

An artificial neural network to classify healthy aging in elderly Brazilians

Abstract

Aging in Brazil, especially in the Amazon, is a complex and irregular process. Something is happening here that cannot be explained simply due to social inequalities. The objective of this study was to present the development of an artificial neural network and the stages of training, validation and testing for the classification of healthy aging among elderly Brazilians. We constructed a protocol for rapid diagnosis and health screening for the elderly. The form was developed offline in Microsoft Excel. Macros (routines capable of performing pre-programmed tasks) were created using Microsoft's Visual Basic for Applications (VBA) language. In the analysis of the confusion matrix, good accuracy were obtained in all stages, training (61.5%), validation (60.0%) and test (80.0%), which indicates that the network learned through the inputs and outputs initially defined and during the sample divisions performed for testing and validation. In the test stage, a ROC curve was obtained with better true positive rates and lower false positive rates, being close to the Y axis (left side), thus indicating better results. We conducted a pilot study with thirty-six community active elderlies from a city in Eastern Amazonia, Brazil. This study was divided into four parts: data collection, data pre-processing, training of an artificial neural network and evaluation methods.

Keywords: aged, aging, elderly, artificial neural network, MATLAB.

Volume 9 Issue 5 - 2020

Ágatha Yasmin de Sousa Araujo,¹ Maylon Sivalcley da Costa Rocha,¹ Elton Rafael Alves,¹ Ana Cristina Viana Campos²

¹Institute of Geosciences and Engineering, Universidade Federal do Sul e Sudeste do Pará, Brazil

²Institute of Studies in Health and Biological, Universidade Federal do Sul e Sudeste do Pará, Brazil

Correspondence: Ana Cristina Viana Campos, Universidade Federal do Sul e Sudeste do Pará (Unifesspa). Folha 31, Quadra 07, Lote Especial, Marabá, Pará, Brazil, ZIP code 68507-590. Tel 55 94 2101-7116. Email campos.acv@gmail.com, anacampos@unifesspa.edu.br

Received: September 30, 2020 | **Published:** October 16, 2020

Introduction

The world's elderly population has been increasing over the years. In developing countries, there is a worrying relationship between the accelerated aging of the population and greater dependence on a family and social structure unable to provide necessary support to the elderly.¹

Scientific evidence on longevity and good health is still scarce. WHO published the world's first report on health. In this document, healthy aging is constructed on a combination of the individual's functional ability, relevant environmental and contextual characteristics and interactions between the individual and those characteristics.²

Healthy aging has been evaluated on domains and scales. A review study found studies that measure healthy aging by assessing activities of daily living (37 studies), cognitive functions (33 studies), by assessing psychological well-being (24 studies), questions about participation in social activities (22 studies), general health status of individuals (16 studies). Fifteen studies applied Health Survey (SF12/36) or developed health scores to assess healthy aging.³

Different indicators and measures have been proposed to assess healthy aging. Most of them developed with and for the elderly population of developed countries.⁴ In Brazil, the age range that defines the elderly population is extensive, which results in a very heterogeneous population. This process is the result of high fertility between the 1950s and 1960s and the reduction in mortality of the elderly population in recent decades.⁵

A Brazilian study indicates that the presence of diseases is common among the elderly, closely influenced by socioeconomic factors and little related to lifestyle.⁶ In another study, racial inequalities were associated with general living conditions among the elderly, and most of the black and brown elderly lived alone, had lower educational levels and lower income.⁷

However, aging in Brazil, especially in the Amazon region, is a complex and irregular process. Something is happening here that cannot be explained simply due to social inequalities. The objective of this study was to present the development of an artificial neural network and the stages of training, validation and testing for the classification of healthy aging among elderly Brazilians.

Material and methods

We conducted a pilot study with thirty-six elderlies living in the community from a city in Eastern Amazonia, Brazil.

All the elderly assisted by a voluntary association were interviewed by trained and calibrated researchers during a conference. The inclusion criteria were: being 60 years old and more; be aware, oriented and able to interact during the interview and be able to move around with or without help. The interviews were conducted in a private, individual room between 10 and 15 minutes.

This study was divided into four parts: data collection, data pre-processing, training of an artificial neural network and evaluation methods.

A) Data collection

We constructed a protocol for rapid diagnosis and health screening for the elderly. The form was developed offline in Microsoft Excel. Macros (routines capable of performing pre-programmed tasks) were created using Microsoft's Visual Basic for Applications (VBA) language.

The variables investigated were: age; sex (male, female); marital status (married, separated, single, widowed); education level (illiterate, basic education, higher education); home arrangement (living alone, lives with relatives/parents, lives with partner); quality of life (very bad, bad, regular, good and very good); number of functional activities of daily living performed with assistance; cognitive deficit (yes, no);

mobility difficulties (yes, no); acute or chronic diseases (yes, no); number of falls in last year, use of medications (0, 1, 2, 3 or more); alcohol consumption (none, less than 1 day a week, 1 day a week, 2 to 3 days a week, 4 to 6 days a week and every day), smoking habit (never smoked, smoked and stopped, smokes less than 1 cigarette a day, smokes 1 or more cigarettes a day); community participation (yes, no) and family support (very bad, bad, regular, good and very good).

Of the 19 (nineteen) questions, only 14 (fourteen) scored. The scores used were used as input for training an artificial neural network, forming a matrix of 14 columns and 36 rows.

Each person received an overall score for the sum of the questions answered. Based on this sum, the median of the scores was calculated. If the score was above the median, healthy aging was classified as good (1); if it was below, it was rated poor (2). The set of 1 and 2 was used as the desired output for in artificial neural network (ANN).

B) Data pre-processing

The input data for ANN training has been normalized to reduce discrepancies between input data. Equation 1 describes the method of normalization used.

$$value_{normalized} = \frac{value_{original} - \min_{input}}{\max_{input} - \min_{input}} \quad (1)$$

This normalization transformed original values of input variables into values of the interval [0, 1].

C) Training of an artificial neural network

An Artificial Neural Network (ANN) tries to imitate the functioning of the human brain, allowing to create computational systems capable of learning, making generalizations and adaptations.⁸ The ANN used in this work consisted of a multilayer perceptron, trained through the backpropagation algorithm.

Cross validation was used as an ANN training stop criterion. So, the data used in this proposal were randomly divided into three groups: 70% were used for training, 15% were used for validation and 15% were used for testing da ANN. Therefore, results of test data were used as overall performance for the model.

The ANN training topology adopted was: fourteen input patterns, two output patterns, sigmoid activation function in the hidden layer and softmax activation function in the hidden layer. The best accuracy rates of test were obtained through forty-five neurons in the hidden layer.

The RNA used in this study was used to classify two categories: “1” the elderly have a good healthy aging and “2” the elderly have poor healthy aging. The program used for training and simulation of the artificial neural network was Matlab®.

D) Evaluation methods

One way to assess ANN was the employment of the rate for adequate classification obtained through a confusion matrix between the predicted class (output class) and the true class (real class)⁸. In the matrix, each column represents a predicted result, while each line represents a true result. The confusion matrix was used to assess models of classification used in this study is exposed in Table 1.

When a positive example was classified as positive by the classifier, it was computed into the matrix as true positive, however if it was classified as negative, then was denominated false negative. When a negative example was classified as negative by the classifier then it was computed as true negative, however if it was classified as positive, then it was classified as false positive.

Table 1 Confusion matrix for the evaluation of classification model

		Predict Class	
		1	2
True Class	1	TP	FP
	2	FN	TN
		P	N

From these crossings we obtained: TP (True Positives), FP (False Positives), FN (False Negatives) and TN (True Negatives). Finally, N represents the number of total negative events and P the number of total positive events, which are analyzed by the classifier.

From Table 1 possible metric analyses of the classifier performance are also exposed: accuracy (A), sensibility (S), specificity (E), efficiency (F), positive predicted value (PPV) and negative predicted value (NPV).

$$A = \frac{TP + TN}{P + N} \times 100 \quad (2)$$

$$S = \frac{TP}{TP + FN} \times 100 \quad (3)$$

$$E = \frac{TN}{TN + FP} \times 100 \quad (4)$$

$$F = \frac{S + E}{2} \times 100 \quad (5)$$

$$PPV = \frac{TP}{TP + FP} \times 100 \quad (6)$$

$$NPV = \frac{TN}{TN + FN} \times 100 \quad (7)$$

The sensitivity corresponds to the hit rate in the positive class, specificity corresponds to the hit rate in the negative class, FP (False Positives) and FN (False Negatives) correspond, respectively, number of examples whose true class is negative but which were incorrectly classified as positive and number of examples whose true class is positive but which were incorrectly classified as negative.

An alternative way of evaluating the performance of a classifier is through Receiver Operating Characteristic Curves (ROC).⁹ The ROC graphic is a two-dimensional graphic, with X and Y axes representing the measures of false positive rate and true positive rate. An ideal classifier is represented by the point (0, 1) in the ROC.

This research approved by a Brazilian Ethical Committee (number 05573218.5.0000.0018). After accepting to participate in the study,

the elderly signed or stamped with digital of the right thumb, the free and informed consent form.

Results & Discussion

Six men (16.7%) and 30 women (83.3%) participated in this phase of the study. The mean age was 72.11 (\pm 9.38) and the median was 73 years old. The profile of the elderly population in Brazil and in the world indicates greater longevity of women, a phenomenon called feminization.¹⁰⁻¹²

After the models were built, we used the test data set to evaluate the accuracy and prediction performance of the models. To verify the accuracy of the responses of the developed neural network, confusion

matrices were plotted for the three stages: training, validation and testing (Figure 1). In the analysis of the confusion matrix, good accuracy was obtained in all stages, training (61.5%), validation (60.0%) and test (80.0%), which indicates that the network learned through the inputs and outputs initially defined and during the sample divisions performed for testing and validation. Parameters calculated for training, validation and testing step are in Table 2.

Our results are similar to other studies,^{13,14} indicating that the ANN model is accurate in classifying dependent variables. An artificial neural network can be defined as computational techniques that present a mathematical model inspired by the intelligent neural structure that acquire knowledge through experience.¹⁵

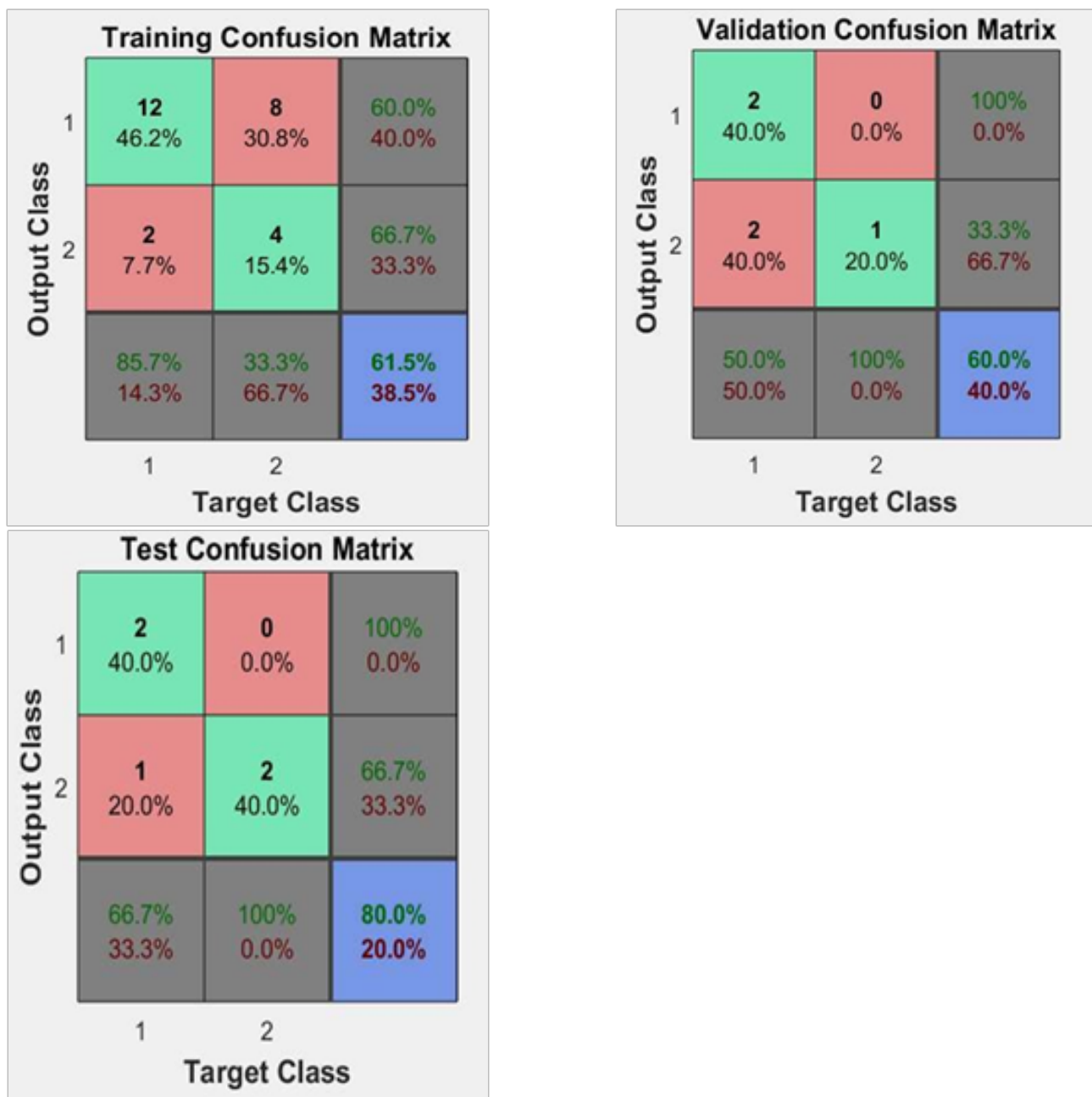


Figure 1 Network confusion matrix for training, validation and test.

Table 2 Results of the parameters of the confusion matrix graph for each phase

Parameters	Training	Validation	Test
Sensitivity	85.7%	50.0%	66.7%
Specificity	33.3%	100.0%	100.0%
Efficiency	59.5%	75.0%	83.4%
Accuracy	61.5%	60.0%	80.0%
Positive Predictive Value	60.0%	100.0%	100.0%
Negative Predictive Value	66.7%	33.3%	66.7%

The overall performance of the network in the test phase deserves to be highlighted, since with only 15% of the data, not previously used, 80% of correctness was obtained for the classification of healthy aging in this sample of elderly Brazilian.

Figure 2 shows the ROC curve (class 1 – blue color and class 2 – brown color) for the sample in each stage (training, validation, and testing). The diagonal line (gray line) of the ROC represents random classifiers, classifiers that are below this line are considered worse than the random. In the training ROC and validation ROC, both classes 1 and 2 showed satisfactory results, because their curves are better than the random classifier (above the gray line). In the test stage, a ROC curve was obtained with better true positive rates and lower false positive rates, being close to the Y axis – left side and point (0,1) – thus indicating better results.

The incorporation of ANN learning in predicting outcome using self-reported predictors may be potentially useful for predicting health aging in elderly Brazilians, with emphasis on quality of life. Increase in the number of elderly people in the world has opened space for the study and evaluation of the quality of life in aging¹⁶. Studies have shown that the quality of life of the elderly is influenced by functional and cognitive ability, social conditions and the presence of chronic diseases and lifestyle.^{17,18}

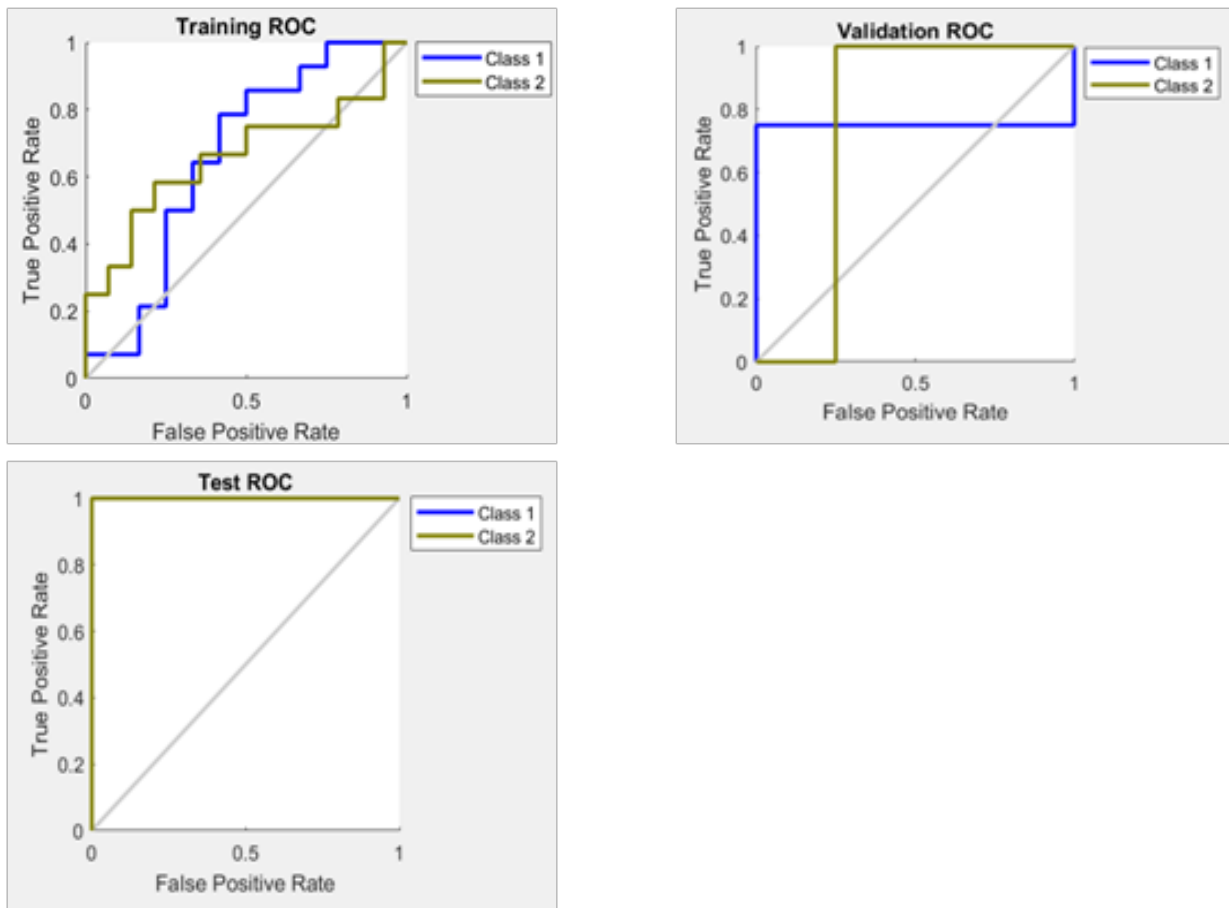


Figure 2 ROC for training, validation and test.

There is sufficient scientific evidence proving the use of ANN in the health status, but few have been applied specifically to the elderly.^{19,20} ANN can be an analysis model for investigating healthy

aging, using simple and inexpensive questions that assess the elderly’s perception about their life and health.

Conclusion

We conclude that the use of artificial neural networks to classify healthy aging in elderly Brazilians was efficient. Our results may contribute to a deeper understanding of the most accurate diagnosis of healthy aging.

Epidemiological studies have a prominent role in studies on aging and the use of technologies can bring answers that go unnoticed in conventional analyzes. Therefore, the application of artificial neural networks will be of great relevance to obtain a more effective classification of data.

Acknowledgments

Pro-reitoria de Pós-Graduação, Pesquisa e Inovação Tecnológica (Propit) da Universidade Federal do Sul e Sudeste do Pará (Unifesspa).

Conflicts of interest

The author declares have no conflict of interest about the publication of this paper.

References

1. Lunenfeld B, Stratton P. The clinical consequences of an ageing world and preventive strategies. *Best Pract Res Clin Obstet Gynaecol.* 2013; 27(5):643–659.
2. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. *Lancet.* 2016;387(10033):2145–2154.
3. Lu W, Pikhart H, Sacker A. Domains and measurements of healthy aging in epidemiological studies: A Review. *Gerontologist.* 2019;59(4):e294–e310.
4. Michel JP, Graf C, Ecarnot F. Individual healthy aging indices, measurements and scores. *Aging Clin Exp Res.* 2019;31(12):1719–1725.
5. Camarano AA, Kanso S, Mello JL. *Como vive o idoso brasileiro?* In: Camarano AA. Os Novos Idosos Brasileiros: Muito Além dos 60? Rio de Janeiro: IPEA, 2004.
6. Melo LA de, Lima KC de. Prevalence and factors associated with multimorbidities in Brazilian older adults. *Ciênc saúde coletiva.* 2020;25(10):3869–3877.
7. Oliveira BLCA de, Silva AM da, Silva RA da, et al. Racial inequalities in the socioeconomic, demographic and health conditions of elderly from Maranhão State, Legal Amazon, Brazil: a population-based study. *Acta Amaz.* 2014;44(3):335–344.
8. Fawcett T. An introduction to ROC analysis. *Pattern recognition letters.* 2006;27(8):861–874.
9. Haykin S. *Neural networks: a comprehensive foundation.* Prentice–Hall, Inc., 2007.
10. Carney GM. Toward a gender politics of aging. *J Women Aging.* 2018;30(3):242–258.
11. Campos ACV. Predictors of healthy aging for Brazilian elderly. *MOJ Gerontol Ger.* 2020;5(5):142–144.
12. Campos ACV, et al. Family functioning of Brazilian elderly people living in community. *Acta paul enferm.,* 2017;30(4):358–367.
13. Renganathan V. Overview of artificial neural network models in the biomedical domain. *Bratisl Lek Listy.* 2019;120(7):536–540.
14. Fei Y, Li WQ. Improve artificial neural network for medical analysis, diagnosis and prediction. *J Crit Care.* 2017;40:293.
15. Hair JF, Black WC, Babin JB, et al. *Multivariate Data Analysis.* Upper Saddle River: Prentice–Hall; 2009.
16. Li CP. Quality of life patterns and survival among older people. *J Nurs Res.* 2013;21(2):94–109.
17. Bonfiglio V, Umegaki H, Kuzuya M. Quality of life in cognitively impaired older adults. *Geriatr Gerontol Int.* 2019;19(10):999–1005.
18. Abolhassani N, Santos–Eggimann B, Bula C, et al. Temporal changes in importance of quality of life domains: a longitudinal study in community-dwelling Swiss older people. *Qual Life Res.* 2019;28(2):421–428.
19. Świetlik D, Białowąs J. Application of Artificial Neural Networks to Identify Alzheimer’s Disease Using Cerebral Perfusion SPECT Data. *Int J Environ Res Public Health.* 2019;16(7):1303.
20. Lugade V, Lin V, Farley A, et al. An artificial neural network estimation of gait balance control in the elderly using clinical evaluations. *PLoS One.* 2014;9(5):e97595.