Genetic Optimization of Ensemble Neural Network Architectures for Prediction of COVID-19 Confirmed and Death Cases

Julio C. Mónica, Patricia Melin, and Daniela Sánchez

Abstract  In this work a genetic algorithm for ensemble neural network architecture optimization applied to COVID-19 time series prediction is proposed. The main objective of this paper is to show the results of the optimized number of neurons in two hidden layers of an ensemble artificial neural network used for time series prediction using a real genetic algorithm. The time series dataset used in this work is the confirmed and death cases of COVID-19 of 12 states of Mexico (and information about the whole country). Being the COVID-19 the pandemic that has been affecting many lives in Mexico, for this reason, this work seeks to find a prediction for confirmed and death cases in this country.

Keywords  Genetic algorithm · Time series prediction · Artificial neural network · Fuzzy logic · NAR · Optimization

1 Introduction

The artificial neural networks (ANNs) and the genetic algorithms (GAs) are two techniques that have their own strengths and weaknesses. The ANNs have the learning but their main problem is how to establish their architecture, this problem can be resolved using an optimization technique such as the GAs. Nowadays, hybrid systems have been of vital importance to solve problems in a better way, it is a valuable application in technological advances, covering a large number of different areas, such as, medicine, finances, industry, among many others (Mahajan 2013).
The proposed method is applied using genetic algorithms to optimize the numbers of neurons in the two hidden layers, having as objective function minimizing the errors of the 10 days ahead prediction of the ensemble artificial neural network. The ANNs have been used in other works, a prediction model was constructed to deal with the fatigue of polymer using artificial neural networks (Yang 2020). For the analysis of the spatial evolution of coronavirus pandemic around the world using Self Organizing Maps (Melin et al. 2020). The ANNs have also been used to identify aggression, hate speech, cyberbullying, etc. on tweets (Sadiq et al. 2020). To optimize a modular granular neural network using a firefly algorithm for human recognition (Sánchez et al. 2017), it was also used for series prediction of the Mexican Stock using Particle swarm optimization of ensemble neural networks with fuzzy aggregation (Pulido et al. 2014). The genetic algorithms have been used for feature selection for early time series classification (Ahn and Hur 2020), among elder drivers to predict the severity of fixed object crashes (Amiri et al. 2020). And have been combined with GA such as in predicting a solar space heating system performance (Jamali et al. 2019), for the electricity consumption of a real-world campus building (Luo et al. 2020), in a method designed for human body shape prediction (Cheng 2020), in a model for the vehicle insurance fraud identification (Yan et al. 2020), for the optimization of modular granular neural networks where a hierarchical genetic algorithm for human recognition was used using the ear biometric measure (Sánchez and Melin 2014) and for the prediction of soil wind erodibility (Kouchami-Sardoo et al. 2020) among many others.

This paper is organized as follows: Sect. 2 describes the concepts used in this work, the Sect. 3 contains the general architecture of the proposed method, Sect. 4 shows the experimental results and Sect. 5 presents the conclusion of this work.

2 Basic Concepts

In this document we present the basic concepts that were used in this work.

2.1 Artificial Neural Networks

Artificial neural networks (ANN) are computing systems that imitate the mechanism of the biological neural network. The ANN can be used in a supervised or unsupervised learning method. The components of the models are nodes, weights and layers. There are many different types of ANN. The FITNET is a commonly used Multi-Layer Perceptron (MLP) which is a class of feedforward artificial neural network containing an input layer, a hidden layer and an output layer (Ali and Shahbaz 2020; Cheshmberah et al. 2020). The general architecture of an ANN is shown in Fig. 1.

Where $X_1$, $X_2$ and $X_n$ are the variables of the inputs layer, $n$ is the number of neurons, $W_{ij}$, $W_{kj}$, $W_{lj}$ are the weights and $y$ is the output (Kim 2020).
2.2 Nonlinear Autoregressive Neural Network

The nonlinear autoregressive neural network (NAR) to estimate values uses past values of the time series. The NAR model consists of one input layer, one or more hidden layers and one output layer. NAR is a recurrent and dynamic network with feedback connections. In time series forecasting NAR is used in one-step-ahead or multi-step-ahead (Khan 2020; Ruiz et al. 2016). In the following Eq. (1) we can see how the NAR model is mathematically expressed.

$$y(t) = F(y(t-1), y(t-2), \ldots, y(t-d))$$

(1)

where $y(t)$ is the output value of the considered time series $y$ at time $t$, and $d$ is the time delay and $F$ denotes the transfer function (Benrhmach 2020; Pan and Pu 2019). In Fig. 2, the NAR neural network is shown.

An ensemble neural network (ENN) is a model for learning where the artificial neural networks are together to perform much better when solving robust problems (Chen et al. 2019; Zhou et al. 2002).

2.3 Fuzzy Logic

Fuzzy logic was introduced by Lotfi A. Zadeh in 1965 with the proposal of fuzzy set theory. In the concept of fuzzy logic, the truth values of each variable may fall in between the range of 0 and 1 unlike in the concept of Boolean logic where the variable values are either 0 (represents false) or 1 (represents true) (Chiueh 1991; Reddy et al. 2017). Fuzzy logic is a methodology that solves computational problems...
based on linguistic variables in an environment of imprecision or incompleteness of information. One of the fuzzy logic misconceptions is that fuzzy logic is fuzzy. Fuzzy logic is a precise conceptual system of reasoning and deduction where the analysis is associated with imperfect information. Fuzzy logic uses two remarkable human capabilities, the first one, the capability to communicate, reason, converse and make rational decisions. The second, the capability to perform a wide variety of physical and mental tasks (Zadeh 2009).

### 2.4 Genetic Algorithms

Genetic algorithm (GA), is one of the optimization algorithms. It is based on the Darwinian theory as their inspiration by natural evolution, the survival of the fittest, this is measured by an evaluation function called fitness function in computational terms (Sánchez 2017). Using the principle of evolution, the genetic algorithm techniques are based on many parameters that are required but the critical are, population size, selection, crossover and mutation rate to obtain the best solutions for optimization problems. A chromosome is a sequence of genes, in Fig. 3 a chromosome with genes is shown (Chiroma et al. 2017).
The population is a group of individuals, a chromosome is an individual. The reason for crossover in genetic algorithms is to help the reproduction of better chromosomes and the mutation of the strings is needed because genetic material may be lost in the crossover process, in order to fix the distortion of genetic information we do it through mutation. In Fig. 4 we show the mutation process using a single point (Chiroma et al. 2017).

3 Proposed Method

The proposed method consists of an ensemble neural network with 3 artificial neural networks, its optimal architecture is achieved using a real genetic algorithm technique. The main goal is to obtain the optimal number of neurons in the two hidden layers. We have a dataset of COVID-19 confirmed and death cases of 12 states of Mexico and the information of the whole country, the data of each state is learned from a Fitnet and two NAR neural networks, where each one obtains a prediction and normalized errors are calculated. The normalized errors are sent to a type-1
fuzzy logic integrator to finally give us a weight for each prediction and calculate a final prediction. The real genetic algorithm optimizes the numbers of neurons in the two hidden layers of each artificial neural network. This model is presented in Fig. 5. The values of the ANN are shown in Table 1. And for the genetic algorithm in Table 2. The training function used for Fitnet neural network is \textit{trainlm (Levenberg-Marquardt backpropagation)} and for NAR is \textit{trainlm} and \textit{trainbr (Bayesian regularization backpropagation)}.

For the fuzzy integrator system, where the inputs $e_1, e_2$ and $e_3$ are the normalized mean square errors of the values in the range between 0 and 1 of the 3 artificial neural networks being used to produce the weights $w_1, w_2$ and $w_3$ and $p_1, p_2, p_3$ are the predicted values of each module respectively. Then we combine the predictions to

![Proposed model](image)

**Fig. 5** Proposed model

| Parameters of the ENN | Values          |
|----------------------|-----------------|
| Modules              | 3               |
| Days to predict      | 10              |
| Error goal           | 0.001           |
| Training function    | Trainlm, Trainbr|
| Hidden Layers        | 2               |
| Neurons for each hidden layer | 1–50          |

**Table 1** Values for the ENN

| Parameters of the GA | Values |
|----------------------|--------|
| Individuals          | 10     |
| Generations          | 30     |
| Selection            | Tournament |
| Mutation Rate        | 0.3    |
| Crossover            | Uniform |

**Table 2** Values for the genetic algorithm
obtain the final prediction $PT$ using the expression in Eq. (2).

$$PT = \frac{w_1 p_1 + w_2 p_2 + w_3 p_3}{w_1 + w_2 + w_3}$$  \hspace{1cm} (2)$$

The fuzzy inference system has 3 rules, 3 inputs $e_1, e_2, e_3$ and 3 outputs $w_1, w_2, w_3$ using Gaussian functions. In Fig. 6, the fuzzy inference system is shown.

The fuzzy system contains 3 rules which are the following:

1. If ($e_1$ is small) and ($e_2$ is medium) and ($e_3$ is large) then ($w_1$ is high) ($w_2$ is medium) ($w_3$ is small).
2. If ($e_1$ is large) and ($e_2$ is small) and ($e_3$ is medium) then ($w_1$ is small) ($w_2$ is high) ($w_3$ is medium).
3. If ($e_1$ is medium) and ($e_2$ is large) and ($e_3$ is small) then ($w_1$ is medium) ($w_2$ is small) ($w_3$ is high).

Based on their corresponding errors these rules express the knowledge of how to combine predictions.

The real genetic algorithm used in this work after the initial population has been created randomly at each generation a selection, crossover and mutation is performed to update the population of solutions and the tournament selection is used to select the individuals from the population to continue with the mating pool where the individuals known as parents are paired and crossover is applied to produce two offspring (Sawyerr et al. 2014). First, the whole population is evaluated based on
the fitness function, then after the individual is ranked, the best fitness is saved to avoid modifications, process known as elitism. When the individuals from each generation of the population are selected for reproduction the uniform crossover combines the two chromosomes known as parents to produce new chromosomes known as offspring, the main goal is to attempt obtaining offspring better than their parents. In Fig. 7 a uniform crossover for a real genetic algorithm is shown.

Where $P_1, P_2$ are the chromosome parents and $O_1, O_2$ are the offspring obtained after the crossover, $r$ are the random numbers generated to determine if the crossover occurs and $\alpha$ is the crossover rate.

In this work the fitness function used is the mean square error of the final prediction to obtain the optimization of the neurons in the two hidden layers of each artificial neural network.

### 4 Results of the Experiment

The optimized results using the real generic algorithm technique and all the numbers of neurons of the results are shown in this section.

#### 4.1 Genetic Algorithms

The architectures obtained by the optimization technique using the genetic algorithm are shown in Table 3 of the COVID-19 confirmed cases, where the results show that
Table 3  The results of the COVID-19 confirmed cases

| States                  | Type of ANN | Neurons | Final error |
|-------------------------|-------------|---------|-------------|
| Baja California        | FITNET      | 39,31   | 59.07       |
|                         | NAR         | 34,42   |             |
|                         | NAR         | 24,32   |             |
| Ciudad de Mexico       | FITNET      | 14,14   | 2718.96     |
|                         | NAR         | 41,35   |             |
|                         | NAR         | 14,10   |             |
| Coahuila               | FITNET      | 27,31   | 38.51       |
|                         | NAR         | 40,14   |             |
|                         | NAR         | 26,15   |             |
| Estado de Mexico       | FITNET      | 42,10   | 45.56       |
|                         | NAR         | 40,24   |             |
|                         | NAR .3      | 50,37   |             |
| Jalisco                | FITNET      | 47,7    | 2.62        |
|                         | NAR         | 1,23    |             |
|                         | NAR         | 41,30   |             |
| Nuevo Leon             | FITNET      | 19,6    | 37.92       |
|                         | NAR         | 39,9    |             |
|                         | NAR         | 34,49   |             |
| Puebla                 | FITNET      | 9,50    | 17.57       |
|                         | NAR         | 34,1    |             |
|                         | NAR         | 6,27    |             |
| Quintana Roo           | FITNET      | 39,48   | 56.94       |
|                         | NAR         | 18,13   |             |
|                         | NAR         | 26,13   |             |
| Sinaloa                | FITNET      | 11,18   | 19.32       |
|                         | NAR         | 24,27   |             |
|                         | NAR         | 21,38   |             |
| Tabasco                | FITNET      | 11,18   | 85.98       |
|                         | NAR         | 24,27   |             |
|                         | NAR         | 21,38   |             |
| Veracruz               | FITNET      | 11,18   | 53.59       |
|                         | NAR         | 25,26   |             |
|                         | NAR         | 21,38   |             |
| Yucatan                | FITNET      | 32,8    | 17          |
|                         | NAR         | 1,7     |             |
|                         | NAR         | 46,31   |             |

(continued)
the state with less final error is Jalisco.

In Table 4, the architecture obtained for the death cases are shown, where the results show that the states with less final error are Coahuila and Jalisco.

The comparison with non-optimized results presented (Pulido et al. 2014) for confirmed cases are shown in Table 5.

The comparison with non-optimized results presented (Melin et al. 2020) for death cases are shown in Table 6.

### 5 Conclusions

The proposed real generic algorithm used to find the optimal neural network ensemble architecture for the COVID-19 time series prediction of the death and confirmed cases of 12 states of Mexico and the total data of the country helped us find the optimal number of neurons in the two hidden layers of the artificial neural networks (FITNET and NAR). The method used as an integration was a fuzzy interference system where the inputs received the prediction errors of the artificial neural networks to give us a final prediction. We obtained good results, as we can notice in the Tables 5 and 6 the optimized results of the final errors gave us a big improvement. One of the aspects of using the real genetic algorithm to consider is the time consuming, nowadays the hybrid systems need to get results in a faster way and that is a weakness while using a genetic algorithm in certain cases. For future work, different optimization techniques will be implemented to compare results and other application areas could be considered, like in (Castillo 1998; Castillo and Melin 2003; Sanchez et al. 2014).
| States            | Type of ANN | Neurons | Final error |
|-------------------|-------------|---------|-------------|
| Baja California  | FITNET      | 16,20   | 11.09       |
|                   | NAR         | 43,1    |             |
|                   | NAR         | 14,33   |             |
| Ciudad de Mexico  | FITNET      | 16,27   | 8.18        |
|                   | NAR         | 16,10   |             |
|                   | NAR         | 27,14   |             |
| Coahuila          | FITNET      | 41,40   | 0.11        |
|                   | NAR         | 44,42   |             |
|                   | NAR         | 39,9    |             |
| Estado de Mexico  | FITNET      | 18,14   | 7.32        |
|                   | NAR         | 20,37   |             |
|                   | NAR .3      | 40,17   |             |
| Jalisco           | FITNET      | 6,15    | 0.11        |
|                   | NAR         | 25,3    |             |
|                   | NAR         | 10,28   |             |
| Nuevo Leon        | FITNET      | 36,8    | 0.19        |
|                   | NAR         | 37,36   |             |
|                   | NAR         | 28,23   |             |
| Puebla            | FITNET      | 7,10    | 2.74        |
|                   | NAR         | 45,49   |             |
|                   | NAR         | 13,15   |             |
| Quintana Roo      | FITNET      | 28,8    | 1.15        |
|                   | NAR         | 11,33   |             |
|                   | NAR         | 9,32    |             |
| Sinaloa           | FITNET      | 37,8    | 2           |
|                   | NAR         | 21,15   |             |
|                   | NAR         | 27,29   |             |
| Tabasco           | FITNET      | 33,10   | 2.37        |
|                   | NAR         | 27,3    |             |
|                   | NAR         | 12,39   |             |
| Veracruz          | FITNET      | 17,50   | 2.19        |
|                   | NAR         | 23,23   |             |
|                   | NAR         | 31,41   |             |
| Yucatan           | FITNET      | 7,42    | 0.20        |
|                   | NAR         | 38,31   |             |
|                   | NAR         | 48,5    |             |

(continued)
**Table 4**  (continued)

| States    | Type of ANN | Neurons | Final error |
|-----------|-------------|---------|-------------|
| Nacional  | FITNET      | 31,19   | 115.65      |
| NAR       | 36,50       |         |             |
| NAR       | 13.6        |         |             |

**Table 5**  Comparison of the COVID-19 confirmed cases

| State                  | Final error  |
|------------------------|--------------|
|                        | Non-optimized| Optimized   |
| Baja California        | 2529.07      | 59.07       |
| Ciudad de Mexico       | 1263297.77   | 2718.96     |
| Coahuila               | 109.83       | 38.51       |
| Estado de Mexico       | 41570.11     | 45.56       |
| Jalisco                | 1055.77      | 2.62        |
| Nuevo Leon             | 131.86       | 37.92       |
| Puebla                 | 5516.75      | 17.57       |
| Quintana Roo           | 7513.09      | 56.94       |
| Sinaloa                | 74.22        | 19.32       |
| Tabasco                | 56.71        | 85.98       |
| Veracruz               | 9528.70      | 53.59       |
| Yucatan                | 3811.82      | 17          |
| National               | 2415010.109  | 12190.84    |

**Table 6**  Comparison of the COVID-19 death cases

| State                  | Final error  |
|------------------------|--------------|
|                        | Non-optimized| Optimized   |
| Baja California        | 1119.39      | 11.09       |
| Ciudad de Mexico       | 202.10       | 8.18        |
| Coahuila               | 22.50        | 0.11        |
| Estado de Mexico       | 2578.22      | 7.32        |
| Jalisco                | 24.95        | 0.11        |
| Nuevo Leon             | 2.69         | 0.19        |
| Puebla                 | 20.73        | 2.74        |
| Quintana Roo           | 254.18       | 1.15        |
| Sinaloa                | 168.05       | 2           |
| Tabasco                | 61.23        | 2.37        |
| Veracruz               | 241.81       | 2.19        |
| Yucatan                | 80.21        | 0.20        |
| National               | 28901.55     | 115.65      |
References

Ahn, G., and S. Hur. 2020. Efficient genetic algorithm for feature selection for early time series classification. Computers & Industrial Engineering 142: 106345.

Ali, S., and M. Shahbaz. 2020. Streamflow forecasting by modeling the rainfall–streamflow relationship using artificial neural networks. Modeling Earth Systems and Environment 6 (3): 1645–1656.

Amiri, A., A. Sadri, N. Nadimi, and M. Shams. 2020. A comparison between artificial neural network and hybrid intelligent genetic algorithm in predicting the severity of fixed object crashes among elderly drivers. Accident Analysis and Prevention 138: 105468.

Benrhmach, G., K. Namir, A. Namir, and J. Bouyaghroumni. (2020). Nonlinear autoregressive neural network and extended Kalman filters for prediction of financial time series. Journal of Applied Mathematics.

Castillo, O., and P. Melin. 1998. A new fuzzy-fractal-genetic method for automated mathematical modelling and simulation of robotic dynamic systems. 1998 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 1998) Proceedings. 2, 1182–1187.

Castillo, O., and P. Melin. 2003. Intelligent adaptive model-based control of robotic dynamic systems with a hybrid fuzzy-neural approach. Applied Soft Computing 3 (4): 363–378.

Chen, Y., H. Chang, J. Meng, and D. Zhang. 2019. Ensemble Neural Networks (ENN): A gradient-free stochastic method. Neural Networks 110: 170–185.

Cheng, P., D. Chen, and J. Wang, (2020). Clustering of the body shape of the adult male by using principal component analysis and genetic algorithm–BP neural network. Soft Computing, 1–19.

Cheshmberah, F., H. Fathizad, G. Parad, and S. Shojaeifar. 2020. Comparison of RBF and MLP neural network performance and regression analysis to estimate carbon sequestration. International Journal of Environmental Science and Technology 17 (9): 3891–3900.

Chiroma, H., S. Abdulkareem, A. Abubakar, and T. Herawan. 2017. Neural networks optimization through genetic algorithm searches: a review. Applied Mathematics & Information 11 (6): 1543–1564.

Chiueh, T. C. (1991, January). Optimization of fuzzy logic implementation. In 1991 Proceedings of the Twenty-First International Symposium on Multiple-Valued Logic (pp. 348–349).

Jamali, B., M. Rasekh, F. Jamadi, R. Gandomkar, and F. Makiabadi. 2019. Using PSO-GA algorithm for training artificial neural net-work to forecast solar space heating system parameters. Applied Thermal Engineering 147: 647–660.

Khan, F.M., and R. Gupta. (2020). ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. Journal of Safety Science and Resilience.

Kim, M.K., Y.S. Kim, and J. Srebric. (2020). Predictions of electricity consumption in a campus building using occupant rates and weather elements with sensitivity analysis: Artificial neural network vs. linear regression. Sustainable Cities and Society, 102385.

Kouchami-Sardoo, I., H. Shirani, I. Esfandiarpour-Boroujeni, A. Besalatpour, and M. Hajabbasi. 2020. Prediction of soil wind erodibility using a hybrid Genetic algorithm—Artificial neural network method. CATENA 187: 104315.

Luo, X., L. Oyedele, A. Ajayi, O. Akinade, J. Delgado, H. Owolabi, and A. Ahmed. 2020. Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings. Energy and AI 2: 100015.

Mahajan, R., and G. Kaur. 2013. Neural networks using genetic algorithms. International Journal of Computer Applications, 77(14).

Melin, P. J. Monica, D. Sanchez, and O. Castillo. 2020a. Multiple ensemble neural network models with fuzzy response aggregation for predicting COVID-19 time series: The case of Mexico. Healthcare 8 (2); 181.

Melin, P., J. Monica, D. Sanchez, and O. Castillo. 2020b. Analysis of spatial spread relationships of coronavirus (COVID-19) pandemic in the world using self organizing maps. Chaos, Solitons & Fractals 138: 109917.
Pan, Y., L. Chen, J. Wang, H. Ma, S. Cai, S. Pu, ... and E. Li. 2019. Research on deformation prediction of tunnel surrounding rock using the model combining firefly algorithm and nonlinear auto-regressive dynamic neural network. *Engineering with Computers*, 1–11.

Pulido, M., P. Melin, and O. Castillo. 2014. Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange. *Information Sciences* 280: 188–204.

Reddy, G.P., Y. Deepika, K.S. Prasad, and G.K. Kumar. 2017. Fuzzy Logics associated with neural networks in the real time for better world. *Materials Today: Proceedings* 4 (8): 8507–8516.

Ruiz, L.G.B., M.P. Cuéllar, M.D. Calvo-Flores, and M.D.C.P. Jiménez. 2016. An application of non-linear autoregressive neural networks to predict energy consumption in public buildings. *Energies* 9 (9): 684.

Sadiq, S., A. Mehmood, S. Ullah, M. Ahmad, G. Choi, and B. On. 2020. Aggression detection through deep neural model on Twitter. *Future Generation Computer Systems* 114: 120–129.

Sánchez, D., P. Melin, J. Carpio, and H. Puga. 2017. Comparison of optimization techniques for modular neural networks applied to human recognition. In *Nature-inspired design of hybrid intelligent systems* (pp. 225–241). Springer, Cham.

Sánchez, D., and P. Melin. 2014. Optimization of modular granular neural networks using hierarchical genetic algorithms for human recognition using the ear biometric measure. *Engineering Applications of Artificial Intelligence* 27: 41–56.

Sanchez, M.A., O. Castillo, J.R. Castro, and P. Melin. 2014. Fuzzy granular gravitational clustering algorithm for multivariate data. *Information Sciences* 279: 498–511.

Sánchez, D., P. Melin, and O. Castillo. 2017b. Optimization of modular granular neural networks using a firefly algorithm for human recognition. *Engineering Applications of Artificial Intelligence* 64: 172–186.

Sawyerr, B., A. Adewumi, and M. Ali. 2014. Real-coded genetic algorithm with uniform random local search. *Applied Mathematics and Computation* 228: 589–597.

Yan, C., M. Li, W. Liu, and M. Qi. 2020. Improved adaptive genetic algorithm for the vehicle insurance fraud identification model based on a bp neural network. *Theoretical Computer Science* 817: 12–23.

Yang, J., G. Kang, Y. Liu, K. Chen, and Q. Kan. 2020. Life prediction for rate-dependent low-cycle fatigue of PA6 polymer considering ratchetting: Semi-empirical model and neural network based approach. *International Journal of Fatigue*, 105619.

Zadeh, L.A. 2009. Toward extended fuzzy logic—A first step. *Fuzzy Sets and Systems* 160 (21): 3175–3181.

Zhou, Z.H., J. Wu, and W. Tang. 2002. Ensembling neural networks: many could be better than all. *Artificial Intelligence* 137 (1–2): 239–263.