XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ - BRAZIL

The Usage of Neural Network on Prediction of Covid-19 on Brazil

Eduardo Queiroz Pereira[1], **Gabriel Furini**[1], **Venicio Silva Araujo**[1], **Heinsten Frederich Leal dos Santos** [1]

[1]*Universidade Federal de Mato Grosso*
Avenida dos Estudandes, no 5.055, Bairro Sagrada Familia, 78736900 - Rondonópolis/MT, Brasil
eduardo.queiroz@aluno.ufr.edu.br,
gabriel.furini@aluno.ufr.edu.br,
venicio.araujo@aluno.ufr.edu.br,
heinsten.leal@ufr.edu.br

Abstract. The development of smart dynamic systems through computational models has been used extensively in recent years to predict behaviors, gaining prominence in the field of medicine with regard to the diagnosis, identification and spread of diseases. Usually associated with a probabilistic or classificatory model, these algorithms gain prominence in image diagnostics, acting in the identification of tumors, and in the prediction of dynamic behavior, such as virus propagation, due to their computational demand and easy data acquisition. Therefore, this work aims to use techniques of artificial neural networks to analyze the propagation behavior of COVID-19 in Brazil. For the architecture of the algorithm, a matrix calculation software implemented with neural network algorithms was used and the prediction model was built by reading data from infected, recovered and deaths from the initial days, with limited temporal sampling of the infection in the world. The neural network used follows the Feed-Foward model, trained with backpropagation through the Levenberg-Marquardt equation and was adjusted to predict the number of infected, recovered and deaths over the days using a pattern of behavior analysis based on the respective rates of infection, recovery and deaths. The results obtained for the forecast were validated through the conference with the data presented by the health departments for the days for which the network was configured to forecast, showing considerable precision, which increases as new data are inserted, thus showing a practical model for recurrence and analysis.

Keywords: Covid-19, Neural Network, Prediction.

1 Introduction

The field of smart algorithms, or algorithms with dynamic answers for non linearies problems, has arisen to offer improved control and prediction on applications where numerical and analytical techniques are either insufficient or impractical. The usage of models that describe systems behaviors has also been highlighted as a powerful control tool once the concept integrated with these smart algorithms working for optimization shows ups an alternative to establish management over the most various of behavior, since the administration of a business, optimization for smart structures or even heath-care monitoring for disease problems. The usage of theses algorithm has become one of the most popular research directions and plays a significant role in many fields, such as machine translation, speech recognition, image recognition, recommendation system, etc. (Sun et al. [1]).

In the scenario of heath-care, techniques as Artificial Neural Network and Genetic Algorithm has been used for lot of modeling works to describe diseases behavior or the pattern for spread of virus of any kind, in the intention to define parameters for the heath departments act to control these propagation, avoiding unnecessary costs. Modeling and optimization are the most important stages in biological processes to increase the yield of the process and to improve a system without increasing the cost (Baş and Boyaci [2]). The difficulties associated with work on analytical models comes from behaviors that shows a large quantity of variables to be describe. The introduction of these algorithm with recurrent techniques, adaptive parameters, and learning based on problem, shows up as a great alternative to modeling systems where analytical techniques would provide too complex equations.(Araujo et al. [3]).

Therefore, this work aims to the usage of an artificial neural network algorithm in intention to construct a model for prediction of behave of the spread of Covid-19 in Brazil along the days of infections since the first one.

The analysed patterns focus on the rate of infections, death and recoveries for the disease in the country in intention to provide an overview of the infection and determine the expected number for each case.

2 Artificial Neural Network

Artificial Neural Networks (ANN) are a group of computational techniques to obtain answers that present a mathematical model that corresponds to the neural structure of living organisms, managing to learn and obtain better results with training and applications (Araujo et al. [4]). Neural networks have a large number of highly interconnected processing elements (nodes) that demonstrate the ability to learn and generalize from training patterns or data. They, like humans, can perform pattern-matching tasks, while traditional computer architecture, however, is inefficient at these tasks. On the contrary, the latter is faster at algorithmic computational tasks.

Neural networks, like fuzzy logic control/decision systems, are excellent at developing human-made systems that can perform the same type of information processing than our brain performs.(Lin and Lee [5]).

There are three typical steps for building this model; the input of data, in general is used arrays to organize this data.Learning, the step where multiple data are crossed and analyzed according to stipulated conditions, producing increasingly accurate and close answers to the mathematical ideal, and data output.

The Figure 1 presents a general view for the architecture of this kind of neural network.

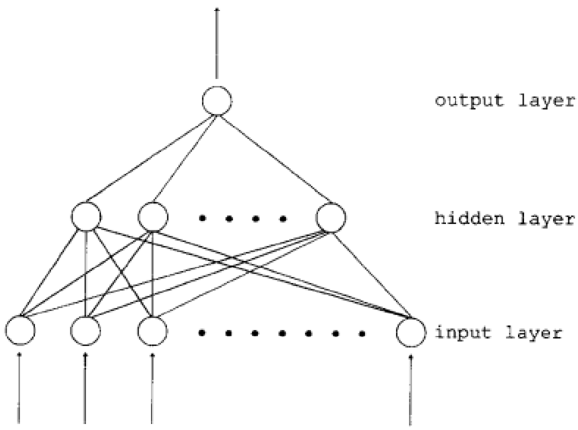

Figure 1. Typical feed-forward neural network structure

This model provides a lot of works in the area of disease diagnosis and prediction as the parkinson's disease prediction using artificial neural network (Sadek et al. [6]) or lung pattern classification for interstitial lung Diseases using a deep convolutional neural network(Anthimopoulos et al. [7]), can be also cited the diagnosis of thyroid disease using artificial neural network methods (Ozyilmaz and Uldirim [8]).

For this paper, to create the Target data, the values of infections, deaths and recoveries of Covid-19 were obtained and organized to been possible establish rates of infection, death and recovery for the virus. A exponential curve representing the behavior of the virus were produced and the cases obtained through theses ratios were defined as the Target data. The respective days of infection for each case was defined as the Input data.

2.1 Adopted Model

For this paper, to create the Target data, the values of cases of infected, deaths and recovered due to Covid-19 were organized according with days of infection, and the infection, recovery and death ratio was obtained for each day, using the relationship between cases from the previous day using equation 1, where the subscript 2 indicates cases for the posterior day and 1 the previously.

$$\alpha = \frac{cases_n}{cases_{n-1}} \tag{1}$$

Theses values were produced and organized in sets of 10 days, and averages of ratios were obtained for these sets. Through these average ratios, the concept of exponential function was used to create an expected curve for cases of infections, deaths and recovery from Covid-19 using the parametric exponential equation, seen in the equation 2, where Y is the number of cases, the subscript n indicates the day associated and β is the averages for the ratios of infections in sets of ten days. The choose of using averages ratios of 10 days was due the intention of minimize effects possible daily errors in the data presented by the health departments, getting a general behavior of the disease excluding points outside the curve.

$$\sum_{n=2}^{\infty} = \beta \cdot Y_{n-1} \tag{2}$$

The curve for cases of infections, deaths and recovery were created using this principle, and the values, respectively, for each case was set as the Target data. The input data was set as the number of days for each number of case, and theses values were processed by the neural network in the intention it finds the pattern between the number of infection, deaths and recoveries for Covid-19 along the days. The choice of the size of the vectors were think aims to balance aspects of precision and execution time of the network, focusing on the converging line of the code.

The choose of using the values for Target data as the ones produced based on the rate of infection and not the values itself was due the intention to make the neural network understand the pattern of the virus and not the pattern of the disclosure of data. The training function of the network updates weight and bias values according to Levenberg-Marquardt optimization. The choose was due the high speed answer of this function and the reliability associated with this back propagation algorithm, working in a maximum of 1000 epochs.

The error function used were the difference between of the values of real infections, deaths and recoveries and the values produced for the exponential equation for those. The code objective were minimize this difference, creating a pattern that describes the progress of the virus, became possible establish prediction on the cases along the days.

This model is not limited to the Brazil, the country the network was originally trained, with the right acquisition and setting of data is possible to use this network for any country or city that presents a pattern defined for the corona virus and it could be validated with comparison between the results predicted by the neural network and the cases of the virus along the days.

The Figure 2 shows a flowchart of the proposal models.

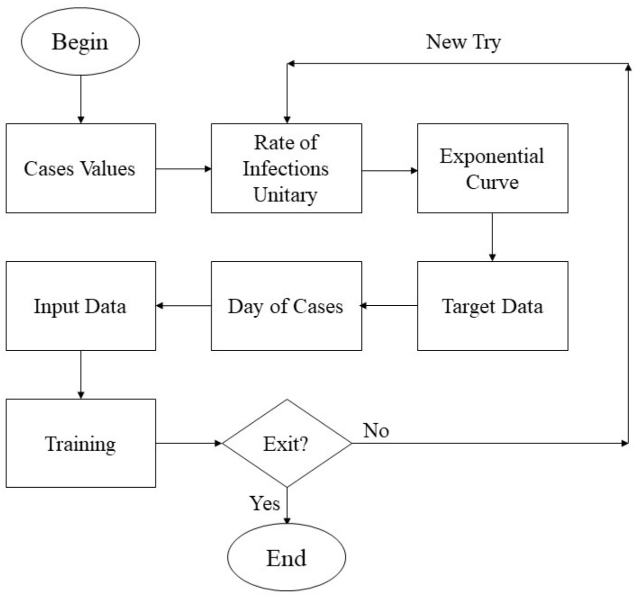

Figure 2. Flowchart of The Artificial Neural Network

3 Results

As this paper works with dynamic values for cases, considering the number of infections, deaths and recoveries increases as long as it is write is necessary to point that the graphics was plotted for the values that was available at time and the neural network itself was trained with that data.

The results for the number of infections was initially made for the first weeks of the pandemic when the neural network had enough data to give considerable results. The first results for infections follows in the Figure 3 with a comparison with the registered cases by the national heath department for the same period.

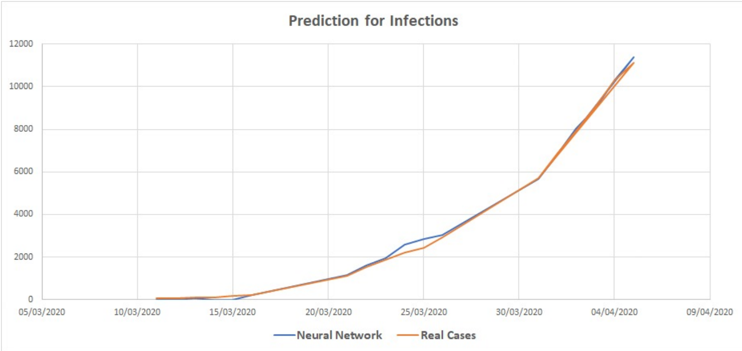

Figure 3. Infections for the First Weeks

As the neural network improve their results with acquisition of data, the Figure 4 was plotted to demonstrate the result for all the available data since the begin of this paper. Is notable as the precision of the neural network were improved since the values for the first weeks, becoming almost indistinguishable the results for the neural network and the provided by the health department.

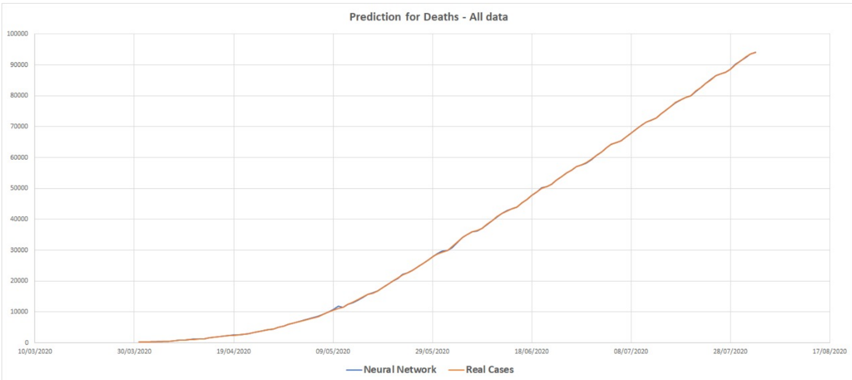

Figure 4. Infections for All Data

In intention to measure the precision of the neural network along the time, a graphic was plotted comparing the percentage error of the network with the acquisition of data, using the number of infections as the reference. It follows in the Figure 5, where it's seen a percentage error close to 20% for the first weeks that converge to 0% with the acquisition of previously values, feature of an well trained neural network.

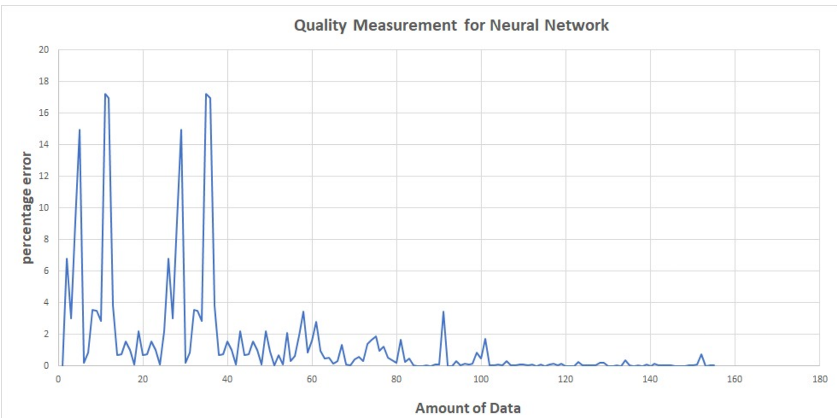

Figure 5. Percentage Error of The Neural Network

The same methodology was adopted for the number of deaths, the values for the first week follows in the Figure 6.

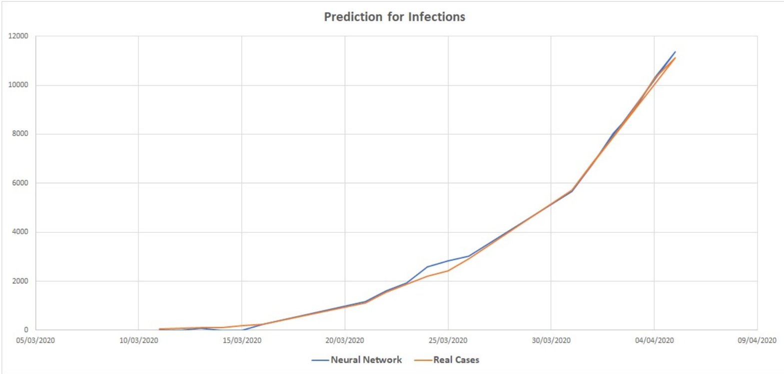

Figure 6. Deaths for the First Weeks

The graphic for deaths for the all data it's seen in the Figure 7.

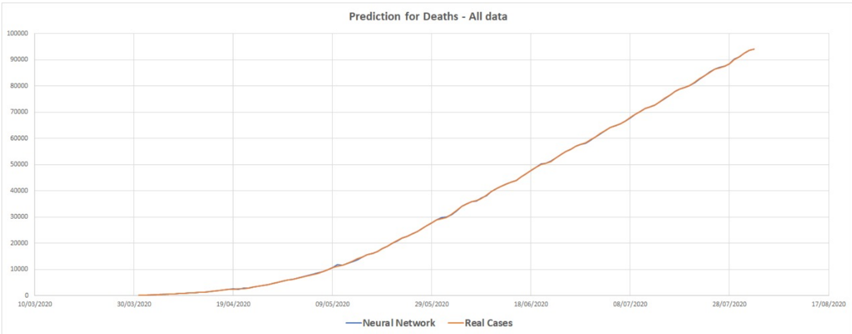

Figure 7. Deaths for All Data

For the recoveries' number provided by the national health department too less data than to the infections or death's number, so, to avoid a poor margin of precision for the first weeks, only an all data graphic was plotted, it follows in the Figure 8.

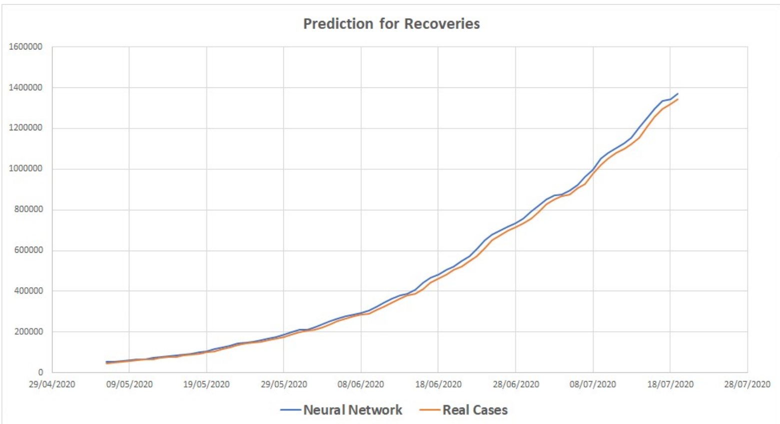

Figure 8. Recoveries for All Data

Was also plotted, to make visual the effects of working with less quantity of data, a graphic comparing the precision for the neural networking when it was set to predict the number of recoveries. It's seen in the Figure 9 a feature of improvement, but not well behaved, on the quality measurement for the neural network associated with this poor data.

It's seen in the Figure 9 a decedent feature for the percentage error of the neural network, it shows the learning feature of this nature of algorithm, although due working with less than half of the data for the infections this graphics presents a less reliable results for consequence, converging from almost 20% to 2% of error of the provided results for recoveries cases.

For the conference of results provided by the neural network, the table 1 to 3 were added to this work with the values for infections, deaths and recoveries since the firsts weeks to the last computed for each case.

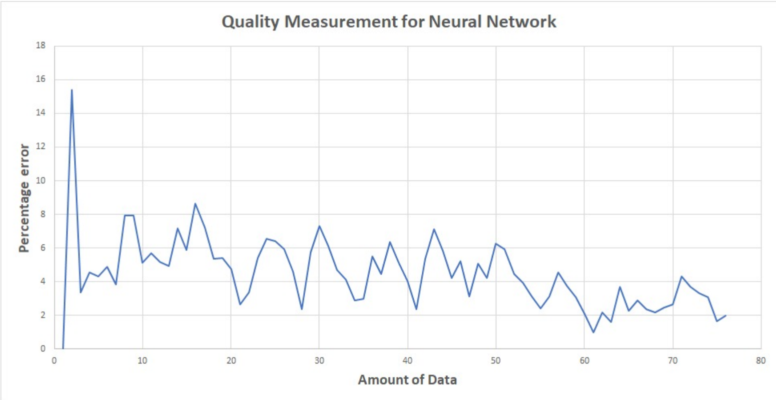

Figure 9. Percentage Error for Recoveries

Week	Neural Network Prediction	Real Cases	Percentage Error
1	2	2	0,000%
3	100,43	121	17,000%
5	3317,55	3903	15,000%
7	20105,19	20727	3,000%
9	56461,185	58509	3,500%
10	95432,04	96396	1,000%
22	2286654	2287475	0,036%

Table 1. Number of Infections by Week

Week	Neural Network Prediction	Real Cases	Percentage Error
4	1	1	0,000%
6	230	201	14,428%
8	1628	1532	6,266%
10	5800	5017	15,607%
12	13612	13149	3,521%
14	25300	25598	1,164%
24	94172	94104	0,072%

Table 2. Number of Deaths by Week

Week	Neural Network Prediction	Real Cases	Percentage Error
11	59221	51370	15,283%
13	125597	116683	7,640%
15	249942	238617	4,7646%
17	474018	463474	2,275%
19	845470	826866	2,250%
21	1233150	1209208	1,980%
22	1398653	1371229	2,000%

Table 3. Number of Recoveries by Week

4 Conclusions

Was observed that the proposed model accurately gauges the number of infections, deaths and recoveries of Covid-19 on Brazil with the correct quantity of data necessary for a good training, being a powerful tool to consult for administration decision made by the health departments.

It's also seen that the neural network increases its performance with the amount of data, being able to achieve a highly precision to stipulate futures values based on previously insertions not getting limited for a country, working in cities, states or wherever presents a clearly behavior for the virus.

An additional advantage of this method is its learning and adaptive features that made the code ignore punctual human errors in the information of values by the health departments, and works analyzing a general pattern for the virus not any kind of pattern for data release.

References

[1] Sun, S., Cao, Z., Zhu, H., & Zhao, J., 2019. A survey of optimization methods from a machine learning perspective. volume 2.

[2] Baş, D. & Boyaci, I. H., 2007. Modeling and optimization i: Usability of response surface methodology. *Journal of Food Engineering*, vol. 78, pp. 836–845.

[3] Araujo, V. S., Silva, G. S., & Santos, H. F. L., 2019a. Neural network and shunt control for beam structure: Multimodal approach. In *25th International Congress of Mechanical Engineering - COBEM 2019*, Uberlândia, Brazil.

[4] Araujo, V. S., Silva, G. S., & Santos, H. F. L., 2019b. Shunt control on smart strucutres using genetic algorithm and neural network method. In *Proceedings of XVIII Internation Symposium on Dynamic Problems of Mechanics - DINAME 2019*, Buzios, Brazil.

[5] Lin, C. & Lee, C., 1991. Neural-network-based fuzzy logic control and decision system. *IEEE Trans. Comput*, vol. 40, pp. 1320–1336.

[6] Sadek, R. M., Mohammed, S. A., Abunbehan, A. R. K., Ghattas, A. K. H. A., Badawi, M. R., Mortaja, M. N., Abu-Nasser, B. S., & Abu-Naser, S. S., 2019. Parkinson's disease prediction using artificial neural network. *International Journal of Academic Health and Medical Research (IJAHMR)*, vol. 3, pp. 1–8.

[7] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., & Mougiakakou, S., 2016. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, vol. 30, pp. 5.

[8] Ozyilmaz, L. & Uldirim, T., 2016. The diagnosis of thyroid disease using artificial neural network methods. volume 4.