ARTIFICIAL NEURAL NETWORKS AS AN ALTERNATIVE AND COMPLEMENTARY TOOL TO HELP IN AUTISM DIAGNOSIS

Data de submissão: 31/10/2023 *Data de aceite: 01/12/2023*

Gabriela Menossi da Silva Floriano
Institute of Chemistry - São Paulo State University - Unesp
Araraquara -SP
http://lattes.cnpq.br/5125792709572256

Érica Regina Filletti
Department of Engineering, Physics and Mathematics
Institute of Chemistry - São Paulo State University - Unesp
Araraquara -SP
http://lattes.cnpq.br/9684427777093476

ABSTRACT: Artificial Neural Networks (ANNs) are computational models based on human brain structures. The ANNs receive data, learn from it and generate coherent responses to unknown data. Autism Spectrum Disorder (ASD), commonly known only as autism, has already been the subject of studies in the area of artificial intelligence to aid in the diagnosis, which, in turn, is very complicated because the symptoms appear from the age of two and are generally very light. This disorder is very complex and therefore it is not always possible to make an accurate diagnosis. Therefore, this work check the possibility of using of ANNs as an alternative and complementary tool to aid in the diagnosis of autism, due to their ability to learn from a reduced amount of data and generate coherent responses to new data. Tests were made using the MATLAB software with the nnstart and the pattern recognition application tool. For the first tests with answers from only 15 real people, 93.3% of correct answers were obtained with ANN. With the addition of fictitious responses totaling 150 responses, it was possible to reach 100% ANN accuracy.

KEYWORDS: Artificial Neural Networks. Learning algorithm. Computational models. Autism.

INTRODUCTION

Artificial Neural Networks are distributed parallel systems composed of simple processing units, called neurons, whose purpose is to calculate mathematical functions, normally non-linear (HAGAN and MENHAJ, 1994; HAYKIN, 1999). They are computational mathematical models inspired by biological networks capable of obtaining knowledge from a database, learning from it and then providing accurate results for other data. (BRAGA, CARVALHO and LUDERMIR, 2000). It is an alternative widely used today due to its power of learning and generalizing information with very satisfactory results.

Autism Spectrum Disorder is a complex cyclical condition that affects the nervous system and generates difficulties in using language and relating to other people, in addition to obsessive desires and repetitive behaviors. It is a psychiatric disorder that develops in childhood, affecting communication and the child's overall social development. (LEVY, 2000; RAPIN and GOLDMAN, 2008).

The diagnosis of autism is made through behavioral analyzes and becomes very subjective as it depends on the perception of a professional according to different possibilities, such as interviewing the child's parents or caregivers or even observing the child in various activities. Therefore, diagnosing the disorder is time-consuming and imprecise. Autism has no cure, but the sooner it is diagnosed, the greater the chances of having effective treatment to improve the patient's quality of life (ALMEIDA, 2018). With this in mind, research began to be carried out on alternatives for diagnosis to become accurate and fast and, thus, an alternative found was to use computational models that involve artificial intelligence.

Currently, when we talk about autism, we already refer to the Childhood Autism Rating Scale, known as the CARS scale, developed by Shopler, Reichler and Renner (1988). It contains fifteen questions and three possible diagnoses: severe autism, mild/moderate autism or non-autistic. The translation into Portuguese was made and validated by Pereira (2007). It is considered a very strong scale in relation to behaviors associated with autism and is found in several languages.

The CARS-BR-Fuzz mathematical model (ZAGO, 2019) proposes a system according to Fuzzy rules, or fuzzy logic (translation into Portuguese) that contains the same questions as the original CARS scale (SHOPLER, REICHLER and RENNER, 1988) and with the same possible diagnoses: severe autism, mild/moderate autism or non-autistic. This type of artificial intelligence is commonly used in cases involving uncertainties within bioscience.

Given the above, the objective of this work was to verify the possibility of using artificial neural networks as an alternative tool to assist in the diagnosis of Autism Spectrum Disorder, based on the CARS-BR scale (SHOPLER, REICHLER and RENNER, 1988), producing a computational mathematical model with rapid diagnosis because it can analyze large amounts of data presenting results according to the same classifications: severe autism, mild-moderate autism and non-autistic. The interesting thing about the study carried out in this work is that despite much research, no results were found from the application of neural networks with the CARS-BR scale and autism, showing the proposed innovation.

METHODOLOGY

In this section, we will explain how the matrices with the databases for ANN training were created, as well as their development.

Database

First, some response data from the CARS-BR scale from fifteen real people published in the work of Zago (2019) was used. This data was transported to a table in Excel 2013 software.

As the ANN developed with the database of fifteen people did not obtain 100% accuracy, more data was created with the responses of fictitious people according to the proposed scale to have a larger data set for the ANN to work with, totaling one hundred and fifty samples of data to be used in the development of the neural model.

The ANN input matrix was composed of the answers to the fifteen questions on the CARS-BR scale, which meant that the ANN had fifteen neurons in the input layer. Thus, for the ANN with data from fifteen real people, the input matrix had dimensions of 15x15; the input matrix for the ANN with real and fictitious responses was 15x150. The ANN output layer had 3 neurons, each responsible for identifying one of the three responses on the CARS-BR scale and, therefore, the output matrix had dimensions of 3x15 for the ANN with data from real people and 3x150 for the ANN with responses of real and fictional people.

Developed neural models

To implement the ANNs, MATLAB version 2019 software was used with the nnstart tool available, choosing the pattern recognition app option for pattern classification, as shown in Figure 1. The training algorithm used was a staggered conjugate gradient (MOLLER, 1993).

The division of all samples was done randomly using the dividerand command in which 70% were for ANN training, 15% for validation and 15% for testing. In the case of ANN for real and fictitious responses, as there were 150 data points in total (15 responses from real people and 135 fictitious responses), 104 responses were used in the training set, and the validation and test sets had 23 responses each one, as shown in Figure 2.

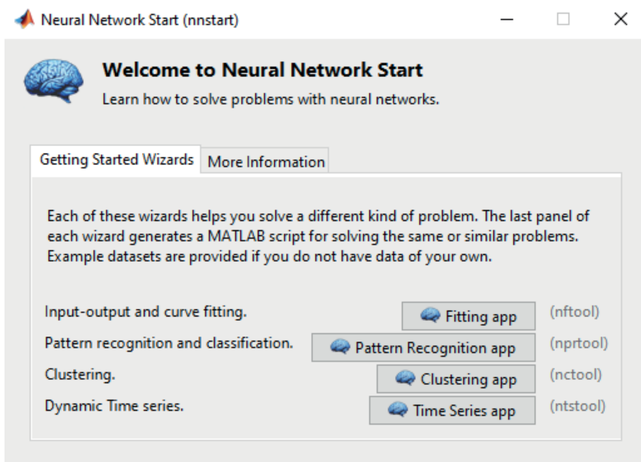

Figure 1: nnstart tool used for testing.

Source: Author herself, 2020.

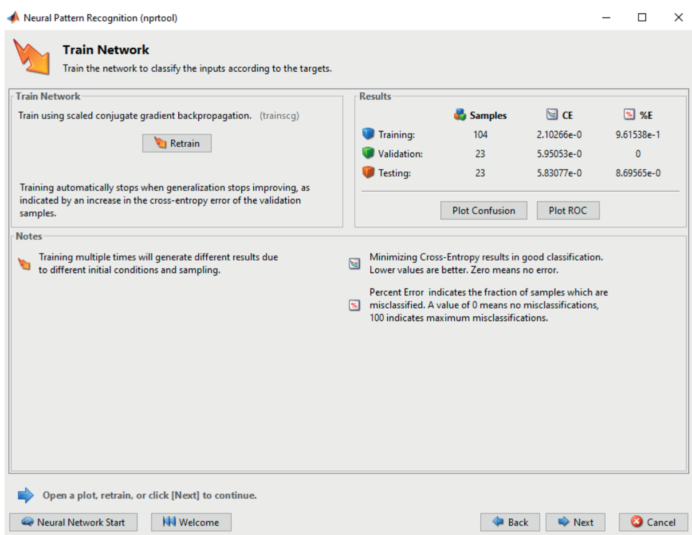

Figure 2**:** ANN training – data division.

Source: Author herself, 2020.

RESULTS AND DISCUSSIONS

Initially, tests were carried out only with data from fifteen real people and then with data from real and fictitious people, totaling 150 samples (15 real people and 135 fictitious people). To develop the two ANNs, tests were carried out by varying the number of neurons in the intermediate layer (between 6 and 10 neurons) to obtain the best result, through trial and error. The ANNs were designed with 15 neurons in the input layer (representing the 15

CARS-BR scale responses) and 3 at the output (each representing a diagnostic class: class 1 – no autism; class 2 – mild/moderate autism; and class 3 – severe autism).

Tests only with answers from real people

Figure 3 shows the confusion matrix with the best results obtained for the responses of real people in training, validation and testing, with 10 neurons in the intermediate layer and with the staggered conjugate gradient algorithm. On the main diagonal, in green, are the ANN´s correct results, and in the red part are the errors.

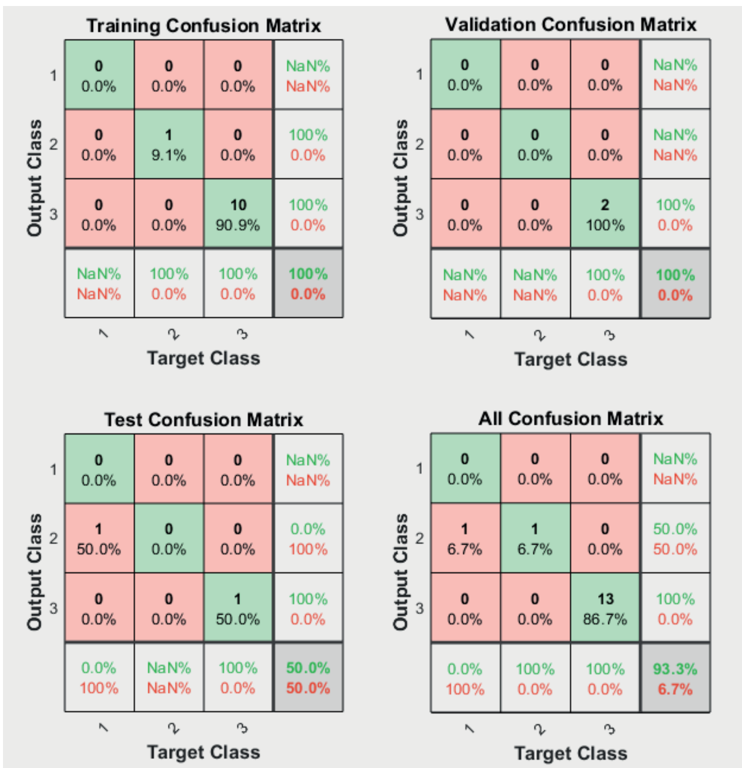

Figure 3: Confusion matrix – real people.

Source: Author herself, 2020.

ANN responses were classified into three distinct categories, according to the three possible diagnoses of the CARS-BR scale, namely: class 1 – without autism; class 2 – mild/moderate autism; and class 3 – severe autism. In the set presented to the ANN, of the eleven samples (70%) that were used for training, all were classified correctly; there were no class 1 samples in the training set. It is worth mentioning that of the fifteen responses from real people, there was only one classified without autism, one classified as mild/moderate autism and thirteen responses in the severe autism class. In validation, the two samples that made up the 15% were class 3 and both were classified correctly. However, in the ANN test, of the two samples (15%) that were class 1 and 3, the one from class 1 (without autism) was wrongly classified as class 2 (mild/moderate autism) and the other was classified correctly, configuring a hit 50% on the test set. This totals an average of 93.3% accuracy for the ANN as a whole. Figure 4 presents the ANN error histogram, which presents a Gaussian distribution centered on zero, as expected.

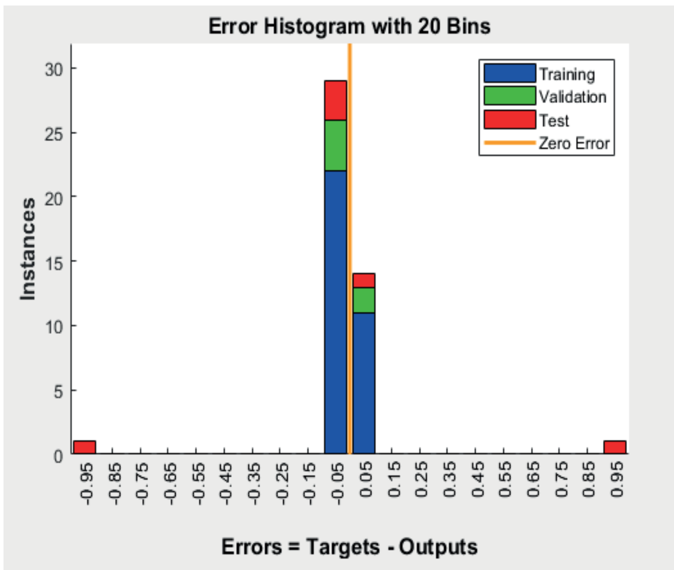

Figure 4: Error histogram – Real people.

Source: Author herself, 2020.

Finally, Table 1 shows the results obtained by ANN compared with the expected results of the set in training, validation and testing, in the case of using only responses from real people. Sample 13, highlighted in the table, was incorrectly classified as mild/moderate autism while its classification is without autism, as already discussed.

Sample	ANN´s results			ANN classification	Expected results			Expected classification
1	0	0	1	3	0	0	1	3
2	0	0	1	3	0	0	1	3
3	0	0	1	3	0	0	1	3
4	0	0	1	3	0	0	1	3
5	0	0	0,999	3	0	0	1	3
6	0	0	1	3	0	0	1	3
7	0	0	0,999	3	0	0	1	3
8	0	0	1	3	0	0	1	3
9	0	0	0,999	3	0	0	1	3
10	0	1	0	2	0	1	0	2
11	0	0	1	3	0	0	1	3
12	0	0	1	3	0	0	1	3
13	0	0,999	0	2	1	0	0	1
14	0	0	1	3	0	0	1	3
15	0	0	1	3	0	0	1	3

Table 1: Obtained and expected results from the database – Real people

Source: Autor herself, 2020.

Tests with real and fictitious people

In order to improve the ANN's correct percentage and have a more balanced database, 135 responses from fictitious people were created for the ANN training, which added to the 15 responses from real people, totaled 150 responses. For the test with real and fictitious people, the number of neurons in the intermediate layer was also varied between 6 and 10 until finding the expected result of 100% accuracy, which was obtained with 10 neurons in the intermediate layer, as shown in Figure 5.

In the database presented to RNA, of the 104 samples (70%) that went to training, 104 were correctly classified according to their classes. In validation, the 23 samples that made up the 15% were correctly classified as well as the 23 test samples. This totals an average of 100% accuracy for the ANN as a whole.

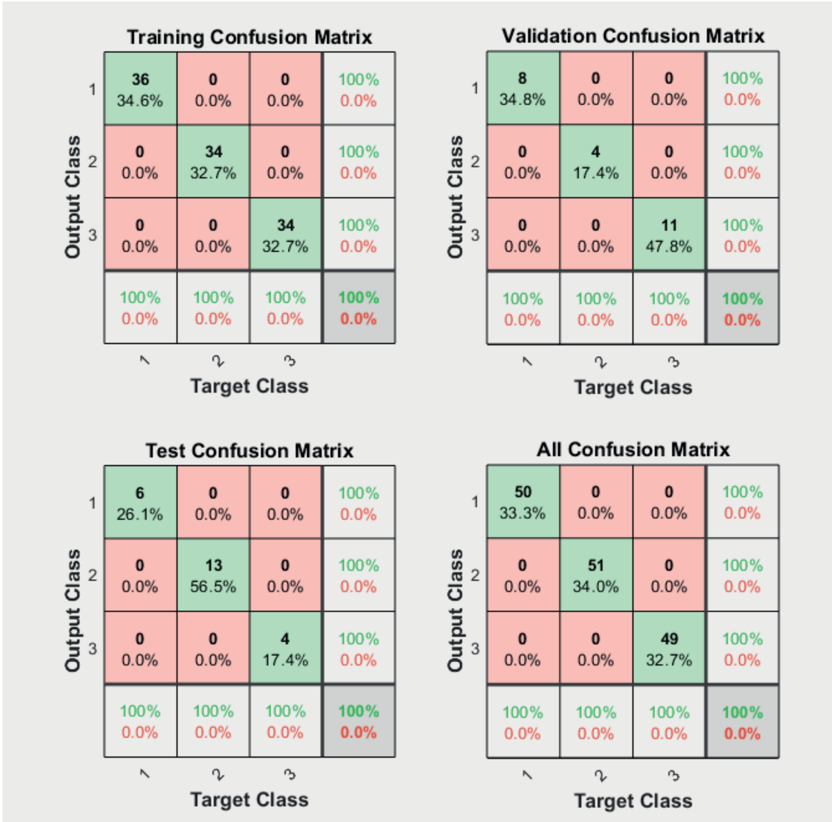

Figure 5: Confusion matrix – real and fictitious people.

Source: Author herself, 2020.

Again, when we analyze the error histogram, Figure 6, we observe that it resembles a Gaussian curve centered at zero.

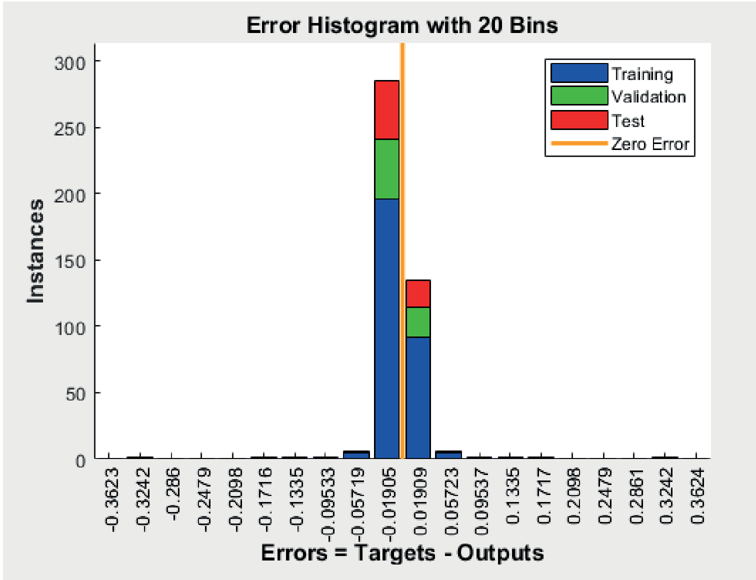

Figure 6: Error histogram – Total sample.

Source: Author herself, 2020.

Figure 7 shows the architecture of the ANN that obtained the results presented, with 10 neurons in the intermediate layer.

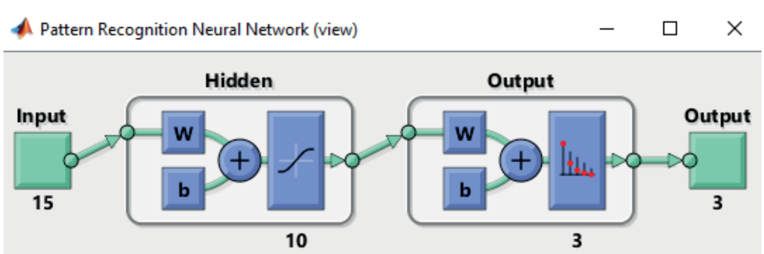

Figure 7: Classification ANN scheme.

Source: Author herself, 2020.

FINAL CONSIDERATIONS

The first tests carried out with only the responses of 15 real people were not enough to achieve 100% accuracy in the development of the ANN as a whole and that is why the data set was increased with responses from fictitious people totaling 150 samples. When this new set of data was provided to the ANN, the expected result was obtained without errors in training, validation and testing and a histogram of errors with the characteristic of a Gaussian curve.

Therefore, the addition of fictitious data helped to improve the performance and results of the ANN, reaching the ideal situation, with 100% accuracy for the 150 samples.

Therefore, it can be concluded that artificial neural networks are viable to assist in the diagnosis of autism as a complementary tool to make the diagnosis faster and more accurate. To do this, it is necessary to have a good database for training the ANN.

REFERENCES

ALMEIDA, M. S. R. **Instrumentos diagnósticos para avaliar o autismo – TEA.** [S.l.], 2018.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. (2000). **Redes Neurais Artificiais: Teoria e Aplicações**. LTC - Livros Técnicos e Científicos Editora S. A. Rio de Janeiro –RJ, 2000.

HAGAN, M. T.; MENHAJ, M. B. Training feedforward networks with the Marquardt algorithm. IEEE Transactions on Neural Networks, v. 5, n. 6, p. 989-993, 1994.

HAYKIN, S. (1999). **Neural Networks: a comprehensive foudation**. 2nd ed. New Jersey: Prentice Hall.

LEVY, P. de Q. **Protocolo para estudo das pessoas com autismo. Autismo e Genética**, Universidade de Lisboa, v. 31, n. 2, p. 149–154, 2000.

MOLLER, Martin Fodslette. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**, v. 6, p. 525–533, 1993.

RAPIN, I.; GOLDMAN, S. **A escala cars brasileira: uma ferramenta de triagem padronizada para o autismo.** Jornal da Pediatria, v. 84, n. 6, 2008.

PEREIRA, A. M. **Autismo Infantil: Tradução e Validação da Cars (Childhood Autism Rating Scale) para uso no Brasil**. Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, 2007.

SHOPLER, R.; REICHLER, R.; RENNER, B. **The childhood autism rating scale (cars)**. Western Psychological Services, 1988.

ZAGO, L.A. **Modelagem Matemática por meio de Sistemas Fuzzy: um instrumento para avaliação de Autismo**. Dissertação (mestrado). Universidade Federal de São Carlos- UFSCAR, Sorocaba. 2019.