Adaptive Model for Diseases Number Prediction Based on Neuro - Fuzzy Technique

ALEXANDER ROTSHTEIN, MORTON POSNER
Dept. of Industrial Engineering and Management
Jerusalem College of Technology – Machon Lev
21 Havaad Haleumi, 91160, Jerusalem
ISRAEL

HANNA RAKYTYANSKA
Dept. of Applied Mathematics
Vinnitsa State Technical University
95 Khmelnitske sh., 21021 Vinnitsa
UKRAINE

Abstract: - A neuro-fuzzy method is proposed for diseases number prediction on the basis of expert regularities that can be revealed in available experimental data. The neuro-fuzzy technology allows to organize the training process in real time and to use all of new experimental data for on-line improvement of the prediction model.

Key-Words: - Fuzzy logic, linguistic approximation, neuro-fuzzy network, training of neuro-fuzzy network

1 Introduction

The prediction of the number of diseases of some type or other is a necessary element of organization of medical-preventive measures. From a formal viewpoint, this problem is related to a wide class of problems of predicting discrete sequences [1] originating not only in medicine but also in engineering, economics, sociology, etc. The nontrivial nature of the prediction of discrete sequences is due to the fact that, in contrast to well-algorithmisized interpolation procedures, the prediction requires the extrapolation of data on the past to data on the future. In this case, it is necessary to take into account an unknown low governing a process generating discrete sequences.

A great number of papers deal with the development of mathematical models of prediction. The methods based on probabilistic-statistical means are most widely used; however, their use requires a considerable amount of experimental data, which are not always available under the conditions of even recent events. Interest has recently been revived [2] on the use of artificial neural networks for the solution of prediction problems. The networks are considered as universal models akin to the human brain, which are trained to recognize unknown regularities. However, a large sample of data is required in the case of training neural networks, as well as in the case of using probabilistic-statistical methods. Moreover, a trained neural network does not permit to explicitly interpret the weights of arcs.

An approach of diseases number prediction based on expert fuzzy IF-THEN rules has been proposed in [3]. The problem of IF-THEN rules training has been solved by off-line optimization technique (gradient or genetic algorithm). This paper proposes neuro-fuzzy model for diseases number prediction which gives the possibility of training fuzzy knowledge bases in real time, i.e. on-line training.

2 Linguistic Model of Prediction

We consider information on the incidence of appendicular peritonitis disease according to the data of the Vinnitsa clinic of children's surgery in 1982-2000 that are presented in Table 1.

Table 1. Distribution of the diseases number

Year	1982	1983	1984	1985	1986	1987	1988	1989
Number of diseases	109	143	161	136	161	163	213	220
Year	1990	1991	1992	1993	1994	1995	1996	1997
Number of diseases	162	194	164	196	245	252	240	225
Year	1998	1999	2000					
Number of diseases	160	185	174					

Analyzing the disease dynamics in Fig.1, it is easy to observe the presence of four-year cycles whose third position occupies a leap year.

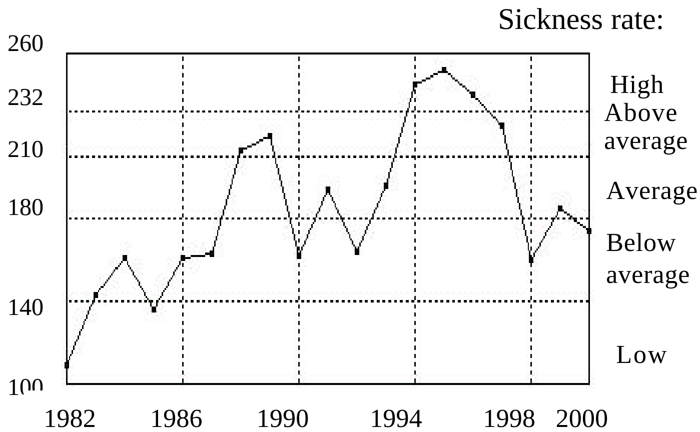


Fig. 1. Disease dynamics

These cycles will be denoted as follows:

$$\ldots\ x_4^{i-1} \} \{ x_1^i \quad x_2^i \quad x_3^i \quad x_4^i \} \{ x_1^{i+1} \ \ldots$$

leap year

where i is the number of a four-year cycle, x_1^i is the number of diseases during two years prior to a leap year, x_2^i is the number of diseases during one year prior to a leap year, x_3^i is the number of diseases during a leap year, and x_4^i is the number of diseases during the year next to a leap year.

The regularities that can be seen in Fig. 1 are easily written in the form of four expert opinions in a natural language. These opinions are IF-THEN rules that relate the sickness rates in the ith and $(i+1)$ th cycles [3]:

$$F_1 = \left|\begin{array}{l} \text{IF } x_1^l = \textit{low} \text{ AND } x_2^l = \textit{below average} \\ \text{THEN } x_3^i = \textit{below average}, \\ \text{IF } x_1^l = \textit{below average} \text{ AND } x_2^l = \textit{below average} \\ \text{THEN } x_3^i = \textit{above average} \\ \text{IF } x_1^l = \textit{below average} \text{ AND } x_2^l = \textit{average} \\ \text{THEN } x_3^i = \textit{below average} \\ \text{IF } x_1^l = \textit{high} \text{ AND } x_2^l = \textit{high} \\ \text{THEN } x_3^i = \textit{high} \end{array}\right.$$

$$F_2 = \left|\begin{array}{l} \text{IF } x_1^l = \textit{low} \text{ AND } x_2^l = \textit{below average}, \\ \text{THEN } x_4^l = \textit{low} \\ \text{IF } x_1^l = \textit{below average} \text{ AND } x_2^l = \textit{below average}, \\ \text{THEN } x_4^l = \textit{above average} \\ \text{IF } x_1^l = \textit{below average} \text{ AND } x_2^l = \textit{average}, \\ \text{THEN } x_4^l = \textit{average}, \\ \text{IF } x_1^l = \textit{high} \text{ AND } x_2^l = \textit{high} , \\ \text{THEN } x_4^l = \textit{above average} \end{array}\right.$$

$$F_3 = \left|\begin{array}{l} \text{IF } x_4^l = \textit{low} \text{ THEN } x_1^{l+1} = \textit{below average} \\ \text{IF } x_4^l = \textit{above average} \\ \text{THEN } x_1^{l+1} = \textit{below average} \\ \text{IF } x_4^l = \textit{average} \text{ THEN } x_1^{l+1} = \textit{high} \end{array}\right.$$

$$F_4 = \left|\begin{array}{l} \text{IF } x_4^l = \textit{low} \text{ AND } x_1^{l+1} = \textit{below average}, \\ \text{THEN } x^{l+1}_2 = \textit{below average} \\ \text{IF } x_4^l = \textit{above average} \text{ AND } x_1^{l+1} = \textit{below average} \\ \text{THEN } x^{l+1}_2 = \textit{average} \\ \text{IF } x_1^l = \textit{average} \text{ AND } x_1^{l+1} = \textit{high} \\ \text{THEN } x^{l+1}_2 = \textit{high} \end{array}\right.$$

The network of relations in Fig. 2, which combines the rules generated above, shows that it is possible, by the first two years of the ith cycle to predict the situation for the next four years: for the last two years of the ith cycle and for the first two years of the succeeding $(i+1)$th cycle.

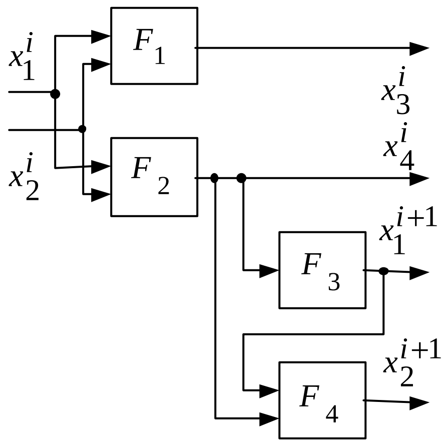


Fig. 2. A network of relations for prediction

To use the expert-linguistic opinions $F_1 \div F_4$, we apply the methods of the theory of fuzzy sets [4]. According to this theory the linguistic estimates “low”, “below average”, and others are formalized with the help of membership functions. We will specify these functions in the following form [5]:

$$m^T(x) = \frac{1}{1+\left(\frac{x-b}{c}\right)^2}$$

where b and c are parameters that are first chosen by an expert and then are adjusted according to experimental data and $m^T(x)$ is the number within the range of [0,1] characterizing a subjective measure of conformity of the value of x to a linguistic estimate T.

The parameters b and c chosen by an expert for various linguistic estimates and used in rules $F_1 \div F_4$, are presented in Table 2. The membership functions obtained in this case are shown in Fig.3

Table 2. Parameters of membership functions before training

Linguistic estimates of variables $x_1^i \div x_4^i$	Parameter	
	b	c
low (L)	100	50
below average (bA)	160	30
average (A)	195	25
above average (aA)	222	20
high (H)	260	30

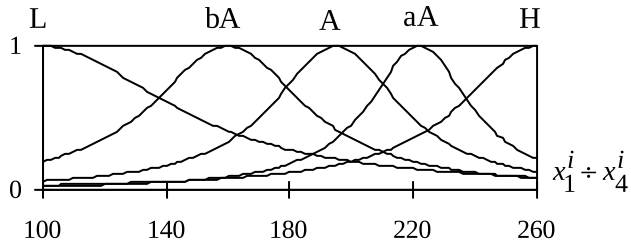
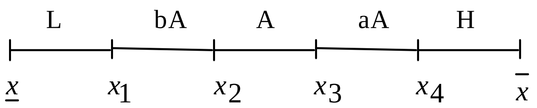


Fig. 3. Membership functions of linguistic estimates before training

In addition to the two-parameter membership functions chosen above, other functions can also be used, for example, triangular or trapezoidal ones [4], containing three and four adjustable parameters.

We denote by $\left[\underline{x}, \overline{x}\right]$ the range of possible values of the diseases number. Let us subdivide this range into the following five parts associated with the linguistic estimations: low(L), below average (bA), average (A), above average (aA), and high (H).



Then, using the fuzzy-logic operations *min (AND) and max (OR),* and the principle of weighted sum for transformation of a membership function into a precise number, we can write a model of prediction in the following explicit form:

$$F_1:\begin{cases} x_3^i = \dfrac{x_1 m^{bA}\left(x_3^i\right) + x_3 m^{aA}\left(x_3^i\right) + x_4 m^{H}\left(x_3^i\right)}{m^{bA}\left(x_3^i\right) + m^{aA}\left(x_3^i\right) + m^{H}\left(x_3^i\right)}, \\ m^{bA}\left(x_3^i\right) = \max\begin{pmatrix} \min\left(m^{l}\left(x_1^i\right), m^{bA}\left(x_2^i\right)\right), \\ \min\left(m^{bA}\left(x_1^i\right), m^{A}\left(x_2^i\right)\right) \end{pmatrix} \\ m^{aA}\left(x_3^i\right) = \min\left(m^{bA}\left(x_1^i\right), m^{bA}\left(x_2^i\right)\right) \\ m^{H}\left(x_3^i\right) = \min\left(m^{H}\left(x_1^i\right), m^{H}\left(x_2^i\right)\right) \end{cases} \qquad (1)$$

$$F_2:\begin{cases} x_4^i = \dfrac{\underline{x}\, m^{L}\left(x_4^i\right) + x_2 m^{A}\left(x_4^i\right) + x_3 m^{aA}\left(x_4^i\right)}{m^{L}\left(x_4^i\right) + m^{A}\left(x_4^i\right) + m^{aA}\left(x_4^i\right)} \\ m^{L}\left(x_4^i\right) = \min\left(m^{L}\left(x_1^i\right), m^{bA}\left(x_2^i\right)\right) \\ m^{A}\left(x_4^i\right) = \min\left(m^{bA}\left(x_1^i\right), m^{A}\left(x_2^i\right)\right) \\ m^{aA}\left(x_4^i\right) = \max\begin{pmatrix} \min\left(m^{bA}\left(x_1^i\right), m^{bA}\left(x_2^i\right)\right), \\ \min\left(m^{H}\left(x_1^i\right), m^{H}\left(x_2^i\right)\right) \end{pmatrix} \end{cases} \qquad (2)$$

$$F_3:\begin{cases} x_1^{i+1} = \dfrac{x_1 m^{bA}\left(x_1^{i+1}\right) + x_4 m^{H}\left(x_1^{i+1}\right)}{m^{bA}\left(x_1^{i+1}\right) + m^{H}\left(x_1^{i+1}\right)} \\ m^{bA}\left(x_1^{i+1}\right) = \max\left(m^{L}\left(x_4^i\right), m^{aA}\left(x_4^i\right)\right) \\ m^{H}\left(x_1^{i+1}\right) = m^{A}\left(x_4^i\right) \end{cases} \qquad (3)$$

$$F_4:\begin{cases} x_2^{i+1} = \dfrac{x_1 m^{bA}\left(x_2^{i+1}\right) + x_2 m^{A}\left(x_2^{i+1}\right) + x_4 m^{H}\left(x_2^{i+1}\right)}{m^{bA}\left(x_2^{i+1}\right) + m^{A}\left(x_2^{i+1}\right) + m^{H}\left(x_2^{i+1}\right)} \\ m^{bA}\left(x_2^{i+1}\right) = \min\left(m^{L}\left(x_4^i\right), m^{bA}\left(x_1^{i+1}\right)\right) \\ m^{A}\left(x_2^{i+1}\right) = \min\left(m^{aA}\left(x_4^i\right), m^{bA}\left(x_1^{i+1}\right)\right) \\ m^{H}\left(x_2^{i+1}\right) = \min\left(m^{A}\left(x_4^i\right), m^{H}\left(x_1^{i+1}\right)\right) \end{cases} \qquad (4)$$

Using the obtained model we can receive rough forecasting of diseases number as shown in Fig. 4.

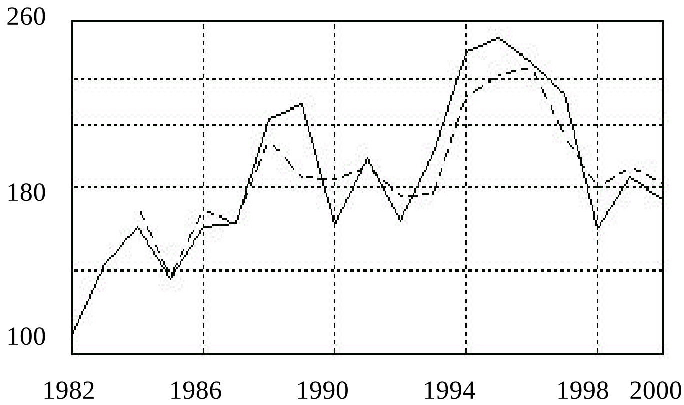


Fig. 4. Comparison of the experimental data (___) and prediction model (-----) before training

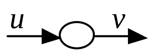
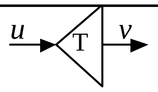
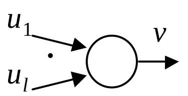
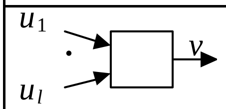
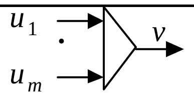
To increase the precision of prediction it is necessary to train the model by available experimental data.

3 Neuro-fuzzy Model of Prediction

For on-line training of above mentioned fuzzy model we use the neuro-fuzzy method of nonlinear function identification proposed in [6]. According to this method a given system of fuzzy IF-THEN rules need to be transform into the neuro-fuzzy network using the elements presented in the table 3. A neuro-fuzzy model of prediction, based on the elements from table3 is presented in Fig.5. As is seen from Fig.5 the neuro-fuzzy network has the following five layers:

(1) – the inputs of the model of prediction;
(2) – fuzzy terms used in knowledge bases $F_1 \div F_4$;
(3) – conjunctions rows of knowledge bases;
(4) – the rules united in the classes $[\underline{x}, x_1, x_2, .., \overline{x}]$;
(5) – defuzzyfication operation.

Table 3. Elements of Neuro-fuzzy Network

Network Node	Node Name	Function
	Input	$v=u$
	Fuzzy Term	$v=m^{T}(u)$
	Fuzzy Rule	$v=\prod_{i=1}^{l} u_i$
	Class of Rules	$v=\sum_{i=1}^{l} u_i$
	Defuzzy-fication (x_j – center of class)	$v=\dfrac{\sum_{j=1}^{5} u_j x_j}{\sum_{j=1}^{m} u_j}$

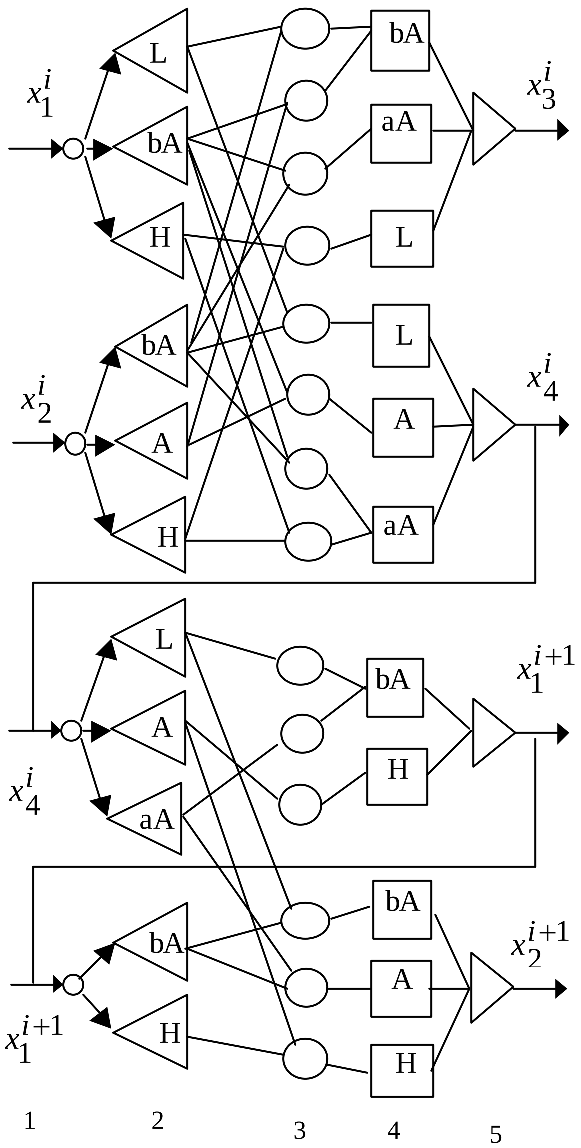


Fig. 5. Neuro-fuzzy model of prediction

The following weights are assigned to the edges of the graph: unity (the edges between the first and second layers, fourth and fifth layers); the membership function specifying the grade of membership of an input in a fuzzy term (the edges between the second and third layers); the weights of rules (the edges between the third and fourth layers).

In defining the elements “fuzzy rule” and “class of rules” in Table3 the operations *min* and *max* from formulas (1-4) are replased by the arithmetic operations of multiplication and addition The possibility of such a replacement is justified in [4]. In this article such a replacement enables us to obtain analytical expressions that are easily differentiated.

4 On-line Training of the Neuro-fuzzy Model of Prediction

The essence of training [6] lies in the selection of the weights of edges that minimize the distinction between the theoretical and experimental number of diseases:

$$\sum_{i=1}^{N}\left(x_3^i-\hat{x}_3^i\right)^2+\sum_{i=1}^{N}\left(x_4^i-\hat{x}_4^i\right)^2+$$
$$+\sum_{i=1}^{N-1}\left(x_1^{i+1}-\hat{x}_1^{i+1}\right)^2+\sum_{i=1}^{N-1}\left(x_2^{i+1}-\hat{x}_2^{i+1}\right)^2=\min_{w,b,c},$$

where $x_3^i, x_4^i, x_1^{i+1}, x_2^{i+1}$ are predicted numbers of diseases depending on the parameters b and c of the membership functions and rules weights;

$\hat{x}_3^i, \hat{x}_4^i, \hat{x}_1^{i+1}, \hat{x}_2^{i+1}$ are experimental numbers of diseases;

N is the number of cycles used to train the model.

The following system of recursive relations is used for on-line training:

$$w_{jk}(t+1)=w_{jk}(t)-\eta\frac{\partial E_t}{\partial w_{jk}(t)}, \quad (5)$$

$$c_{1-4}^{ip}(t+1)=c_{1-4}^{ip}(t)-\eta\frac{\partial E_t}{\partial c_{1-4}^{ip}(t)}, \quad (6)$$

$$b_{1-4}^{ip}(t+1)=b_{1-4}^{ip}(t)-\eta\frac{\partial E_t}{\partial b_{1-4}^{ip}(t)}, \quad (7)$$

minimizing the criterion

$$E_t=\frac{1}{2}(\hat{x}_t-x_t)^2,$$

which is used in the theory of neural networks, where

$\hat{x}_t$ and x_t are theoretical and experimental number of diseases at the tth step of training;

$w_{jk}(t)$ is weight of kth rule, combined diseases numbers in the years n_{jk} in relation F_j, $j = \overline{1,4}$

$c^{ip}_{1-4}(t)$, $b^{ip}_{1-4}(t)$ are parameters of membership function of variable x^i_{1-4} to pth fuzzy term at the tth step of training;

$\boldsymbol{h}$ is parameter of training which can be chosen in accordance with the recommendations of [7].

The algorithm of training neuro-fuzzy networks is similar to the back-propagation rule and consists of two stages. At the first stage, a model-based value of the object output (*x*) corresponding to given network architecture is computed. At the second stage, the value of the discrepancy (E_t) is computed and weights of interneural connections are recalculated. The formulas for computation of partial derivatives in recursive relations (5)-(7) are presented in the Appendix. After training of the prediction model for $N = 4$, which means using the data obtained over the years 1982-1997, the expert-linguistic regularities F_1 -F_4 weights presented in Table 4 are evaluated. It was supposed before training that rules weights were equal to 1. Parameters of membership functions after training are presented in Table 5.

Table 4. Weights of the expert-linguistic regularities

Rules weights in F_1				Rules weights in F_3		
w_{11}	w_{12}	w_{13}	w_{14}	w_{31}	w_{32}	w_{33}
1.000	0.999	0.564	0.885	1.000	1.000	0.668
Rules weights in F_2				Rules weights in F_4		
w_{21}	w_{22}	w_{23}	w_{24}	w_{41}	w_{42}	w_{43}
1.000	1.000	1.000	1.000	1.000	0.992	0.965

Table 5. Parameters of membership functions after training

Linguistic estimates of variables $x^i_1 \div x^i_4$	Parameter	
	b	*c*
low (L)	99.944	8.194
below average (bA)	145.813	19.504
average (A)	194.949	6.999
above average (aA)	234.001	10.636
High (H)	249.134	42.742

As is seen from Tables 2 and 5 after training the neuro-fuzzy network we have the greatest changes in the parameters ñ of membership functions. This can be explained by the fact that in forming the fuzzy knowledge base the expert has specified sufficiently exact positions of the maxima of membership functions (*b*) and weights of the rules (*w*). The choice of large values of the parameters ñ by the expert testifies to a considerable uncertainty in estimating fuzzy terms. A decrease in the values of the parameters ñ in the course of training has resulted in a concentration (compression) of membership functions which testifies to the removal of the uncertainty in estimating fuzzy terms. Membership functions after training are presented in Fig.6. The following values were taken into consideration: $\underline{x}=100$, $x_1=140$, $x_2=180$, $x_3=210$, $x_4=232$, $\overline{x}=260$.

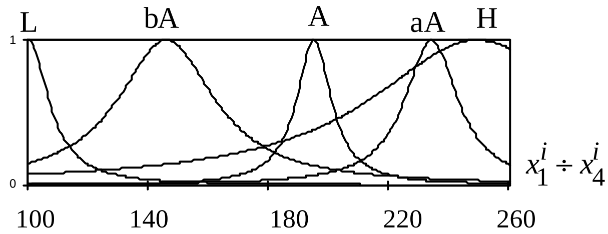


Fig. 6. Membership functions of linguistic estimates after training

The training time estimated on a Ñeleron 450 processor was equal to 2,5 minutes for 5000 iterations. It has allowed to train the neuro-fuzzy model of prediction in real time. The training was performed until the prognose produced by the neuro-fuzzy network was sufficiently close to experimental data. The application of tuned membership functions allows one to obtain a prediction model that is sufficiently close to the experimental data (Fig. 7).

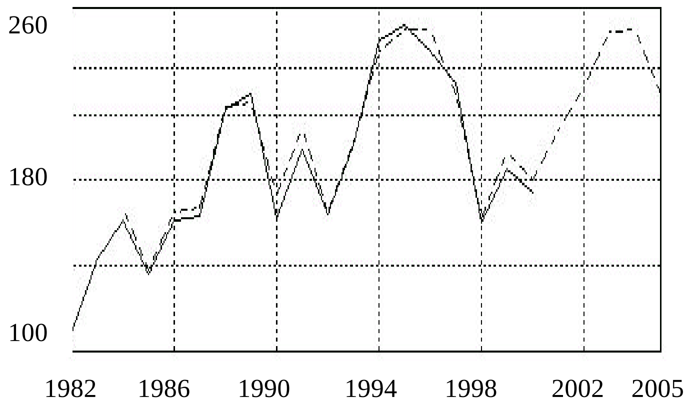


Fig. 7. Comparison of the experimental data (___) and prediction model (-----) after training

Since experimental values of the numbers of diseases in the years 1998-2000 have not been used in training the model, the proximity of the theoretical and experimental results for these years demonstrates the sufficient quality of the constructed prediction model from the practical viewpoint. A comparison of the results of simulation with the experimental data and also a prediction of the number of appendicular peritonitis diseases until 2005 is presented in Table 6.

Table 6. Diseases number evolution prediction

Year	1982	1983	1984	1985	1986	1987	1988	1989
Experiment	109	143	161	136	161	163	213	220
Theory			167	138	165	167	214	216
Error			6	2	4	4	1	4
Year	1990	1991	1992	1993	1994	1995	1996	1997
Experiment	162	194	164	196	245	252	240	225
Theory	173	204	165	197	240	250	250	220
Error	11	10	1	1	5	2	10	5
Year	1998	1999	2000	2001	2002	2003	2004	2005
Experiment	160	185	174					
Theory	162	193	180	203	223	249	250	220
Error	2	8	6					

5 Appendix

The partial derivatives in relations (5)-(7) characterize the sensitiveness of the error (E_t) to a change in parameters of a neuro-fuzzy network and are computed as follows:

$$\frac{\partial E_t}{\partial w_{jk}} = \varepsilon_1 \varepsilon_2 \varepsilon_3 \frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial w_{jk}},$$

$$\frac{\partial E_t}{\partial c_{1-4}^{ip}} = \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial c_{1-4}^{ip}},$$

$$\frac{\partial E_t}{\partial b_{1-4}^{ip}} = \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial b_{1-4}^{ip}},$$

where

$$\varepsilon_1 = \frac{\partial E_t}{\partial x_{1-4}^i} = x_t - \hat{x}_t ,$$

$$\varepsilon_2 = \frac{\partial x_{1-4}^i}{\partial \mu^{ip}\left(x_{1-4}^i\right)} = \frac{x_p \sum\limits_p \mu^{ip}\left(x_{1-4}^i\right) - \sum\limits_p x_p \mu^{ip}\left(x_{1-4}^i\right)}{\left(\sum\limits_p \mu^{ip}\left(x_{1-4}^i\right)\right)^2},$$

$$\varepsilon_3 = \frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial \left(\prod\limits_{l=1}^{n_{jk}} \mu_l^{ip}\left(x_{1-4}^i\right)\right)} = w_{jk} ,$$

$$\varepsilon_4 = \frac{\partial \left(\prod\limits_{l=1}^{n_{jk}} \mu_l^{ip}\left(x_{1-4}^i\right)\right)}{\partial \mu^{ip}\left(x_{1-4}^i\right)} = \frac{1}{\mu^{ip}\left(x_{1-4}^i\right)} \prod_{l=1}^{n_{jk}} \mu_l^{ip}\left(x_{1-4}^i\right),$$

$$\frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial w_{jk}} = \prod_{l=1}^{n_{jk}} \mu_l^{ip}\left(x_{1-4}^i\right),$$

$$\frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial c_{1-4}^{ip}} = \frac{2 c_{1-4}^{ip}\left(x_{1-4}^i - b_{1-4}^{ip}\right)^2}{\left(\left(c_{1-4}^{ip}\right)^2 + \left(x_{1-4}^i - b_{1-4}^{ip}\right)^2\right)^2}$$

$$\frac{\partial \mu^{ip}\left(x_{1-4}^i\right)}{\partial b_{1-4}^{ip}} = \frac{2\left(c_{1-4}^{ip}\right)^2\left(x_{1-4}^i - b_{1-4}^{ip}\right)}{\left(\left(c_{1-4}^{ip}\right)^2 + \left(x_{1-4}^i - b_{1-4}^{ip}\right)^2\right)^2}$$

6 Conclusion

In this paper a neuro-fuzzy method is proposed for the prediction of the evolution of diseases on the basis of expert information on regularities that can be revealed in available experimental data. The use of expert information in the form of the natural language rules IF-THEN formalized by means of fuzzy logic allows to construct models of prediction in the case of relatively small (in comparison with statistical methods) samples of experimental data. An advantage of the prediction model proposed is the transparency and invariability of the rules IF-THEN , in which the membership functions of fuzzy terms can be trained as new experimental data are accumulated. The neuro-fuzzy technology used for diseases number evolution prediction allows to organize the training process in real time and to use new experimental data for on-line improvement of the prediction model. The method proposed can be recommended for the prediction of discrete sequences not only in medicine but also in physico-technical applications, economics, sociology, etc.

References:

[1] A.G. Ivakhnenko, *Prediction of Random Processes*, Naukova Dumka: Kiev, 1971.

[2] F.Mayoraz, Th. Cornu, and L. Vulliet, Prediction of slope movements using neural networks, *Fuzzy Systems Artif. Intel.*, Vol.4, No.1, 1995, pp. 9-17.

[3] Rotshtein A.P., Loiko E.E., Katel'nikov D.I., Prediction of the diseases number on the basis of expert-linguistic information, *Cybernetics and Systems Analysis*,Vol.35, No.2,1999, pp.335-342.

[4] Zimmermann H.-J., *Fuzzy Set Theory and Its Application*, Kluwer: Dordrecht, 1991.

[5] Rotshtein A., Design and Tuning of Fuzzy Rule-Based Systems for Medical Diagnosis. In N.-H. Teodorescu(ed.): *Fuzzy and Neuro-Fuzzy Systems in Medicine*, CRC Press, 1998, pp. 243-289.

[6] Rotshtein A.P., Mityushkin Y.I., Neurolinguistic identification of nonlinear dependences, *Cybernetics and Systems Analysis*, Vol.36, No.2, 2000, pp.179-187.

[7] Tsypkin Ya., *Information Theory of Identification*, Nauka: Moscow, 1984.