POZNAN UNIVERSITY OF TECHNOLOGY ACADEMIC JOURNALS
No 80 Electrical Engineering 2014

Stanisław PŁACZEK*

SELECTED PROBLEM OF STRUCTURE OPTIMIZATION FOR ARTIFICIAL NEURAL NETWORKS WITH FORWARD CONNECTIONS

The problem of Artificial Neural Network (ANN) structure optimization related to the definition of optimal number of hidden layers and distribution of neurons between layers depending on selected optimization criterion and inflicted constrains. The article presents the resolution of the optimization problem. The function describing the number of subspaces is given, and the minimum number of layers as well as the distribution of neurons between layers shall be found.

KEYWORDS: Artificial Neural Network, network structure, structure optimization

1. NETWORK STRUCTURE

Artificial Neural Network (ANN) is implemented as universal approximator function with multidimensional variables. The function can be displayed as:

$$Y = F(X) \tag{1}$$

where: X – input vector, Y – output vector.

Selection of neural network structure, aimed at resolution of a specific problem is a challenging task. It is necessary to consider the following issues:

- Structure of Neural Network, including the number of hidden layers and distribution of neurons between layers. Usually, the size of input and output layer is defined by, respectively, dimension of vectors X and Y.
- Structure of activation function, considering requirements of learning algorithm.
- Methods of data transfer between layers.
- Optimization criteria and type of learning algorithm.

The most popular artificial neural network structure is the network with Direct Connection. This structure consist of at least one hidden layer. Data is fed from proceeding layer to the succeeding one.

The following paper consists of an analysis of the Cross Forward Connection, Fig. 1. In this structure, the input signal is passed on to each layer in the network. Therefore, a layer j = 1, 2, 3….W, where W is the output layer, has two inputs:

* Akademia Finansów i Biznesu Vistula w Warszawie.

- Vector X, dimension N_0.
- Vector V_{j-1} output of proceeding layer, dimension N_i.

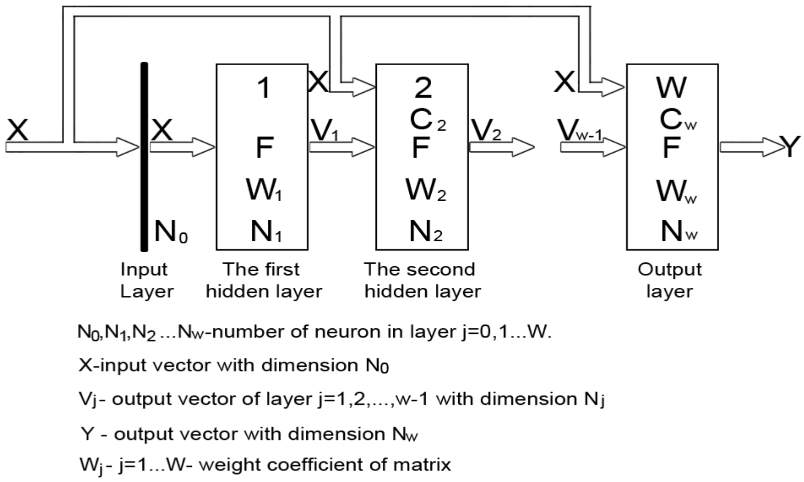

Fig. 1. Forward connection ANN structure

2. ANALYSIS OF ANN STRUCTURE

Describing ANN structure as ANN (2- 3 -1) it could be stated that the network structure includes:
$N_0 = 2$; number of neurons in input layers,
$N_1 = 3$; number of neurons in hidden layers,
$N_2 = 1$; number of neurons in output layer.

For analysis of ANN structure, division of initial future space X of dimensionality N_0 by each network's hidden layer could be applied. Every hidden layer includes Ni neurons, where: i- 1,2,….W-1, number of hidden layers.

In accordance with [2], [8], N_i neutrons in the hidden layer divide initial future space to $\phi(N_o,N_i)$ subspaces of dimensionality N_0-1.

$$\phi(N_0, N_i) = C_{N_i-1}^{N_0} + 2\sum_{k+0}^{N_0-1} C_{N_i-1}^{k} \tag{2}$$

where

$$C_n^k = \frac{n!}{k!(n-k)!} \tag{3}$$

Number of subspaces formed by division of N_0 dimensional input vector X by N_i neurons present in hidden layer Table 1.

Table. 1. Number of subspace $\phi(N_o,N_i)$

N_i / N_0	1	2	3	4	5	6	7	8	9	10
1	2	2	2	2	2	2	2	2	2	2
2	3	4	4	4	4	4	4	4	4	4
3	4	7	8	8	8	8	8	8	8	8
4	5	11	15	16	16	16	16	16	16	16
5	6	16	26	31	32	32	32	32	32	32
6	7	22	42	57	63	64	64	64	64	64
7	8	29	64	99	120	127	128	128	128	128
8	9	37	93	163	219	247	255	256	256	256
9	10	46	130	256	382	466	502	511	512	512
10	11	56	176	386	638	848	968	1013	1023	1024

Example 1. Two-dimensional input space $N_0 = 2$, is divided into 7 subspaces by $N_i = 3$ neurons in hidden layer Fig. 2.

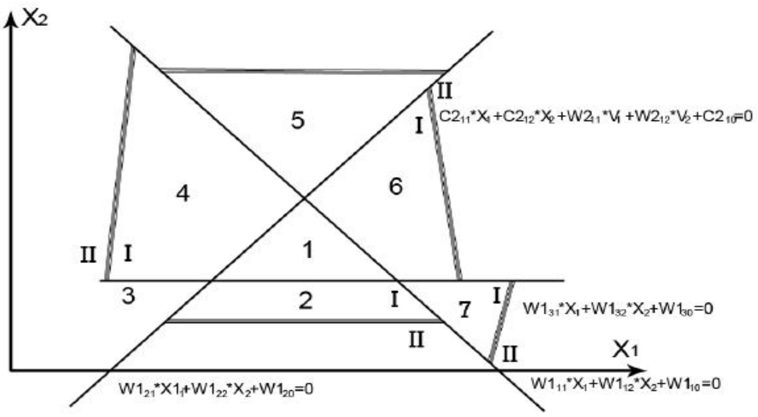

Fig. 2. Division of two-dimensional input space by three neurons

Each subspace formed by the hidden layer is further divided into subspaces by subsequent hidden layer. With number of layers W > 2, total number of subspaces for all ANN could be calculated.

$$\psi(N_0, W) = \prod_{i=1}^{W} \phi(N_0, N_i) \quad (4)$$

Example 2. For ANN (2-3-1), total number of subspaces equals 14.

$$\psi(2,2) = \prod_{i=1}^{W} \phi(2, N_i) = \phi(2,3)\phi(2,1) = 7 \cdot 2 = 14$$

3. ISSUE OF STRUCTURE OPTIMIZATION

Total number of ψ subspaces for Cross Connection Network is given. Minimum number of W layers and N_i neuron distribution between layers shall be found.

$$\min N = \min_{w-1, N_i} \sum_{i=1}^{w-1} N_i \tag{5}$$

and

$$\psi = \prod_{i=1}^{W-1} \phi_i(N_0, N_i) \text{ is given} \tag{6}$$

$$\phi(N_0, N_i) = C_{N_i-1}^{N_0} + 2\sum_{k+0}^{N_0-1} C_{N_i-1}^{k} \tag{7}$$

Our objective is to minimize the total number of neurons and layers. This complex problem could be solved regarding the relation between dimensionality of feature space, N_0, and number of neurons in each of N_i hidden layers. The aforementioned problem could be divided into two sub – problems:

- $N_i > N_0$ – number of neurons in each hidden layer is greater than the number of neurons in input space,
- $N_i \le N_0$ – Dimensionality of input vector is greater than/equal to the number of neurons in each hidden layer,
  where: i = 1,2,3…W-1 or alternative notation i =1,2,3…H where H- is the number of hidden layers.

4. NUMBER OF ψ SUBCPACES IS GIVEN. IT IS NECESSERY TO FIND THE NUMBER OF HIDDEN LAYERS H AND DISTRUBUTION OF NEURONS BETWEEN LAYERS N_i

In this case the target function is defined

$$\min\{N\} = \min_{W-1, N_i} \{\sum_{i=1}^{W-1} N_i\} \tag{8}$$

and

$$\psi = \prod_{i=1}^{w-1} \phi_i(N_0, N_i) \tag{9}$$

is given constrains

$$\phi(N_0, N_i) = C_{N_i-1}^{N_0} + 2\sum_{k+0}^{N_0-1} C_{N_i-1}^{k} \tag{10}$$

and constrains

$$N_i > N_0 \qquad \text{for all} \qquad I = 1,2,\ldots W\text{-}1 \ \text{layers} \tag{11}$$

For Kuhn – Tucker condition needs, constrains (11) could be written

$$N_i \geq N_0 + 1 \tag{12}$$

To find solution, Kuhn – Target condition could be applied. Taking into account (8), (9), (12) Lagrange equation is written.

$$L = \sum_{i=1}^{H} N_i + \lambda_0 \left[\psi - \prod_{i=1}^{H} \phi_i(N_0, N_i) \right] + \sum_{i=1}^{H} \lambda_i (N_i - N_0 - 1) \tag{13}$$

where: H = W - 1. Set of equations could be written

$$\frac{\partial L}{\partial N_i} = 1 - \lambda_0 \frac{\partial \phi}{\partial N_i} \prod_{i=1}^{H} \phi(N_0, N_i) + \lambda_i = 0 \qquad \text{for } i = 1,2,\ldots H \tag{14}$$

$$\frac{\partial L}{\partial \lambda_0} = \psi - \prod_{i=1}^{H} \phi(N_0, N_i) = 0 \tag{15}$$

$$\frac{\partial L}{\partial \lambda_i} = N_i, - N_0 - 1 = 0 \qquad \text{for } i = 1,2,\ldots H \tag{16}$$

From (16) N_i could be found

$$N_i = N_0 + 1 \qquad \text{for } i = 1,2,\ldots H \tag{17}$$

Using (14) and (15) formula (14) could be rewritten

$$\frac{\partial L}{\partial N_i} = 1 - \lambda_0 \frac{\partial \phi}{\partial N_i} \psi + \lambda_i = 0 \qquad \text{for } i = 1,2,\ldots H \tag{18}$$

and finally

$$\lambda_i = \lambda_0 \frac{\partial \phi}{\partial N_i} \psi - 1 \qquad \text{for } i = 1,2,\ldots H \tag{19}$$

The aforementioned means that

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_H \tag{20}$$

Using (18) and (15) formula (15) could be rewritten

$$\psi = \phi(N_0, N_0 + 1)^H \tag{21}$$

And finally

$$H = \frac{\ln \psi}{\ln[\phi(N_0, N_0 + 1)]} \tag{22}$$

Using (18) and (8) the minimum sum of neurons distribution

$$N = H(N_0 + 1) \tag{23}$$

Example 3. Number of subspaces $\psi = 100$ and feature space dimensionality is $N_0 = 2$. Find number of hidden layers H and total number of neurons accomplishing the number of ψ subspaces.

From (17) $\qquad N_i = N_0 + 1 = 2 + 1 = 3$

From (23) calculate

$$H = \frac{\ln \psi}{\ln[\phi(N_0, N_{0+1})]} = \frac{\ln 100}{\ln[\phi(2,3)]} = 2.4 \approx 3$$

From (23) calculate $N = 3 \cdot 3 = 9.$

5. CONCLUTION

The article describes Artificial Neural Network structure and structure optimization based on selected criteria. Cross Connection ANN allows not only to analyze ANN structure but also to synthesize the structure using target functions and constrains. Significant role is played by the dimensionality of input vector X. This vector creates initial feature space for ANN allowing all hidden layers and neurons in those layers to divide input space into subspaces. Those parameters characterize ANN with respect to learning ability and resolution of the analyzed problem. Each problem could be solved using different ANN structures, i.e. various number of hidden layers and different distribution of neurons between layers. Unfortunately, these parameters can potentially have an influence on complexity and time consumption of the learning algorithm. Optimization of parameters is related to minimization of number of layers and total neuron number in ANN structure. Therefore, this operation is vital and conduces to acceleration of processes, including the learning process.

The article gives special attention to minimization of the aforementioned parameters, taking into account dimensionality of vector X as well as relations between dimensionality of neurons in both, hidden layers N_i and initial feature space N_0.

Received ANN structures are distinguished by particular symmetry in the structure of hidden layers. Hidden layers should have the same number of neurons, connected with the dimensionality of vector X. The aforementioned affects the ANN structure; ANN can have a greater number of hidden layers simultaneously fulfilling the criterion of number of ψ subspaces.

Continuing the subject, it is necessary to connect characteristics of given vectors X and Y with requirements regarding dimensionality of the ψ subspace.

REFERENCES

[1] S. Osowski, Sieci neuronowe do przetwarzania informacji, Oficyna Wydawnicza Politechniki Warszawskiej, Warszaw 2006.
[2] O. B. Lapunow, On possibility of circuit synthesis of diverse elements, Mathematical Institut of B.A. Steklova, 1958.
[3] Toshinori Munakate, Foundational of the New Artificial Intelligence, Second Edition, Springer 2008.
[4] Colion Fyle, Artificial Neural Network and Information Theory, Department of Computing and Information System, The university of Paisley, 2000.

[5] Joarder Kamruzzaman, Rezaul Begg, Artificial Neural Network in Finance and Manufacturing, Idea Group Publishing, 2006.
[6] A.Marciniak, J.Korbicz, J.Kus, Wstępne przetwarzanie danych, Sieci Neuronowe tom 6. Akademicka Oficyna Wydawnicza EXIT 2000.
[7] L. Rutkowski, metody i techniki sztucznej inteligencji, Wydawnictwo naukowe PWN, Warszawa 2006.
[8] S. Placzek, B. Adhikari, Analysis of Multilayer Neural Network with Direct Connection Cross-forward Connection, Conference CS&P 2013 Warsaw University, Warszawa 2013.