Biomedical Statistics and Informatics
2022; 7(1): 7-11
http://www.sciencepublishinggroup.com/j/bsi
doi: 10.11648/j.bsi.20220701.12
ISSN: 2578-871X (Print); ISSN: 2578-8728 (Online)

Sensitivity Analysis by Variables in the Classification of Renal Insufficiency Using Artificial Neural Networks VS Multinomial Logistic Regression

Elena Martín Pérez[1, *], Quintín Martín Martín[2]

[1]Institute of Legal Medicine of Zamora, Zamora, Spain
[2]Department of Statistics, University of Salamanca, Salamanca, Spain

Email address:
elenamrtnprz@gmail.com (E. M. Pérez), qmm@usal.es (Q. M. Martín)
[*]Corresponding author

To cite this article:
Elena Martín Pérez, Quintín Martín Martín. Sensitivity Analysis by Variables in the Classification of Renal Insufficiency Using Artificial Neural Networks VS Multinomial Logistic Regression. *Biomedical Statistics and Informatics.* Vol. 7, No. 1, 2022, pp. 7-11.
doi: 10.11648/j.bsi.20220701.12

Received: February 18, 2022; **Accepted**: March 14, 2022; **Published**: March 23, 2022

Abstract: Purpose: The aim of the research is to see to what extent Neural Networks (non-parametric technique) could be used as opposed to Multinomial Logistic Regression (parametric technique) in the analysis of sensitivity by variable in the classification of "Renal insufficiency" (three categories: NOT, MODERATE and ADVANCED). The analysis of sensitivity by variable, in our case, consists of eliminating from the model the three most influential variables (one by one): Blood creatine, Urine creatine and Urea, in that order. Once a variable is removed from the model, it will not reenter the model. Methods: This study collects data from the University Hospital of Salamanca (Spain), configuring a file of renal insufficiency data with 184 cases and 9 variables. First, we do descriptive-exploratory analysis of data for renal insufficiency data, obtained experimentally through a pilot survey. The comparison between ANNs and MLR is carried out by classification in the categories NOT, MODERATE and ADVANCED. Results: The descriptive-exploratory analysis of data shows the high value of the Coefficient of variation, Kurtosis and Skewness of the variables Blood creatine, Urine creatine and Urea, all of them well above 66%, 10 and 2.70, respectively. The study shows that when all the variables in the model are considered, the highest classification percentage concerns the Multinomial Logistic Regression (MLR), while, for the analysis of sensitivity by variables, the classification percentages are favourable to the Artificial Neural Network model (ANN). Conclusions: The joint classification percentages in the analysis of sensitivity by variables are favourable to the artificial neural network model (perceptron). That is, the non-parametric technique (ANNs) would surpass the parametric technique (MLR) in the classification of a patient's renal insufficiency.

Keywords: Renal Insufficiency, Exploratory Data Analysis, Artificial Neural Network, Multinomial Logistic Regression

1. Introduction

The classification or prediction of diseases through multinomial logistic regression (MLR) and Artificial Neural Networks (ANNs) has been discussed in several articles [1-7]. Our work wants to contribute in this line by providing a sensitivity analysis by variable with the comparison between both methods (ANNs and MLR). The aim of this paper is to highlight the usefulness of both methods (ANNs and MLR) in the classification of renal insufficiency.

The application of neural networks to the field of Medicine has been improving due to the increase in research in the field of artificial neural networks with new networks adapted to each field of study [8-15]. We will focus on the use of the multilayer perceptron and the multinomial logistic regression for the classification of the renal insufficiency levels NOT, MODERATE and ADVANCED.

The use of artificial neural networks (ANNs) in the prediction-classification of the state of renal insufficiency (NOT, MODERATE and ADVANCED) of patients is a heuristic search that usually gives good results in the field of Medicine in general.

First, we do a descriptive-exploratory analysis of the data in order to have information on the distribution of the data, in particular the "Age" with respect to the variable "Renal insufficiency" with three categories: NOT, MODERATE and ADVANCED.

By the appropriate combination of the activation function of the hidden layer and the output layer, we obtain the artificial neural network (perceptron) that we will apply to classify in the category of renal insufficiency in which a patient is.

The results obtained are good. In the network training phase, the ranking percentage for the three categories is: NOT (96.9), MODERATE (61.5) and ADVANCED (91.2); and for the reserve phase (the one used for the classification): NOT (96.4), MODERATE (33.3) and ADVANCED (100).

The study shows that, in the reserve phase, the classification percentages are high for the NOT and ADVANCED categories.

2. Material and Methods

The data from 184 patients were collected at the University Hospital of Salamanca (Spain) taking as variables: Age (years), Sex, Hematocrit (%), Blood creatine (mg/dl), Urine creatine (mg/dl), Albumin (g/dl), Body Mass Index (kg/m ^ 2) and Renal insufficiency.

2.1. Artificial Neural Network (ANN)

The artificial neural network that we will use in this study is the multilayer perceptron. In our case, this network is composed of an input layer, a hidden layer and an output layer. In Figure 1, we can see a typical perceptron formed by an input layer, a hidden layer and an output layer (N-H-M). Having studied the most effective initial structure of the network, in our case expressed in variables is (9-7-3).

The inputs in the network are the variables $x_1, x_2, x_3,..., x_N$ (independent variables), and the weights w_{ji} (importance of the connections between the neurons of the input layer and the hidden layer), the weights w_{kj} (importance of the connections between the neurons of the hidden layer and the output layer) and the output variables $y_1, y_2, y_3,..., y_M$ (dependent variables) form the neural network (perceptron). In our case, we will only have one qualitative dependent variable with three categories [16-22].

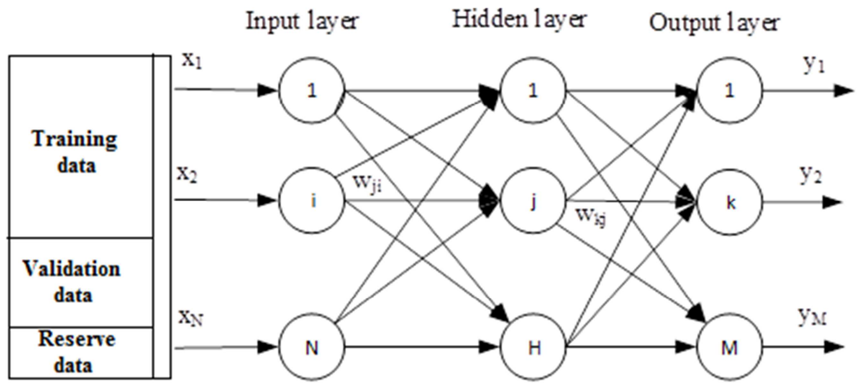

Figure 1. *Multilayer perceptron (N-H-M).*

For the analysis of the data we will create a partition variable: training, test and reserve samples. The training sample comprises the records of the data used to train the neural network; a certain percentage of cases in the dataset must be assigned to the above-mentioned sample in order to obtain a model. The validation data is an independent set of data records used to track errors during training, to avoid excess. It is highly recommended to create a training sample. Network training will generally be more effective if the test sample is smaller than the training sample. The reserve sample is another independent set of data records used to predict or classify the final neural network; the reserve sample error provides an estimate of the predictive capacity of the model, because reserved cases are not used to create such a model. In our case, the respective percentages for the three phases are 60%, 0% and 40%.

In this type of architecture, the connections between neurons are always forward, that is, they go from the neurons of a certain layer to those of the next; there are neither lateral connections between neurons belonging to the same layer nor backward connections, ranging from one layer to the previous one. Therefore, information is always transmitted from the input layer to the output layer.

The method we have followed to apply neural networks to the study of the classification of renal insufficiency is to establish the random seed to be able to replicate the studies and a partition variable to assign the training, test and reserve groups. In our case, due to the number of cases, we will use the training and reserve phases. To achieve the structure of the neural network (non-parametric technique) more efficient in the classification of the data, the activation functions of the input-hidden-output layers will be modified, trying to find the best relationship between layers (activation functions), so that it provides the best result in the output of the network. [23-29]. In our case: Hidden Layer = Sigmoid and Output Layer = Identity.

2.2. Multinomial Logistic Regression (MLR)

Multinomial Logistic Regression (MLR) is useful for situations in which you want to be able to classify subjects based on values of a set of predictor variables. This type of regression is similar to logistic regression, but it is more general because the dependent variable is not restricted to two categories. In our case the dependent variable (Renal insufficiency) has three categories: NOT, MODERATE and ADVANCED.

3. Results

The descriptive-exploratory analysis of data by means of the variables provides us with a series of values that will help us understand some of the results obtained later.

Table 1. *Descriptive statistics.*

	Age (years)	Ht (%)	Blood creatine (mg/dl)	Urine creatine (mg/dl)	Urea (mg/dl)	Albumin (g/dl)	Body Mass Index (kg/m ^ 2)
Mean	64,86	41,198	1,4339	75,5333	63,50	4,3750	30,3214
Std. Deviation	13,667	6,022	1,3956	50,221	50,781	0,5836	5,1155
Coefficient of variation	21,07%	14,61%	97,32%	66,49%	79,97%	13,35%	16,85%
Skewness	-0,583	-0,373	3,862	2,757	3,229	-1,947	0,424
Kurtosis	0,320	0,031	18,905	10,446	15,491	5,182	0,346
Range	75	30,90	10,77	322,70	406	3,60	27,6021

In this table we can see the following: the high value of the Coefficient of variation of the variables Blood creatine, Urine creatine and Urea, all of them well above 66%; its high degree of concentration (leptokurtic Kurtosis) around the mean, all of them exceed the value of 10, and its high skewness, positive or to the right, above 2.70. These three variables will be subject to variable sensitivity analysis.

The variable "Renal insufficiency" presents the following values for the three categories compared to the variable Sex:

Table 2. *Renal insufficiency.*

Renal insufficiency	N	Women	Men	Mean (years)
NOT	125	53	72	61,28
MODERATE	16	10	6	69,06
ADVANCED	43	18	25	73,72

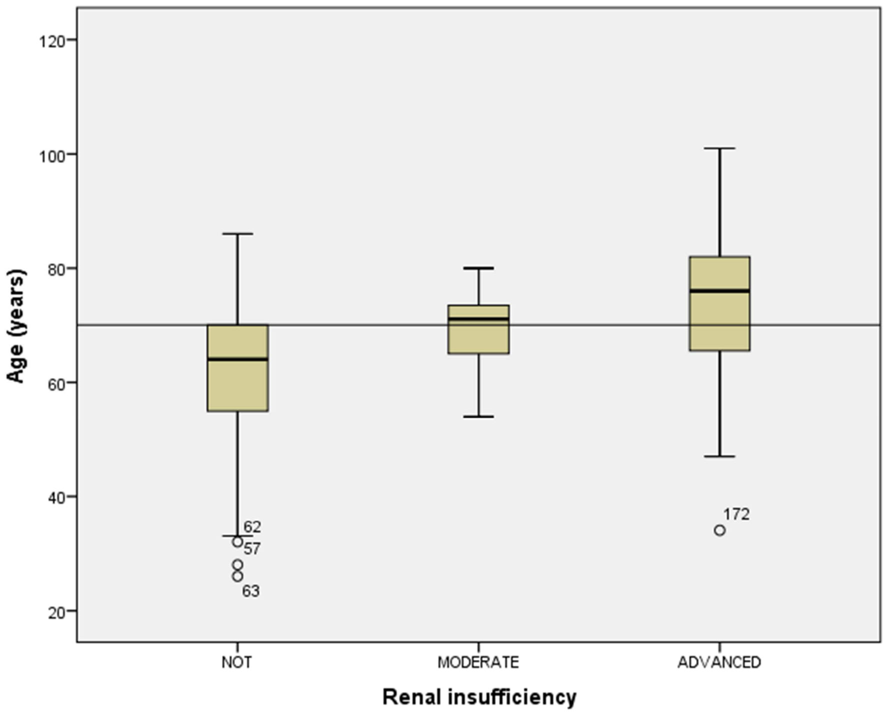

Figure 2. *Box-Plot chart.*

The box diagram (Box-Plot) collects, in a visual way, the distribution of the variable "Age" according to the category. For the category that does not present renal insufficiency (NOT), it has the lowest average age (61.28 years), although the subjects 62, 57 and 63 show anomalous values; for the category with moderate renal insufficiency (MODERATE), they have a mean age of 69.06 years, without anomalous values, and, for the category with advanced renal insufficiency (ADVANCED), their average age is 73.72 years, presenting an anomalous value in the subject 172. The "Interquartile Range" for the "NOT" and "ADVANCED" categories is the same (17) and for the "MODERATE" category it is the lowest value (9). Both categories have a negative or left-wing asymmetry. This study compares the results of artificial neural network (ANN) classification (initial multilayer perceptron: 9-7-3, excluding the bias in the hidden layer) with multinomial logistic regression (MLR).

Because not much data is available (184), a partition variable has been created for the artificial neural network in which the data is distributed in 60% for network training and 40% for the backup phase (network application). The comparison of the ANNs is carried out once performed the training, with the data of the reserve (Test data). The architecture of the artificial neural network (perceptron) used in the study is Hidden Layer = Sigmoid and Output Layer = Identity.

The sensitivity analysis for both the artificial neural network and the multinomial logistic regression model has been performed by suppressing, one by one, the independent variables (covariates for the artificial neural network). Once suppressed, they no longer enter the model, so that we can see the joint effect of the suppressed variables. The variables that come out of the model, considered here to be the most influential ones according to Table 1, are Blood creatine (*), Urine creatine (**) and Urea (***). The results are shown in Table 2.

Table 3. *ANN classification (Hidden Layer = Sigmoid and Output Layer = Identity) vs MLR. Dependent Variable: Renal insufficiency.*

Observed	Predicted							
	Percent Correct ANNs	**Percent Correct MLR**	**Percent Correct ANNs (*)**	**Percent Correct MLR (*)**	**Percent Correct ANNs (**)**	**Percent Correct MLR (**)**	**Percent Correct ANNs (***)**	**Percent Correct MLR (***)**
NOT	96,3	98,4%	96,3	96,0%	98,1	97,6%	94,4	94,4%
MODERATE	25,0	75,0%	25,0	56,3%	75,0	50,0%	25,0	31,3%
ADVANCED	88,9	93,0%	94,4	93,0%	83,3	88,4%	61,1	60,5%
Overall Percent	90,8	95,1%	92,1	91,8%	93,4	91,3%	82,9	81,0%

As we can see in Table 3, the joint classification percentages, in the analysis of sensitivity by means of variables, are favourable to the artificial neural network model (perceptron).

4. Conclusion

Of the results obtained in the classification of a patient's renal insufficiency by means of the two techniques: the non-parametric technique (ANNs) and the parametric technique (MLR), we can conclude that the result is favorable to the technique (ANNs) in the analysis of sensitivity by variable. That is, when a patient's data is entered either of the two models, regardless of any of the three variables analyzed in the model, the best result is obtained with the technique (ANNs).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Tu, J. V. (1996). Advantages and Disadvantages of Using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes. *Journal of Clinical Epidemiology*, 49 (11), 1225-1231.

[2] Lang, E. W., Pitts, L. H., Damron, S. L. & Rutledge, R. (1997). Outcome after Severe Head Injury: An Analysis of Prediction Based upon Comparison of Neural Network versus Logistic Regression Analysis. *Neurol Res,* 19, 274-80.

[3] Terrin, N., Schmid, C. H., Griffith, J. L., D'Agostino, R. B. & Selker, H. P. (2003). External Validity of Predictive Models: A Comparison of Logistic Regression, Classification Trees, and Neural Networks. *Journal of Clinical Epidemiology*. 56, 721-729.

[4] Song, J. H, Venkatesh, S. S., Conant, E. A., Arger, P. H. & Sehgal, C. M. (2005). Comparative Analysis of Logistic Regression and Artificial Neural Network for Computer-aided Diagnosis of Breast Masses. *Acad Radiol,* 12 (4), 487-95.

[5] Chen, S. T., Hsiao, Y. H., Huang, Y. L., Kuo, S. J., Tseng, H. S., Wu, H. K. et al. (2009). Comparative Analysis of Logistic Regression, Support Vector Machine and Artificial Neural Network for the Differential Diagnosis of Benign and Malignant Solid Breast Tumors by the Use of Three-dimensional Power Doppler Imaging. *Korean Journal of Radiology*, 10 (5), 464-471.

[6] Shafiei, E. et al. (2017). Comparison of Artificial Neural Network and Logistic Regression Models for Prediction of Psychological Symptom Six Months after Mild Traumatic Brain Injury. *Iran J Psychiatry Behav Sci*.

[7] Martín Pérez, E., Caldero Alonso, A. & Martín Martín, Q. (2021). Classification of Psychiatric Disorders Using Multinomial Logistic Regression vs Artificial Neural Network. *JP Journal of Biostatistics,* 18, (3), 395-408.

[8] Ayer, T., Chatwal, J., Alagoz, O., Kahn, C. E., Jr., Woods, R. W. & Burnside, E. S. (2010). Informatics in Radiology: Comparison of Logistic Regression and Artificial Neural Network Models in Breast Cancer Risk Estimation. *Radiographics*, 30 (1), 13-22.

[9] Bartfay, E., Mackillop, W. J. & Pater, J. L. (2006). Comparing the Predictive Value of Neural Network Models to Logistic Regression Models on the Risk of Death for Small-cell Lung Cancer Patients. *European Journal of Cancer Care* (England), 15 (2), 115-124.

[10] Caocci, G., Baccoli, R., Vacca, A., Mastronuzzi, A., Bertaina, A., Piras, E. et al. (2010). Comparison between an Artificial Neural Network and Logistic Regression in Predicting Acute Graft-vs-Host Disease after Unrelated Donor Hematopoietic Stem Cell Transplantation in Thalassemia Patients. *Experimental Hematology*, 38 (5), 426-433.

[11] Chen, H., Zhang, J., Xu, Y., Chen, B. & Zhang, K. (2012). Performance Comparison of Artificial Neural Network and Logistic Regression Model for Differentiating Lung Nodules on CT Scans. *Expert Systems with Applications*, 39, 11503-11509.

[12] Dreiseitl, S. & Ohno-Machado, L. (2002). Logistic Regression and Artificial Neural Network Classification Models: A Methodology Review. *Journal of Biomedical Informatics*, 35 (5-6), 352-359.

[13] Jaimes, F., Farbiarz, J., Álvarez, D. & Martínez, C. (2005). Comparison between Logistic Regression and Neural Networks to Predict Death in Patients with Suspected Sepsis in the Emergency Room. *Critical Care*, 9 (2), R150-156.

[14] Kazemnejad, A., Batvandi, Z. & Faradmal, J. (2010). Comparison of Artificial Neural Network and Binary Logistic Regression for Determination of Impaired Glucose Tolerance/Diabetes. *Eastern Mediterranean Health Journal*, 16 (6), 615-620.

[15] Lin, C. C., Ou, Y. K., Chen, S. H., Liu, Y. C. & Lin, J. (2010). Comparison of Artificial Neural Network and Logistic Regression Models for Predicting Mortality in Elderly Patients with Hip Fracture. *Injury*, 41 (8), 869-873.

[16] Lin, S. P., Lee, C. H., Lu, Y. S. & Hsu, L. N. (2006). A Comparison of MICU Survival Prediction Using the Logistic Regression Model and Artificial Neural Network Model. *Journal of Nursing Research*, 14 (4), 306-314.

[17] Ottenbacher et al. (2004). Comparison of Logistic Regression and Neural Network Analysis Applied to Predicting Living Setting after Hip Fracture. *Annals of Epidemiology*, 14 (8), 551-559.

[18] Pai, D. R. et al. (2012). Experimental Comparison of Parametric, Non-parametric, and Hybrid Multigroup Classification. *Expert Systems with Applications*, 39 (10), 8593-8603.

[19] Sargent, D. J. (2001). Comparison of artificial neural networks with other statistical approaches: Results from medical data sets. *Cancer*, 91 (sup 8), 1636–1642.

[20] Masic, N. & Pfurtscheller, G. (1993). Neural Network Based Classification of Single-trial EEG Data. *Artificial Intelligence in Medicine*, 5, 1993, 503-513.

[21] Dorffner, G. & Porenta, G. (1994). On Using Feedforward Neural Networks for Clinical Diagnostic Tasks. *Artificial Intelligence in Medicine*, 6, 417-435.

[22] Dybowski, R. & Gant, V. (2001). Clinical Applications of Artificial Neural Networks. *Cambridge University Press*.

[23] Mazurowski, M. A. et al. (2008). Training Neural Network Classifiers for Medical Decision Making: The Effects of Imbalanced Datasets on Classification Performance. *International Joint Conference on Neural Networks IJCNN'07*. Orlando, Florida, USA, 21 (3), 427-436.

[24] Ince, T. et al. (2010). Evaluation of Global and Local Training Techniques over Feed-forward Neural Network Architecture Spaces for Computer-aided Medical Diagnosis. *Expert Systems with Applications*, 37, 8450-8461.

[25] Amato F. et al. (2013). Artificial Neural Networks in Medical Diagnosis. *Journal of Applied Biomedicine*, 11, 47-58.

[26] Peek, N. et al. (2015). Thirty Years of Artificial Intelligence in Medicine (AIME) Conferences: A Review of Research Themes. *Artificial Intelligence in Medicine*, 65, 61-73.

[27] Shaikhina, T. & Khovanova, N. A. (2017). Handling Limited Datasets with Neural Networks in Medical Applications: A Small-data Approach. *Artificial Intelligence in Medicine*, 75, 51-63.

[28] Corsinia, M.-M., Schmittb, A. & Bruzekb, J. (2005). Process Variability on the Human Skeleton: Artificial Network as an Appropriate Tool for Age at Death Assessment. *Forensic Science International*, 148, 163-167.

[29] Barni, M., Faila, P. & Lazzareti, R. Privacy-preserving ECG (2011). Classification with Branching Programs and Neural Networks. *IEEE Transactions on Information Forensics and Security*, 6 (2), 452-468.