LEARNING THE NEURAL NETWORKS BY THE SET OF PATTERNS HAVING THE FORM OF FUZZY DATA

Vladimír MIKULA
Institute of Radioelectronics
Technical University Brno
Antonínská 1, 662 09 Brno
Czech Republic

Abstract

This paper deals with the possibility of learning the neural networks by the use of training patterns having the form of both the crisp numerical data as well as fuzzy numbers.

Keywords:

neural networks,training patterns,crisp data, fuzzy numbers, fuzzy logic, fuzzy IF-THEN rules, classifiers

1. Introduction

Neural networks are obviously used for pattern recognition and their classification into defined classes. One of the main properties of neural networks is their ability to learn by the sets of training patterns, having obviously the character of crisp numerical data. However, more general method of learning the neural networks,studied in recent time, enables to use at the input of neural networks both the crisp numerical data, as well as fuzzy numbers, that can be obtained as the expert knowledge, expressed by the set of rules of the type IF (antecedent) THEN (consequent), where one operates with linguistic values, defined by the dependence of the membership degree upon the real variable.This interesting idea was introduced by ISCHIBUCHI, FUJIOKA and TANAKA in [3] and it was studied also in our institute in the frame of the grant No 102/93/1266 (supervised by Prof. J. Pospíšil), supported by the Grant Agency of the Czech Rupublic.

2. Learning the Neural Networks using the Rules of Fuzzy Logic

Neural networks used for the classification of patterns can be of various paradigms [1]. Let us assume the often case of neural network - so called multilayer perceptron (MLP) with one input (fan-out) layer of neurons, one or several hidden layers of performing (computing) neurons and one layer of output neurons, their number c depends upon the number of output classes (Fig.1). Such a network operates in feed-forward mode of operation (without the feed-back), having two phases, namely the phase of learning (or training) the network by the set of training patterns, and the phase of the use of trained-up network for recognizing or classification of unknown input patterns into the different classes.

Let such a network is under the influence of n-dimensional input patterns, generally represented by fuzzy vectors X, for which it is valid the proposition

$$X = (X_1, X_2, \dots X_n) \text{ belongs to class } C_k, \qquad (1)$$

where $k = 1,2,\dots,c$ is the index of the class of classification.

In general case the input fuzzy vector contains the crisp numerical data as well as fuzzy data obtained by the expert method as a set of rules of the type

IF (x_1 is X_1 AND x_2 is X_2 AND...AND x_n is X_n)

THEN ($x = (x_1, x_2, \dots, x_n)$ belongs into the class C_k). (2)

Here $x_1, x_2, \dots, x_n$ denotes the input variables,

$X_1, X_2, \dots, X_n$ are the input fuzzy sets,that can have the character of either fuzzy numbers or sometimes the character of linquistic values expressed by the labels, indicating the properties as "low","medium","large", etc.We can assume them as fuzzy numbers, having for example triangular form of the function of membership degree [2] (Fig.2). Crisp values can be supposed as a special kind of triangular fuzzy numbers with the accurate peak coordinate x with $\mu(x)$=1 and the spread a=0.

For the membership degree of fuzzy vector x it is then valid (with respect to the operation AND in relation(2)):

$$\mu_X(x) = \min\{\mu_{X_1}(x_1), \dots, \mu_{X_n}(x_n)\}, \qquad (3)$$

where $x = (x_1, x_2, \dots, x_n)$ is the vector of input variables.

Such a neural network for the purpose of classifications has to transform the input fuzzy vectors into the output fuzzy numbers. The input-output relations of individual neurons in the feed-forward neural network will be in the form

$$Y_i = X_i, \qquad (4)$$

where $i = 1,2,\dots, n$
for the input (fan-out) layer of neurons,

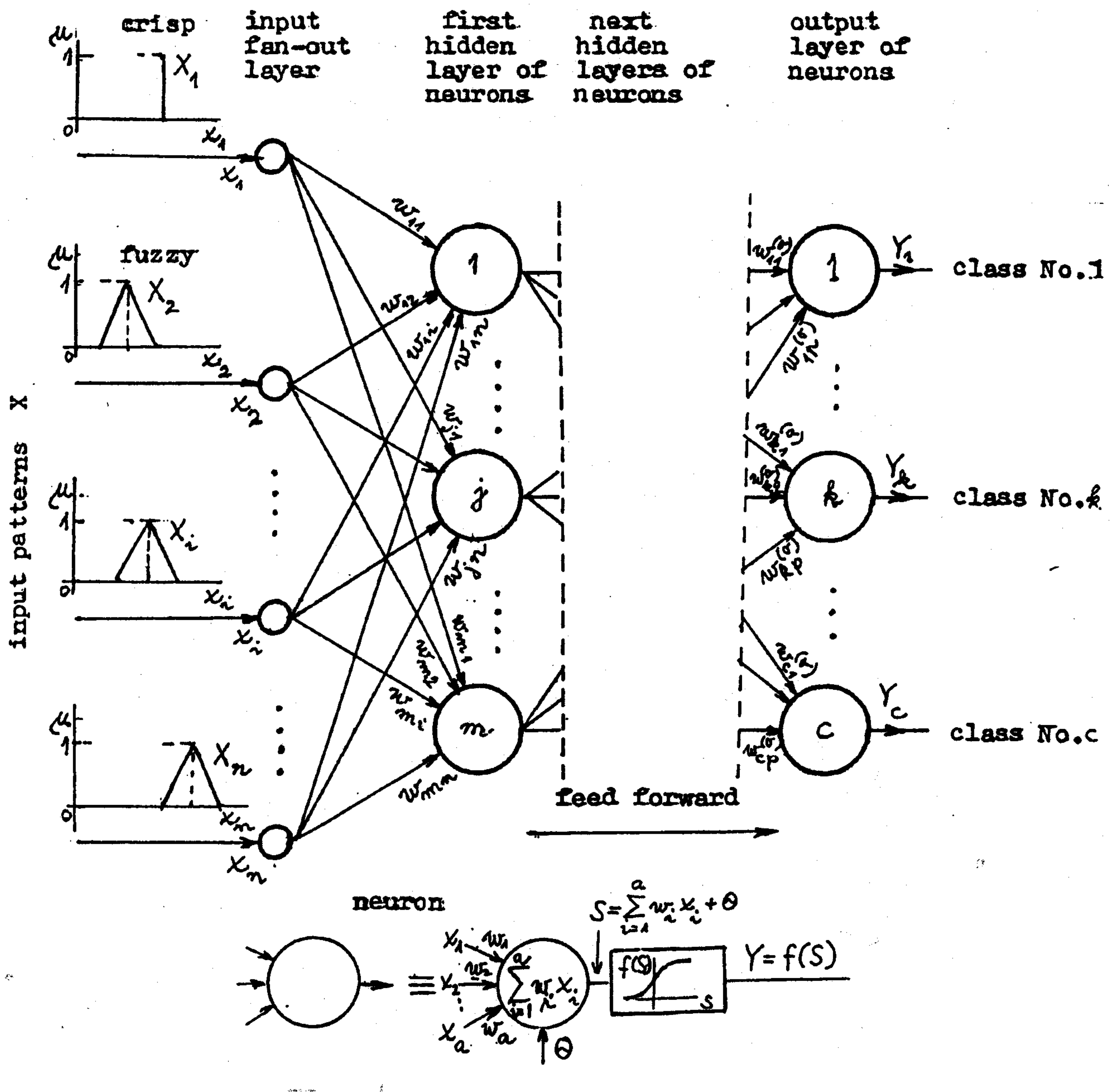

Fig.1 Feed-forward neural network of the type of multilayer perceptron with input patterns having the form of crisp numerical data as well as fuzzy numbers.

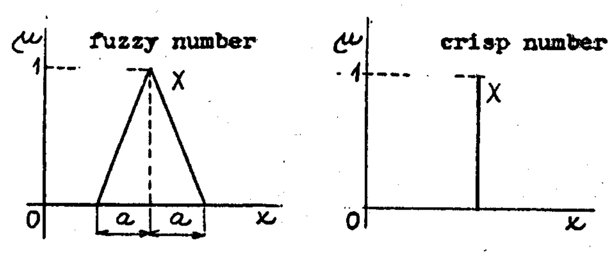

Fig.2 Components of input patterns

$$Y_j = f(S_j) = f(\sum_{i=1}^{n} w_{ji} X_i + \theta_j), \qquad (5)$$

where $j = 1,2,..,m$,

for the 1[st] hidden layer of performing neurons, similarly for next hidden layers, up to the output layer, where

$$Y_k = f(S_k) = f(\sum_{q=1}^{p} w_{kq}^{(o)} Y_q + \theta_k), \qquad (6)$$

where $k = 1,2,..,c$ and $q = 1,2,...,p$ is the denotation of neurons in last hidden layer, (o) means "output". Here $w_{j1}, ..., w_{kq}^{(o)}$ (the synaptic weight coefficients) and $\theta_j, ..., \theta_k$

(threshold values) are real parameters adapted during the trainig phase of the network.

The input quantities X_i, nonlinear transfer functions $f(S)$ and the output quantities are fuzzy numbers. The activation function of performing neurons has obviously sigmoidal form

$$f(S) = 1/(1 + e^{-\lambda S}), \tag{7}$$

where S represents the weighted sum of inputs

$$S_j = \sum_{i=1}^{n} w_{ji} X_i + \theta_j, \tag{8}$$

up to

$$S_k = \sum_{q=1}^{p} w_{kq}^{(o)} Y_q + \theta_k \tag{9}$$

and λ is the parameter determining the slope of the sigmoid in the point of inflection. The computing of actual values of the outputs Y_k needs the performing of the arithmetic sumation of fuzzy numbers. This can be done using the reprezentation of fuzzy sets by their α-level cuts. [2],[4],[5].

For α-cuts of the fuzzy set X it is valid

$$[X]_\alpha = \{x, \text{ at } \mu_X(x) \geq \alpha \text{ and } 0 \leq \alpha \leq 1\}. \tag{10}$$

The fuzzy set X can be expressed as

$$X = \bigcup_{\alpha_i=0}^{1} \alpha_i [X]_{\alpha_i}, \tag{11}$$

where the symbol $\bigcup$ means the "maximum" or disjunction operation in fuzzy logic (Fig.3).

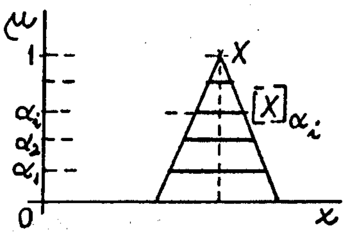

Fig.3 Reprezentation of fuzzy set by its α-cuts

As one can see from Fig.3. the fuzzy numbers are the convex fuzzy sets, their α-cuts are continual intervals. For arithmetic operations. especially for arithmetic sumation of fuzzy numbers, as needed in relations (5) and (6), one can use the interval sumation (+). where it is valid [4]:

$$X_1(+)X_2 = [x_{L1}, x_{U1}] \, (+) \, [x_{L2}, x_{U2}] = [x_{L1} + x_{L2}, x_{U1} + x_{U2}] \tag{12}$$

where L and U denotes the lower and the upper limit,respectively (Fig.4)

For multiplication of fuzzy number by constant value k it is valid

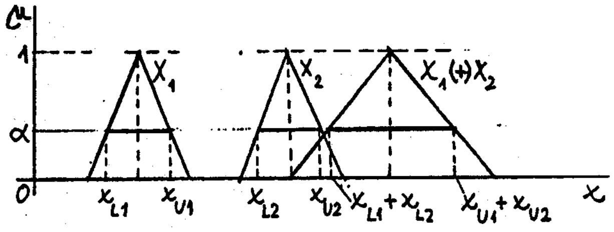

Fig.4 The arithmetic sumation of fuzzy numbers using interval arithmetic

$$kX_1 = k[x_{L1}, x_{U1}] = [kx_{L1}, kx_{U1}] \quad \text{for } k \geq$$
$$\text{or } = [kx_{U1}, kx_{L1}] \quad \text{for } k < 0 \tag{13}$$

Similarly, the activation function $f(S)$ can be found by interval operations

$$f(S) = f[S_L, S_U] = [f(S_L), f(S_U)] \tag{14}$$

It can be seen from (7), that the shape of $f(S)$ will be strongly influenced by the shape of sigmoid.

At the training of the multilayer perceptron with supervision we are using for the correction (adaptation) of the parameters of the network, expecially of the weight coefficients the algorithm based on the back-propagation of the error signal, obtained by comparing the actual output Y_K with the so called desired (target) output Y_{tk}. This target output of the k-th output neuron for certain fuzzy input vector X is defined by the relation

$$Y_{tk} = \begin{cases} 1 \text{ if } X \text{ belongs to the class } C_k \\ 0 \text{ if } X \text{ does not belong to the class } C_k. \end{cases} \tag{15}$$

At the training of the network we want to find the setting up of weight coefficients w, eventually also the biases θ using the BP method, so that the changes of weights will be determined according to the relations of the type

$$\Delta w(t+1) = \eta\left(-\frac{\partial E_\alpha}{\partial w}\right) + \gamma \Delta w(t), \tag{16}$$

where t is a present and $t+1$ is the next turn of presentation of the set of patterns, η is the learning constant ($\eta < 1$) and γ is the momentum constant ($\gamma < 1$) [1].

The symbol E_α denotes the error function of the type

$$E_\alpha = \max\left\{\tfrac{1}{2}(Y_{tk} - Y_k)^2\right\}, \tag{17}$$

where $Y_k \in [Y_k]_\alpha$.

The total error function E, its value indicates the degree (quality) of the training-up of the neural network as the classsifier of input patterns, will be expressed by the relation

$$E = \sum_{\alpha_{min}}^{\alpha_{max}} \alpha E_\alpha = \sum_{\alpha_{min}}^{\alpha_{max}} \sum_{k=1}^{c} \alpha \max\left\{\tfrac{1}{2}(Y_{tk} - Y_k)^2\right\}. \tag{18}$$

3. Conclusion

The neural network is learnt up, when the value of total error function E will decrease to the required satisfactory low level (for example to 0.01). Then it is possible to exploit the network for the classification of input patterns, having the character of both the crisp (numerical) as well as fuzzy numbers.

4. References

[1] KOSKO,B.: Neural Networks and Fuzzy Systems. Prentice-Hall Int., Inc. New York, 1992.

[2] NOVÁK,V.: Fuzzy množiny a jejich aplikace (in Czech) (Fuzzy Sets and their Aplications). Matematický seminář, Praha, SNTL 1990.

[3] ISCHIBUCHI,H.-FUJIOKA,R.-TANAKA,H.: Neural Networks That Learn from Fuzzy If-Then Rules. IEEE Transactions on Fuzzy Systems. 1 (1993) No 2, pp. 85-87.

[4] KAUFMÁNN,A.: Iniciations Élementaires aux Sous-ensambles flous. Tome I. École polytechnique féderale de Lausanne, 1992.

[5] UEHARA,K.-FUJISE,M.: Fuzzy Inference Based on Families of α-level Sets. IEEE Trans. on Fuzzy Systems 1(1993) No.2, pp.111-124.

About author...

Vladimír Mikula was born in Kochanovce, Slovakia in 1929. He received the ME degree in electrical engineering from the Military Technical Academy Brno in 1954, where he also obtained his CSc.(PhD) degree in electronics in 1965. Since 1967 he is with the Institute of Radioelectronics, Faculty of Electrical Engineering and Computer Science, Technical University of Brno, where he was head of Department in 1981-1990. He has been appointed the Professor in Radioelectronics in 1979. His pedagogical and scientific interests are: non-linear circuits, pulse digital techniques and at present time the neural networks and fuzzy systems.

RADIOELECTRONICS ANNUAL MEETING

The 18th Annual Meeting of radioelectronics oriented departments of Czech and Slovak Technical Universities was held in May 24-26, 1995 in Cikháj at the Bohemian-Moravian Highlands. The organizer was the Institute of Radioelectronics, Technical University Brno in co-operation with the Radioengineering Society. Over 40 participants of the following departments took part in the meeting:

Department of Radioelectronics, Technical University Košice
Department of Radioelectronics, Military Academy Liptovský Mikuláš
Department of Radioelectronics, Slovak Technical University Bratislava
Department of Electronics, Technical University Ostrava
Department of Radar, Military Academy Brno
Department of Electronics and Electrical Engineering, Military Academy Brno
Institute of Biomedical Engineering, Technical University Brno
Department of Radioelectronics, Czech Technical University Praha
Department of Electromagnetic Field, Czech Technical University Praha
Department of Circuit Theory, Czech Technical University Praha
Department of Applied Electronics, West Bohemian University Plzeň

The special guests were Prof.Dr. Stanley Novak from Instituto Militar de Engenharia Rio de Janeiro, Brasil and Dr. Petr Moos, dean of the Faculty of Transport, Czech Technical University Praha.

The participants presented results of the pedagogical and research activities of their departments and talked about various problems of their own work. New ideas for mutual co-operations as well as for some teaching and research projects were elaborated. They were very useful discussions between colleagues and friends in the beautiful spring country.

The next 19th annual meeting will be organized by the Department of Electronics, TU Ostrava in the year 1996.

Jiří Svačina, Technical University Brno