Interval Computing in Neural Networks

Raquel E. Patiño Escarcina[1], Benjamín R. Callejas Bedregal[1], and Aarão Lyra[2]

[1] Universidade Federal do Rio Grande do Norte
Laboratorio de Lógica e Inteligência Computacional LABLIC - DIMAP
59078-970 Natal, RN, Brasil
raquel@ppgsc.ufrn.br, bedregal@dimap.ufrn.br
[2] Universidade Potiguar
59056-000 Natal, RN, Brasil
aarao@unp.br

Abstract. In many applications data is naturaly interval and it is important to have tools that work with this kind of data. This work presents the *Interval computing in Neural Networks*. The neural network proposed is based on Punctual Neural Networks, and try to be a solution for the problems in calculus precision error and treatment of interval data without modifie it. Beyond it, it is believed that interval connections between neurons permit the number of the epochs needed to converge to be lower than the needed in punctual networks without loss efficiency.
The interval computing in a neural network with supervisioned training was tested in comparison with a traditional neural network and the result is show. The result shows that the behavior of the neural network can be better than traditional and including guaranteed about the computational errors.

1 Introduction

Artificial Neural Networks are used to solve problems in different science fields. They represent a simple way to find solutions for highly complex problems. In real-life situations, this problems have data sets of physical quantites but it is not so easy to obtain measures that represent these physical quantities exactly. In many cases, it is extremely important to know how different the actual value of data can be from our real data. To achieve reliability, both mathematical and computationally, all quantities are represented by the smallest machine representable intervals where the physical quantities belong to the intervals. This concept was originally introduced by [9], to the field of reliable numerical computations. Later the concept of the interval value has been successfully applied in regression analysis [12], fuzzy-interval [6], principal Component Analysis [3],etc. Artificial neural networks are viewed as a parallel computational model, with varying degrees of complexity, comprised of densely interconnected adaptive processing units. A very important feature of these networks is their adaptive nature, where learning by example replaces traditional programming in solving problems. Tthe computational process envisioned with artificial neural networks is as follows: an artificial neuron or processing element receives inputs from a number of other neurons or from an external stimulus. A weighted sum of these inputs constitutes the argument to an activation function or transfer function. This activation function is generally nonlinear. The resulting value of the activation function is the output of the neuron, this output gets distributed or fanned out along weighted connections to other neurons. The actual manner in which these connections are made defines the information flow in the network and is called architecture of the network. The weighted connections in these architectures play an important role, such that these network are also called connectionist models of computation. The method used to adjust the weights in the process of training the network is called the learning rule. In summary, the three essential ingredients of a computational system based on ANNs are the transfer function, the architecture, and the learning rule.

2 Interval Mathematic

The Interval Mathematic is a theory introduced by R. E. Moore [9] in the 60's. It was thought to replay questions about accuracy and efficiency that appear in scientific computation practice and numeric resolution of problems.
Data with interval values is used in many applications to represent certain degree of uncertainty, for example measures, variables, extremes behaviors, etc. Dealing with uncertainty is still one of the central topics in artificial intelligence. Uncertainty comes as a result of incompleteness of our observations, measurements and estimations of the world.
The interval technique is an alternative to obtain results with warranted limits in the scientific computation because it has rigorous and automatic control of errors. [7]
In interval mathematic, an interval is a closed non-empty set of real numbers, where $\underline{x}$ is the lower bound, $\overline{x}$ is the upper bound and $\underline{x} < \overline{x}$ and This set is a real interval or only interval will be denoted by $X = [\underline{x}, \overline{x}]$. In the following, lower case letters denote elements of the real numbers and capital letters to denote interval quantities and vectors, respectively. An interval X where $\underline{x} = \overline{x}$ is called *Degenerated Interval*. Let A, B intervals, the operations used in this work, are:

– Addition, is defined by the equation:

$$A + B = [\underline{a}, \overline{a}] + [\underline{b}, \overline{b}] = [\underline{a} + \underline{b}, \overline{a} + \overline{b}] \tag{1}$$

– Subtration, is defined by the equation

$$A - B = A + (-B) = [(\underline{a} - \overline{b}); (\overline{a} - \underline{b})] \tag{2}$$

– Multiplication of them is defined by the equation:

$$A \cdot B = [\underline{a}, \overline{a}] \cdot [\underline{b}, \overline{b}] = [min\{\underline{ab}, \overline{a}\underline{b}, \underline{a}\overline{b}, \overline{ab}\}, max\{\underline{ab}, \overline{a}\underline{b}, \underline{a}\overline{b}, \overline{ab}\}] \tag{3}$$

– The distance of the interval A until the interval B is defined by:

$$dist(A, B) = dist([\underline{a}, \overline{a}], [\underline{b}, \overline{b}]) = max\{|\underline{a} - \underline{b}|; |\overline{a} - \overline{b}|\} \tag{4}$$

- The size of the interval A is defined by:

$$size(A) = size([\underline{a}, \overline{a}]) = (\overline{a} - \underline{a}) \tag{5}$$

  An interval X where $size(X) = 0$ is called *Symmetrical Interval*.
- The sign of the interval X is defined by:

$$sign(A) = sign([\underline{a}, \overline{a}]) = (\overline{a} + \underline{a}) \tag{6}$$

More information about interval mathematic and applications in [1], [9], [10], [5]

3 Interval Computing in Neural Networks

In [11] was proposed some methods for processing interval data input in neural networks. These methods are the most basic and intuitive form to work with intervals but these not guarantee that the resultant data set represents the real data set of the problem. In [4] each interval in the input data set is a set of the possible values formed with opinion of the experts about one variable, here was used the mathematic interval to calculated the output of network. This work is a proposal of solution to the problem of interval data but doesnt show other advantages with the interval mathematic. In [2] was proponed a neural network with interval weights but the solution isnt more simple and the network is modeled like the problem of the solution equations. These works represent the preoccupation of the investigators around the importance to have algorithms can work with interval data. This approach is a contribution for solving part of the problems seen before.
A Neural Network is said to be an *Interval Neural Network* if almost one of its input, output and weight sets have interval values.
Interval Neural Networks (IANN) are formed by processing units called Interval Neurons. They are based on the neuron model proposed by McCulloc-Pitts. This neuron is prepared for receiving interval or real data. An Interval Neuron is formed by three functions (figure 1):

- *Normalizer function*($\mathbb{T}$): This function analyze the input nature and normalize it in order to have only interval inputs.
- *Sum function*(Σ): This function is the same as the sum function in the neurons of the traditional neural networks. Join in a linear way inputs with their respective synaptic weights.
- *Activation function*: This function could be any interval derivable linear function. Restrict the output to interval values between [-1 1] or [0 1].

In Interval Neural Networks, neurons are connected as they are in traditional Neural Networks. Interval Neural Networks can be classified in the same way Punctual Neural Networks are.

Fig. 1. Interval Neuron structure

3.1 A One Layer Interval Neural Network

The one layer Interval Neural Network is formed by interval neurons where the activation function is the binary threshold function of the form:

$$y = \begin{cases} +1 & \text{If } \underline{x} + \overline{x} > 0, X_i = [\underline{x}, \overline{x}] \in \mathbb{IR}, \\ 0 & \text{in other case.} \end{cases} \tag{7}$$

When neural networks are used as classifiers, the output set is a binary one. Patterns are represented with 0 or 1. This neural networks is used as classifier then its outputs are binaries too. The output is transformed from $\mathbb{IR} \to \mathbb{R}$ by the activation function, thus, the interval space is divided by the pattern separation and each interval belongs to the pattern where the majority of it belong.
Neurons of the one layer interval neural network are full connected with inputs, each of these connections are represented by an interval that represent at the same time the force between relationships.

Learning in the One Layer Interval Neural Network The one layer interval neural network has to be trained in order to extract knowledge of a data set. Each neuron compute and send it output, that is going to said if the neuron is active or not. With the concepts given in section 2 the outputs are obtained and the weights are modified in order to correct a possible error in the output. In the table 1 the algorithm for training the one layer interval neural network is presented.

4 Experiment Results

In order to analyse, the influence of intervals, let us consider a simple example, the classification of two classes. One punctual neural network with one neuron was trained. Its training set is denoted by P_n where n is a number of inputs for each pattern, the interval data set is denoted by IP_n, $X_i \in IP_n$ and $x_i \in P$ where $i \leq n$, $x_i \in X_i$ and X_i is not a symmetrical interval. IP is the training set for the interval neural network that is the same of the punctual neural network. The neural network was trained one hundred times, the mean of all solutions is included in the mean of all solutions obtained with the interval neural network, in the same number of training. If the set IP_n has symmetrical interval, the mean of all solutions of the punctual network is included in the mean of all solutions of the interval network. The rect formed by the middle points of the weights in the interval network defines the separation of the classes. It is coincident with the rect formed for the weights of the punctual network, because of it, the interval neural network can be considered as a generalization of the punctual neural network.
The number of epochs necessary for training this interval neural network is smaller than the number necessary for punctual neural network without loss

Table 1. Learning algorithm for the one layer interval neural network

Step 1	Initialize inputs weights and bias weigth Initialize inputs weights (W_{ij}) and bias weigth θ_i with small random intervals where $size(W_{ij})$ and $size(\theta_i)$ are minimum.
Step 2	Present an input pattern $[X, d]$ to the network X_i is a set of input intervals that represent the pattern and d_i represent a binary vector that codifies the class of the pattern X_i. The vector $X_i = ([\underline{x}_0, \overline{x}_0], [\underline{x}_1, \overline{x}_1], ..., [\underline{x}_{n-1}, \overline{x}_{n-1}])$.
Step 3	Calculate the ouput of the network for the input X_i $y(t) = f_n(\sum_{i=0}^{N-1} W_{ij}(t)X_i(t) - \theta_j)$ f_n is the interval activation function, the calculus is made by the equation 7
Step 4	Weights modification $W_{ij}(t+1) = W_{ij}(t) + \eta[d(t) - y(t)]X_i$, where $0 \leq i \leq N-1$ d(t)= $\begin{cases} +1 & \text{If the input is of class A,} \\ -1 & \text{If the input is of class B.} \end{cases}$ η is a real value less than 1 and $d(t)$ is the desired output for the actual pattern. If the network classified it correctly, then the weigths are not going to be modified.
Step 5	the total error in the neural network for the pattern i is calculated by $e = \sum_{k}^{m} = 1[d_i - y_k(t)]$, if $i = n$ (number of pattern) and $\sum_{i}^{n} e \leq \alpha$ where α represent the minimum error, the algorithm finishes in other case go to Step 2

efficiency. The interval neural networks can be trainedg with the set P_n and the results was successfully, of this form, again can be said that the IANN is a generalization of the punctual neural networks. The table 2 shows the results of this experience, there are showed results of training the interval neural network and the punctual neural network.

Table 2. IANN vs. Punctual Neural Network with η variable and $\epsilon = 0$.

		$\alpha =,2$	$\alpha =,4$	$\alpha =,8$
Intervalar.	Number of Epoch	$15,34$	$6,96$	$14,10$
	Percentage of Efficiency	100%	100%	100%
Pontual.	Number of Epoch	$14,9$	$15,06$	$14,41$
	Percentage of Efficiency	92%	89%	91%

The rate of learning (η) and the minimum error are the important parameters in the process of training, for this reason, they are very important for evaluate the behavior of interval neural network. The minimum error of the IANN can be very small and the neural can converge faster than the punctual neural with same rate of learning, and the efficient of the interval neural network doesnt lower. The table 3 shows the results of training both networks with different error rate and with $\eta = 0,2$. A new test was did with interval neural network.

Table 3. IANN vs. Punctual Neural Network with ϵ variable.

		$\epsilon = 0$	$\epsilon = 1$	$\epsilon = 2$
Intervalar.	Number of Epoch	$15,34$	$6,61$	$6,70$
	Percentage of efficiency	100%	100%	96%
Pontual.	Number of Epoch	$14,98$	$12,14$	$10,11$
	Percentage of efficiency	92%	91%	85%

for four random real synthetic data set, $P_n^1, P_m^2, P_k^3, P_l^4$ and interval asymmetric data set denoted by $IP_n^1, IP_m^2, IP_k^3, IP_l^4$. The networks were trained and tested. The table 4 shows the results of the training.

Table 4. RNAI vs. Punctual Neural Network with some training sets.

		P_n^1 e IP_n^1	P_m^2 e IP_m^2	P_k^3 e IP_k^3	P_l^4 e IP_l^4
Intervalar.	N. Epoch	$19,47$	$45,29$	$9,64$	$52,24$
	% Efficiency	100%			100%
Pontual.	Nm. de pocas	$173,29$	$46,89$	$11,12$	$54,11$
	% Efficiency	100%	%	%	100 %

With these experiences, can be concluded that the inclusion of the interval mathematic in neural networks has many advantages in addition to the control of computational errors. The interval neural network was evaluated with several data sets, varying training parameters and the number of epochs is smaller than the punctual network, the efficiency was good and some times it was better. It can be very important when dealing with real problems. The results of this investigation are very encouraging for applying interval mathematic to other models of the neural networks.

5 Conclusions

Other works was developed in this area [11] [4] [2] but they are solutions to problems about the interval data like input data for the neural network. This paper propose the inclusion of mathematic interval in the structur of the neural network and shows that the performance of the neural network with interval mathematic is better. It proposes a modification in the neurons but the structure of the network wasnt modified with this form, the network is very simple and the process is natural.
In conclusion, this paper proposes a new model of neural network that generalize the traditional neural network and has new and better characteristics, the process of interval data. This approach guarantee an exact and efficient computation process and find a satisfactory and better solution.

References

1. Alefeld, G.; Herzberger, J.: Inroduction to interval computations. Academic Press, New York, (1983)
2. Beheshti, M.; Berrached, A.; Korvin, A. D.; Hu, C.; Sirisaengtaksin, O.: On Interval Weighted Three-Layer Neural Networks. In The 31st Annual Simulation Symposium, p188-195 (1998).
3. Bock, H. H.; Diday, E. Analysis of Simbolic Data. Exploratory methods for extracting statistical information from complex data. Springer verlag, (2000).
4. Ishibuchi., H.; Nii, M.: Interval-Arithmetic-BAsed Neural Networks. In: of Bern, H. B. U.; of South Florida, A. K. U., (Ed.), Hybrid Methods in Pattern Recognition, 1, (2001).
5. Kearfott, R. B.; Kreinorvich V.: Applications of Interval Computations. Kluwet Academi Publisher, (1996).
6. Macdo Torres Silveira M. M. Teoria Fuzzy Intervalar: Uma proposta de integrao da matematica Intervalar Teoria Fuzzy, (2002). Master's Thesis - Universidade Federal do Rio Grande do Norte.
7. Medeiros dos Santos, J. M. Em direo a uma representao para equaes algbricas: Uma lgica equacional local. Master's thesis (2001), Universidade Federal do Rio Grande do Norte.
8. McCulloch, W. S.; Pitts, W.: A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, v.5, p.115–133, (1943).
9. Moore, R. E.: Interval Analysis. New Jersey, (1966).

10. Moore, R. E.: In: SIAM (Ed.). Methods and Applications of Interval Analysis. Philadelphia, (1979). p.190.
11. Rossi, F.; Conan-Guez, B.: Multilayer Perceptron on Interval Data. In: , (2002). Classification, Clustering, and Data Analysis (IFCS 2002); resumos. Cracow, Poland: Springer, 2002. 427-434, http://apiacoa.org/publications/2002/ifcs02.pdf.
12. Voschinin, A. P.; Dyvak N. P. ; Simoff S. J. Interval Methods: Theory and Application in the Design of Experiments, Data Analysis nd Fittin.In:Letzky E. K., (Ed.), Design of Experiments and Data Analysis: New Trends and Resultss, Antal Publishing Co., Moscow, (1993).