

Neural networks (NN) include a sequence of input, hidden and output layers interconnected in many different ways. The hidden layers allow the network to generate numerous relationships between the inputs and outputs so that the desired outputs can be produced (learning) using a given set of inputs (HASTIE, 2005, RIPLEY, 1996). The number of units in the input layer is determined by the number of prognostic variables selected by the stepwise procedure in the regression models. It is used a single output unit representing the use of the actual technology. A fully connected multilayer feed-forward NN was used, since this type of network provides a flexible way to generalize linear regression functions (RIPLEY, 1996). If we denote the matrix of p units in the input layer by X, and the output unit by Y, then this model can be written more traditionally as $z_j = \sigma\ (\alpha_{j0} + \alpha_j^t X)$, $j = 1, \ldots, m$, $\hat{y} = f\ (\beta_0+\beta^t Z)$. The activation function σ is used to introduce nonlinearity at the hidden layer, and it was chosen to be a sigmoid given by $\sigma\ (z) = 1/(1+e^{-z})$. The parameters α_{ji} and β_j are known as weights, and define linear combinations of the input matrix X and hidden unit output vector Z, respectively. α_{j0} and β_0 are known as biases. The function f allows a final transformation of the output, and it was chosen as f (v) = v for models with quantitative response and f (v) = 1/[1+exp(-v)] for models with responses that lie in [0, 1]. In this model, each hidden unit can be thought of as a nonlinear basis function that creates a new derived variable z_j from a linear combination of the inputs. The responses are then regressed on these transformed data z_j either linearly or via logistic regression. Livingstone; Manallach; Tetko et al. (1997) suggest that, in order to determine the number of units in the hidden layer, the rate between the number of units in the input layer and hidden layer should be, approximately, equal to two. The learning rule used was back propagation of errors, which adjusts the parameters of the network over repeated cycles to reduce the overall error. Least squares were used to learn (fit) the parameters (weights) (HASTIE, 2005). The weight decay parameter (10^{-4} - 10^{-2}) was used to help the optimization process and to avoid over fitting (RIPLEY, 1996). Although, NN does not allow the evaluation of the functional relationship by testing either the individual weights or the global model for statistical significance, it is always possible to calculate some statistical measures that provide a kind of goodness-of-fit of the models. The adjusted coefficient of determination (R^2) was adopted for the technologies GASO, CPAP and VENTIL and the total precision of classification was adopted for the technologies ADRE, ECHO and BICAR.

The prediction performances of the MLR and NN models were compared by constructing a (1-α)% confidence interval for the estimated mean intensity of use. This interval was evaluated by the equation $\overline{Y} \pm t_c\sqrt{S_2 / N}$, where N is the number of neonates in the test group, t_c is the α% critical value for a two sided t distribution on N-1 degrees of freedom, $\overline{Y}$ and S^2 are the estimated mean and variance of the technological use for the test group, respectively. The confidence interval is then compared with the actual mean technology utilization in the test group (SCALON; VIDAL MELO; PANERAI, 1996). The prediction performance of LOG and NN were assessed by using the following measures: sensitivity, specificity and total precision of classification (FLEISS; LEVIN; PAIK, 2003). Both LOG and NN generate an output that is a continuous probability and therefore a cut-off point of p = 0.5 was adopted, implying that values of $p < 0.5$ were classified as "no use" and $p > 0.5$ were interpreted as "use" of the technology. All statistical analyses were performed using functions either available or developed in R (R DEVELOPMENT CORE TEAM, 2006).

Results

The main clinical-epidemiological characteristics of the 193 cases were 105 (54%) male and 88 (46%) female. The mean ± standard deviation of birth weight of the population studied was 2278 ± 905 g (range: 740 to 4500 g) and 110 (57 %) of the neonates were in the low birth weight group (< 2500 g). The gestational age was 36 ± 3 weeks. The length of stay was 9 ± 2 days. Respiratory problems at admission occurred in 114 (59%) of the patients and 29 (15%) died during the stay. The average percent utilization of the six technologies for the 193 neonates was as follows: GASO (68 %), CPAP (25%), VENTIL (18%), BICAR (9%), ADRE (6%), and ECHO (6%).

Tables 1 to 4 show the results of NN and regression models. Tables 5 to 6 show comparisons of prediction performances between NN and regression models. It can be observed that all technologies have demonstrated the superior predictability (actual mean use inside the confidence intervals) of MLR over NN. There are no clear differences among the measures of predictabilities between LOG and NN.

Table 1 – Estimated coefficients and adjusted coefficient of determination (R^2) of the multiple liner regression models for the use of GASO, CPAP and VENTIL for 148 newborns of the adjustment group. The p- values are given in parentheses.

VARIABLES	GASO	CPAP	VENTIL
CONSTANT	0.191 (0.000)	0.422 (0.000)	0.148 (0.001)
GA	-0.206 (0.000)	-0.287(0.000)	-0.135 (0.004)
APG1	-0.147 (0.001)		-0.076 (0.072)
APG5		-0.194(0.001)	
RESP	0.062 (0.011)	0.104 (0.000)	
INFEC	0.073 (0.058)		
AGE			-0.108 (0.026)
ABOR	0.051 (0.093)		0.077 (0.001)
PRE		-0.116(0.084)	
HYPER	-0.058 (0.024)		
CEPHA	-0.085 (0.006)	-0.072(0.035)	-0.063 (0.029)
CYANO	0.069 (0.005)		0.043 (0.062)
F	7.568 (0.000)	12.77 (0.000)	5.130 (0.000)
R^2	0.30 (0.000)	0.31 (0.000)	0.18 (0.000)

Table 2 – Final weights of the hidden layer units (H) and output (O) of the neural networks models for the utilization of GASO, CPAP and VENTIL for 148 newborns of the adjustment group. Adjusted coefficient of determination (R^2) is the goodness-of-fit measure.

	GASO					CPAP				VENTIL			
VARIABLES	H1	H2	H3	H4	O	H1	H2	H3	O	H1	H2	H3	O
BIAS	2.46	-0.17	-0.16	1.99	0.04	-1.21	-1.05	-0.17	2.02	-3.68	3.83	2.37	-3.49
GA	-7.91	-6.63	-9.15	-7.17	5.40								
					-4.54				-3.64				5.95
					4.00				0.82				-2.51
APG1	-2.32	-1.58	-2.89	-1.97	4.47					-2.96	2.64	1.71	
APG5						3.58	-2.21	-0.82					
RESP	2.03	1.69	2.61	1.96		-2.40	1.24	0.62					
INFEC	0.31	0.21	0.38	0.35									
AGE										-0.73	-0.87	-0.72	-0,73
ABOR	0.93	0.42	0.84	-1.64									
PRE													
HYPER	-1.53	-0.91	-1.92	-1.59									
CEPHA	-2.43	-1.66	-3.81	-2.88		0.26	0.52	0.32		-3.12	2.14	0.69	
CYANO	-5.42	-1.40	2.18	-1.35						4.41	-2.79	-0.33	
R^2	0.74 (0.000)					0.59 (0.000)				0.58 (0.000)			

Table 3 –Estimated coefficients of the logistic regression models for the utilization of ECHO, BICAR and ADRE in the adjustment group of 148 newborns. TPC is the total precision of classification. HL is the level of significance of the Hosmer-Lemeshow test. P-values for the estimated coefficients are given in parentheses.

VARIABLES	ECHO	BICAR	ADRE
CONSTANT	-10.977 (0.001)	-2.312 (0.170)	-0.991 (0.009)
APGAR5	-7.538 (0.023)	-2.570 (0.011)	-3.814 (0.009)
CARDIO	2.905 (0.003)		
ABORTION		1.177 (0.063)	
ABP			1.908 (0.053)
NEURO		1.745 0.013)	
CYANO	2.214 (0.072)		
HL	0.610	0.700	0.320
TPC	0.96	0.91	0.95

Table 4 – Final weights of the hidden layer units (H) and output (O) of the neural networks models for the utilization of ECHO, BICAR and ADRE for 148 newborns of the adjustment group. TPC is the total precision of classification for measuring goodness-of-fit.

VARIAB	ECHO			BICAR			ADRE		
	H1	H2	O	H1	H2	O	H1	H2	O
BIAS	2.41	3.80	2.28	4.55	3.36	3.37	0.60	-1.97	2.70
CYANO	-2.40	0.56	-5.69			-6.50			-3.75
NEURO			-7.76	-2.38	1.85	-6.32			-4.94
APG5	-2.74	-4.96		-3.65	-6.54				
ABP							-2.89	-0.52	
BLOOD									
CARDI	1.67	-5.56							
ABOR				-2.91	3.51				
TPC	0.97			0.95			0.97		

Table 5 – Observed mean use and 95% confidence interval for the estimated mean of use from neural networks (NN) and multiple linear regression (MLR) for the utilization GASO, CPAP and VENTIL in the test group of 45 newborns.

TECNOLOGIES	OBSERVED	MLR	NN
GASO	3.4	(0.50; 7.72)	(0.78; 2.84)
CPAP	0.73	(-0.47; -0.23)	(-0.11; 0.24)
VENTIL	1.07	(0.34; 1.25)	(-0.14; 0.60)

Table 6 – Sensitivity, specificity and total precision of classification (TPC) of neural networks (NN) and logistic regression (LOG) for the utilization of ECHO, BICAR and ADRE in the test group of 45 newborns.

MEASURES	ECHO		BICAR		ADRE	
	NN	LOG	NN	LOG	NN	LOG
SENSITIVITY	0.25	0.43	0	0.71	0.25	0.12
SPECIFICITY	0.98	0.98	0.95	1.00	0.98	1.00
TPC	0.97	0.96	0.95	0.91	0.97	0.95

Discussion

Clinicians have traditionally carried out predictions of the utilization of health technologies for neonates of risk. Unfortunately, clinicians are not always aware of the important relation between variables because the complexity of the system may be beyond the analytical capabilities of a physician (BAXT, 1994). Thus, methodologies based on regression models have been proposed for predicting technology utilization in NICU and they are proved potentially useful for health care planning (ALMEIDA et al., 1991; SCALON; VIDAL MELO; PANERAI, 1996; SCALON; FREIRE; CUNHA, 1998).

Regression models, in general, would correctly provide predictions of technology utilization. However, it is always possible that these traditional methods will break down because of the influence of a potential complex multitude of clinical variables. It is advocated that NN allow the recognition of patterns of complex relationship within data sets that may not be detected with conventional regression models. Thus, it is not surprise that NN have been applied to the analysis of data in many different settings, including diagnosis and prediction of outcome in medicine (ABALAVANAN; CARLO, 2001; BAXT, 1994; CHACON; MAUREIRA, 2004). The present work has addressed the comparison between regression models and NN for predicting the use of health technologies in a NICU.

We started the analysis by choosing the appropriate regression model based on the measure of technology utilization. MLR is the appropriate tool for modeling technologies with quantitative measure (DRAPER; SMITH, 1998) while LOG is the appropriate method of choice for technologies with binary use (HOSMER; LEMESHOW, 2000). Results presented in Table 1 suggest that all MLR models are highly statistical significant as indicated by the F-test ($p < 0.001$) and the regression coefficients for all prognostic variables are significantly different from zero ($p < 0.10$). About 30% of the variance (the value given by (R^2) in technology utilization can be attributed to the prognostic variables selected by the stepwise process. Table 1 also shows evidence that there are a correlation between the severity of initial status and utilization of technologies. For example, the use of GASO tends to decrease as the gestational age and Apgar score at one-minute post delivery increase. The use of CPAP is related to the presence of respiratory complications diagnosed at admission. The negative parameters for gestational age and Apgar scores indicate that the use of technologies decreases with higher gestational ages and Apgar scores. These results suggest that the MLR models are quite compatible with the clinical practice.

On the other hand, Table 2 shows that NN are a black box. Although there are weights (without statistical significances) associated with each prognostic variable, they are, generally, not so useful in explaining the level of contribution of each variable (RIPLEY, 1996). However, NN were capable to explain much more variability of technologies. NN presented larger values of the TPC for all technologies. LOG models presented larger values of sensitivity in two out of three technologies. In summary, these results show that the predictive performance of both models was essentially the same for predicting the probability of use of such technologies within the NICU.

It can also be observed in Table 6 that both methodologies provided values of specificity that are higher than the values of sensitivity for all technologies. The reason why this occurs is that we are modeling technologies with low intensity of use, that is, the majority of the neonates do not use this type of technology and, therefore, models tend to predict the "non-use" of the technology. This means that even in the case of these sporadically used technologies it is possible to obtain models with good predictive performances.

The results presented above suggest that NN present worse (or equal) predictive performance for predicting the use of health technologies within the NICU. However, it is either necessary to point out some important aspects that health managers have to keep in mind when assessing the feasibility of predicting technology utilization by using NN or regression models.

First, the better accuracy performance of regression models may be explained by the fact that the technological use prediction problem at hand appears to not exhibit non-linearity with the prognostic variables. One reason this occurs is that the majority of the prognostic variables selected by the stepwise procedure is binary. This means that their contribution to the model must be on a linear scale. Trying to model them in a different way will not contribute to the predictive performance of the model.

Second, a single cohort of newborns with a relative small sample of complete medical records from a specific NICU was investigated and, therefore, this could cause problem to the NN. It is well know that neural networks are best trained with larger data sets, preferably with thousands of cases. Since this study has relative small groups (adjustment and test), it would be expected to be better off using regression models (HASTIE, 2005). This could explain the bad predictive performance of NN in the test group with just 45 cases.

Third, while stepwise is the common method to select variables that may be relevant for regression models, an equivalent method to find relevant variables for NN do not exist, although some pruning algorithms have been shown to improve predictive ability for NN by the removal of redundant input variables (TETKO; VILLA; LIVINGSTONE, 1996). Thus, the use of a particular method for the identification of the best-input variables for NN might improve their predictive performances.

Finally, it is worth to point out that, in general, when NN gets bad results it is possible to have three other causes besides those discussed above: not sufficient input/output patterns to train neural network (many times, it is impossible to get more patterns), one mean square error greater than necessary (this aspect is difficult to resolve too in some cases) or one incorrect number of neurons in the hidden layer net (HORNIK; STINCHCOMBE; WHITE, 1989; TETKO; VILLA; LIVINGSTONE, 1996). These three aspects can be gotten only empirically and these are the major difficulties to use neural network. When these three problems are resolved simultaneously, a very difficult task, the interpolation gets by NN is as near as possible of perfect.

Conclusions

In this paper, the potential of NN were evaluated as an alternative to traditional regression models for predicting technology utilization in a NICU. Results indicate that the predictive performance of NN is worse than that of regression models. The advantage of NN is their ability to approximate any continuous function and one does not have to guess the functional form of the model. The disadvantage is that is difficult to interpret NN parameters. In regression models, we are able not only to interpret the coefficients in relation to the clinical aspects of the risk factors but also to test both the individual coefficients and the whole model for statistical significance. Another disadvantage of NN is that the convergence to a solution can be slow and depends on the NN's parameters such as initial starting weights and number of hidden units, to name but a few. Thus, with no complex interaction or non-linearity, the linear additive structure of regression models would be more appropriate for our problem, despite the observed skewness of technology utilization in some technologies.

ABSTRACT

Models that accurately predict technology utilization can be potentially beneficial for health care planning. At present, available statistical methodologies for predicting technology utilization rely on regression models. Over the last few years, neural networks have being applied as a tool for modeling in areas where regression models are traditionally employed. This paper compares neural networks and regression models for predicting the use of health technologies in a Neonate Intensive Care Unit. Results indicate that neural networks may not be superior to regression models for predicting the use of health technology when there is no clear non-linear relationship between prognostic variables and technology utilization.

KEY-WORDS

Health technology. Neonatal intensive care. Neural networks. Regression models.

ACKNOWLEDGEMENTS

This work was supported by a grant from Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG). The author would like to thank the reviewer for the comments and suggestions which have improved the paper and Dra. Maria Elizabeth Lopes Moreira (IFF/FIOCRUZ) and Dr. Sergio Miranda Freire (LAMPADA/UERJ) for their suggestions during the planning stages of this study.

REFERENCES

ABALAVANAN, N.; CARLO, W. Comparison of the prediction of extremely low birth weight neonatal mortality by regression analysis and neural networks. *Early Human Development*, v. 65, n. 2, p. 123-137, 2001.

ALMEIDA, R. T. et al. Analysis of the use of multiple technologies in neonatal care. *International Journal of Technology Assessment in Health Care*, v. 7, n. 1, p. 22–29, 1991.

BAXT, W. G. Neural network trained to identify the presence of myocardial infarction bases some decisions on clinical associations that differ from accepted clinical teaching. *Medical Decision Making*, v. 14, n. 2, p. 217-222, 1994.

BENDEL, R. B.; AFIFI, A. A. Comparison of stopping rules in forward stepwise regression. *Journal of the American Statistical Association*, v. 72, n. 1, p. 46 –53, 1977.

CHACON, M.; MAUREIRA, B. Causal networks for modeling health technology utilization in intensive care units. *Lecture Notes in Computer Science*, v. 3287, n.1, p. 645-653, 2004.

DRAPER, N. R.; SMITH, H. *Applied regression analysis.* New York: John Wiley & Sons, 1998. 736p.

FLEISS, J. L.; LEVIN, B.; PAIK, M. C. *Statistical methods for rates and proportions*. New York: John Willey & Sons, 2003. 800p.

HASTIE, T. Neural networks. In: ARMITAGE, P. (Edit). *Encyclopedia of biostatistics*. New York: John Wiley & Sons, 2005. 6100p.

HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximations. *Neural Networks*, v. 2, n. 5, p. 359–366, 1989.

HOSMER, D. W.; LEMESHOW, S. *Applied logistic regression*. New York: John Willey & Sons, 2000. 392p.

JENNET, B. *High technology medicine: benefits and burdens*. Oxford: Oxford University Press, 1986. 318p.

LIVINGSTONE, D. J.; MANALLACK, D. T.; TETKO, I. V. Data modelling with neural networks: advantages and limitations. *Journal of Computer-Aided Molecular Design,* v. 11, n. 2, p. 135–142, 1997

R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Version 2.3.1 Vienna: R Foundation for Statistical Computing, 2006. Disponível em: www.r-project.org. Acesso em: 15 maio 2007.

RIPLEY, B. D. *Pattern recognition and neural networks*. Cambridge: Cambridge University Press, 1996. 504p.

SCALON, J. D., VIDAL MELO, M. F., PANERAI, R. B. A method for predicting the utilization of medical technology in neonatal care. *Technology and Health Care,* v. 3, n. 4, p. 273–278, 1996.

SCALON, J. D., FREIRE, S. M., CUNHA, T. A. Validation of models for predicting the use of health technologies. *Medical Decision Making,* v. 18, n. 3, p. 311- 319, 1998.

TETKO, I. V., VILLA, A. E. P., LIVINGSTONE, D. J. Neural networks studies: 2. Variable selection. *Journal of Chemical Informatics and Computing Science,* v. 36, n. 4, p. 794-803, 1996.

TU, J. V. Advantages and disadvantages of using artificial neural network versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology,* v. 49, n. 11, p. 1225-1231, 1996.

João Domingos Scalon
Departamento de Ciências Exatas Universidade Federal de Lavras
Lavras – MG
Caixa Postal 3037
CEP 37200-000
Fone: (35) 3829-1643
e-mail: scalon@ufla.br

Tramitação

Artigo recebido em:	08/06/2006
Aceito para publicação em:	16/05/2007