Third Int. Workshop on GNs, Sofia, 1 Oct. 2002, 32-36

GENERALIZED NET MODEL OF BACKPROPAGATION LEARNING ALGORITHM

Maciej Krawczak[1] and Hristo Aladjov[2]
[1] Systems Research Institute – Polish Academy of Sciences
Wyzsza Szkola Informatyki Stosowanej i Zarzadzania
Warsaw, Poland,
e-mail: *krawczak@ibspan.waw.pl*
[2] CLBME - Bulg. Academy of Sciences, Acad. G. Bonchev Str., 105 Block, Sofia-1113, Bulgaria,
e-mail: *aladjov@clbme.bas.bg*

Abstract: The backpropagation algorithm and its modifications are frequently used approaches for feed-forward neural networks training. It is shown in the present paper how generalized net model [1,2] of pure backpropagation algorithm can be constructed. The basic idea of this model is to be used as frame in which, with minor changes different algorithm definitions can be introduced, tested and compared.

Keywords: Generalized Nets, Neural Networks, Machine Learning, Backpropagation

§1. Introduction

The idea for describing different areas of artificial intelligence by means of common formalism dates back form early nineteens [3]. In [4] an universal generalized net model of neural network was defined. Unlike that general model here we shall focus our attention on training algorithm implementation.

§2. A generalized net model of backpropagation algorithm

The separate neurons of the neural net are represented by α tokens. They enter the net trough place l_1 with the initial characteristic

$$\langle \text{NNI,LI,SI,f} \rangle,$$

where:
NNI is Neural Network Identifier,
LI is identifier of the neuron layer in the network,
SI is Self Identifier of the Neuron,
f = f(net,λ)is a neuron activation function;

$$Z_1 =< \{l_1, l_2\}, \{l_2, l_3\}, \begin{array}{c|cc} & l_2 & l_3 \\ \hline l_1 & W_{1,2} & W_{1,3} \\ l_2 & W_{2,2} & W_{2,3} \end{array}, \vee(l_1, l_2) >,$$

where
$W_{1,2} = \neg W_{1,3} =$ "if there is only one α_i token with given LI", i.e.,

$$\forall(\alpha_j \in K_{l_1})(pr_2 X_{\alpha_i^0} \neq pr_2 X_{\alpha_j^0}, j \neq i)$$

where K_{l_p} is the subset of tokens that are currently in place l_p
$W_{2,2} =$ "if there are more than one token α with one and the same LI", i.e.,

$$\exists(\alpha_i \in K_{l_1} \& \alpha_j \in K_{l_2})(pr_2 X_{\alpha_i^0} = pr_2 X_{\alpha_j^0})$$

K_{l_p} is the subset of tokens that are currently in place $l_p(p = 1, 2, ..., |K_{l_p}|)$,
$W_{2,3} =$ "if all α tokens form a given layer is combined into one token LI" i.e.,

$$\neg\exists(\alpha_i \in K_{l_1} \& \alpha_k \in K_{l_2})(pr_2 X_{\alpha_i^0} = pr_2 X_{\alpha_k^0})$$

$$\& \neg\exists(\alpha_i \in K_{l_1} \& \alpha_j \in K_{l_1})(pr_2 X_{\alpha_i^0} = pr_2 X_{\alpha_j^0}, i \neq j)$$

When enter in place l_2 all α tokens form one and the same layer unite into one single token. Token α obtains the characteristic

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n\rangle, \langle f_1, f_2, ..., f_n\rangle\rangle$$

in place l_2 and l_3, where n is the number of neurons in layer LI.

Token β with the initial characteristic "performance index E of the neural net" and its threshold value E_{max}, i.e.,

$$\langle NNI, E, E_{max}\rangle$$

enters place m_1, where NNI again is Neural Network Identifier.

$$Z_2 = < \{l_3, m_1\}, \{l_4, m_2\}, \begin{array}{c|cc} & l_4 & m_2 \\ \hline l_3 & true & false \\ m_1 & false & true \end{array}, \wedge(l_3, m_1) > .$$

Token α obtains as new characteristic in place l_4

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n\rangle, \langle f_1, f_2, ..., f_n\rangle, W\rangle$$

where W is the is the initialized matrix of weights between neurons of layer LI and layers $LI - 1$ and $LI + 1$. The β token obtains characteristic

$$\langle NNI, 0, E_{max}\rangle,$$

in place m_2.
Tokens γ_p enter place n_1 with initial characteristic

$$\langle X_p(0), D_p, p\rangle,$$

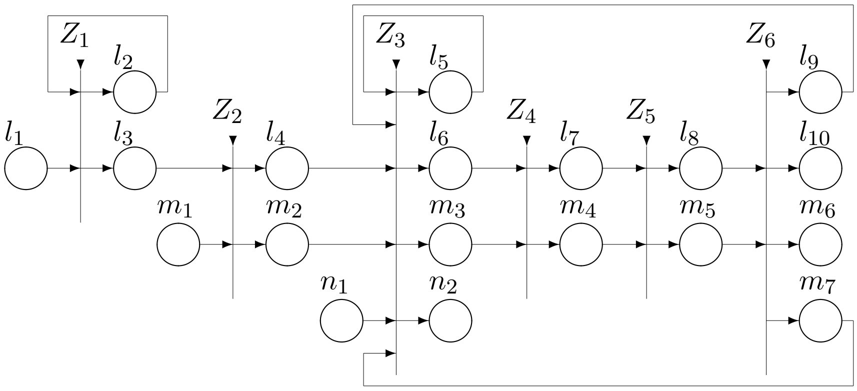
where $X_p(0)$ is the inputs vector of the neural network and D_p is the vector of desired network outputs,and $p, p = 1, 2, ..., P$ is the number of training patterns.

Next transition describes the process of signal propagation within the neural network. After the pattern p is applied to the network inputs $X_p(0)$ the outputs of all layers are calculated subsequently layer by layer.

$$Z_3 =< \{l_4, l_5, l_9, m_2, m_7, n_1\}, \{l_5, l_6, m_3, n_2\}, \begin{array}{c|cccc} & l_5 & l_6 & m_3 & n_2 \\ \hline l_4 & W_{4,5} & W_{4,6} & false & false \\ l_5 & W_{5,5} & W_{5,6} & false & false \\ l_9 & W_{9,5} & W_{9,6} & false & false \\ m_2 & false & false & true & false \\ m_7 & false & false & true & false \\ n_1 & false & false & false & W_{1,2} \end{array},$$

$$\wedge(\vee(l_4, l_5, l_9), \vee(m_2, m_7), n_1) >,$$

where
$W_{4,5} = W_{5,5} = W_{9,5} =$ "Previous layer do not have defined outputs"
$W_{4,6} = W_{5,6} = W_{9,6} = \neg W_{4,5}$.
$W_{1,2} =$ "All layers outputs have assigned values for the current pattern".



All tokens from α-type obtain the characteristic

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n\rangle, \langle f_1, f_2, ..., f_n\rangle, W, X)\rangle$$

where X is the vector of outputs of neurons of LI-th layer in place l_5 they obtain no new characteristic and in place l_6,characteristic with calculated neuron outputs of the layer

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n\rangle, \langle f_1, f_2, ..., f_n\rangle, W, X')\rangle$$

where X' are the new values of outputs.

β-token obtain no new characteristic in place m_3 and transferred γ token preserve its characteristic

$$\langle X_p(0), D_p, p\rangle,$$

in place n_1, where

$$X_p = \{x_{p,1}, x_{p,2}, ..., x_{p,N(0)}\} and D_p = \{d_{p,1}, d_{p,2}, ..., d_{p,N(0)}\}.$$

Transition

$$Z_4 =< \{l_6, m_3\}, \{l_7, m_4\}, \begin{array}{c|cc} & l_7 & m_4 \\ \hline l_6 & true & false \\ m_3 & false & true \end{array}, \wedge(l_3, m_1) >$$

describes the first stage of estimation and weight adjustment process related to the performance index computation. As a result of this computation the β token obtain the new value of performance index estimation

$$\langle NNI, E', E_{max} \rangle$$

in place m_3, where

$$E' = pr_2 X^{beta}_{cur-1} + \frac{1}{2} \sum_{j=1}^{N(L)} (d_{p,j} - x_{p,j(L)})^2$$

. α tokens obtain no new characteristic in place l_7.

In the next transition

$$Z_5 =< \{l_7, m_4\}, \{l_8, m_5\}, \begin{array}{c|cc} & l_8 & m_5 \\ \hline l_7 & true & false \\ m_4 & false & true \end{array}, \wedge(l_7, m_3) >$$

delta factors between desired and obtained network outputs for each layer LI are computed.

Tokens α obtain the characteristic

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n \rangle, \langle f_1, f_2, ..., f_n \rangle, W, X, \Delta_p \rangle$$

in place l_8 and token β obtains no new characteristic in place m_5.

Transition Z_6 describes the process of weight adjustments. It has the form

$$Z_6 =< \{l_8, m_5\}, \{l_9, l_{10}, m_6, m_7\}, \begin{array}{c|cccc} & l_9 & l_{10} & m_6 & m_7 \\ \hline l_8 & W_{8,9} & W_{8,10} & false & false \\ m_5 & false & false & V_{5,6} & V_{5,7} \end{array},$$

$$\wedge(l_8, m_5) >,$$

where
$W_{8,9} =$ "there are more unapplied patterns"
$W_{8,10} = \neg W_{8,9}$,
$V_{5,6} =$ "if the perforans index is below given threshold E_{max}"
$V_{5,7} = \neg W_{5,6}$.

Token α obtains the characteristic with updated neuron connection weights W'

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n \rangle, \langle f_1, f_2, ..., f_n \rangle, W' \rangle$$

in place l_9 and final connection weight W:

$$\langle NNI, LI, \langle SI_1, SI_2, ..., SI_n \rangle, \langle f_1, f_2, ..., f_n \rangle, W \rangle$$

in place l_{10}. In place m_6 and place m_7 β token obtains characteristic

$$\langle NNI, E, E_{max} \rangle$$

describing achieved performance index E.

Conclusion:

In the present paper is described generalized net model of backpropagaton algorithm for neural network training. This algorithm can be easily modified and tested by changing relatively small portion of the generalized net formal description.

References:

[1] K. Atanassov, Generalized nets, World Scientific, Singapore, 1991.

[2] Atanassov K., Generalized index matrices, Comptes rendus de l'Academie Bulgare des Sciences, vol.40, 1987, No.11, 15-18.

[3] Atanassov, K., 1998. Generalized Nets in Artificial Intelligence. Vol. 1: Generalized nets and Expert Systems. "Prof. M. Drinov" Academic Publishing House, Sofia.

[4]. Atanassov K., Aladjov, H., 2000. Generalized nets in artificial intelligence. Vol. 2, "Prof. Marin Drinov" Academic Publishing House, Sofia.