PREDICTION OF THE RESULTS OF FOOTBALL GAMES BASED ON FUZZY MODEL WITH GENETIC AND NEURO TUNING

Alexander Rotshtein, Morton Posner*, Hanna Rakytyanska***

Dept. of Industrial Engineering and Management Dept. of Applied Mathematics
Jerusalem College of Technology - Machon Lev
21 Havaad Haleumi, 91160, Jerusalem, Israel
rot@mail.jct.ac.il posner@mail.jct.ac.il

Vinnitsa State Technical University
95 Khmelnitske sh., 21021 Vinnitsa, Ukraine
h_rakit@hotmail.com

The model of the football game result prediction, which uses information about previous results of opponent teams is proposed. The model is based on the method of non-linear dependencies identification using fuzzy knowledge bases. The reasonable results of simulation can be reached by fuzzy rules tuning using tournament tables data. The tuning procedure implies choosing fuzzy terms membership functions pirameters and rules weights by means of genetic and neural optimization technique combination. The proposed model can be used for commercial programs making prediction of the football games results in bookmaker's offices.

Key words: sport game prediction model, fuzzy IF-THEN rules, tuning of the fuzzy model, genetic algorithm, neural network

1. Introduction

Football is the most popular sport game attracting maximum number of fans. Prediction of football matches results arouses interest from two points of view: the first one is demonstration of the power of different mathematical methods [5,7], the second one is the desire of making money by telling beforehand any winning result

Models and PC-programs of sport prediction are being developed for many years (see, for example, http://dmiwww.cs.tutfl/riku). Most of mem use stochastical methods of uncertainty description: regressive and autoregressive analysis [4,8,20], Bayessian approach in combination with Markov chains and method Monte-Carlo [2, 6, 16, 17]. The specific features of these models are: sufficiently great complexity, a lot of assumptions, the need in big number of statistical data. Besides that, the models can not always be easily interpreted. Some several years passed before some models using neural networks for the results of football games prediction appeared [1, 9, 19]. They can be considered as universal approximators of non-linear dependencies trained by experimental data. These models also need a lot of statistical data and do not allow to define the physical meaning of the weights between neurons after training.

In the practice of prediction making the football experts and fans usually make good decisions using simple reasonings on the common sense level, for example.

IF *team* T_1 *constantly won in previous matches*
AND *team* T_2 *constantly lost in previous matches*
AND *in previous matches between teams* T_1 *and* T_2 *team* T_1 *won,*
THEN *win of team* T_1 *should be expected.*

Such expressions can be considered as concentration of accumulated experts experience and can be formalized using fuzzy logic [21]. That is why it is quite natural that we get use of such expressions as a support for building a model of prediction.

The method of non-linear dependencies identification based on fuzzy logic has been proposed in [10,11]. Different theoretical and practical aspects of this method are considered in [12-15]. In this paper we describe application of fuzzy knowledge bases and method [10,11] for football games results prediction.

The process of modeling has two phases. In the first phase we define the fuzzy model structure, which connects the football game result to be found with the results of previous games for both teams. For such modeling we can use generalized fuzzy approximator, proposed in [10, 11]. The second phase consists of fuzzy model tuning, i.e. of finding optimal parameters using available experimental data. For tuning we use a combination of genetic algorithm and neural network. The genetic algorithm provides rough finding of the area of global minimum of distance between model and experimental results. We use the neural approach for the fine model parameters tuning and for their adaptive correction while new experimental data is appearing.

2. Fuzzy model of prediction

2.1 The Structure of the Model

The aim of modeling is to calculate the result of match between teams T_1, and T_2, which is characterized as the difference of scored and lost goals y. We assume that $y \in [\underline{Y}, \overline{Y}] = [-5,5]$. For prediction model building we will define the value of y on the following five levels:

d_1 is big loss *(BL)*, $y = -5,-4,-3$;
d_2 is small loss *(SL)*, $y = -2,-1$;
d_3 is draw (D), $y=0$;
d_4 is small win *(SW)*, $y = 1,2$;
d_5 is big win *(BW)* $y = 3,4,5$.

Let us suppose that the football game result (y) is influenced by the following factors:

$x_1, x_2, ..., x_5$ are the results of five previous games for team T_1;
$x_6, x_7, ..., x_{10}$ are the results of five previous games for team T_2;
x_{11}, x_{12} the results of two previous games between teams T_1 and T_2.

It is obvious, that values of factors $x_1, x_2, ..., x_{12}$ are changing in the range from -5 to 5.

The hierarchical interconnection between output variable y and input variables $x_1, x_2, ..., x_{12}$ is represented as a tree shown in Fig. 1.

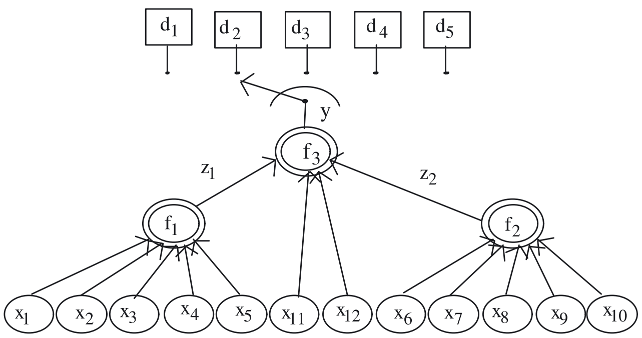


Figure 1. Structure of the Prediction Model

This tree is equal to system of correlations

$$y = f_3(z_1, z_2, x_{11}, x_{12}), \tag{1}$$

$$z_1 = f_1(x_1, x_2, ..., x_5), \tag{2}$$

$$z_2 = f_2(x_6, x_7, ..., x_{10}), \tag{3}$$

where z_1 (z_2) is football game prediction for team T_1 (T_2) based on the previous results $x_1, x_2, ..., x_5$ $x_6, x_7, ..., x_{10}$.

The variables $x_1, x_2, ..., x_{12}$, и z_1 (z_2) considered as linguistic variables [21], which can be evaluated using above mentioned fuzzy terms: *EL, SL, D, SW* and *BW*.

To describe the correlations (1)-(3) we shall use expert matrixes of knowledge (Tables 1,2). These matrixes correspond to fuzzy IF-THEN rules received on the common sense level. An example of one of these rules for Table 2 is given below:

IF $(x_{11}=BW)$ AND $(x_{12}=BW)$ AND $(z_1=BW)$ AND $(z_2=BL)$
OR $(x_{11}=SW)$ AND $(x_{12}=BW)$ AND $(z_1=SW)$ AND $(z_2=D)$
OR $(x_{11}=BW)$ AND $(x_{12}=D)$ AND $(z_1=BW)$ AND $(z_2=SL)$
THEN $y=d_5$.

Table 1.

Knowledge about correlations (2) and (3)

$x_1\,(x_6)$	$x_2\,(x_7)$	$x_3(x_8)$	$x_4\,(x_9)$	$x_5\,(x_{10})$	$z_1\,(z_2)$
BL	BL	BL	BL	BL	
BW	SL	BL	SL	BW	BL
SW	BL	SL	SL	SW	
SL	SL	SL	SL	SL	
D	SL	SL	D	D	SL
SW	D	SL	SL	SW	
D	D	D	D	D	
SL	SW	SW	D	SL	D
D	D	SW	SW	D	
SW	SW	SW	SW	SW	
D	BW	BW	SW	D	SW
SL	SW	SW	BW	SL	
BW	BW	BW	BW	BW	
SL	BW	SW	BW	SL	BW
BL	SW	BW	SW	BL	

Table 2.

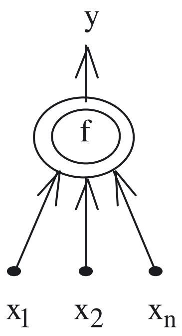
Knowledge about correlation (1)

x_{11}	x_{12}	z_1	z_2	y
BL	BL	BL	BW	
BW	D	BL	D	d_1
SW	BL	SL	SL	
SW	SL	D	SL	d_2
D	SL	SL	D	
SW	D	SL	SL	
D	D	D	D	
SL	SW	SW	D	d_3
SL	D	SW	SW	
SL	SW	SW	BW	
D	BW	BW	SW	d_4
SL	SW	SW	BW	
BW	BW	BW	BL	
SW	BW	SW	D	d_5
BW	D	BW	SL	

2.2 Fuzzy Approximator

For application of fuzzy rules bases (Tables 1,2) we use the generalized fuzzy approximator (Fig. 2), proposed in [10,11].

Figure 2. Generalized Fuzzy Approximator



Ibis approximator describes the dependence $y=f(x_1,x_2,...,x_n)$ between inputs $x_1,x_2,...,x_n$ and output y with use of expert matrix of knowledge (Table 3).

Table 3.

Expert matrix of knowledge

ı	IF <inputs>				THEN <output>	Weight of rule
	x_1	x_2	…	x_n	y	
11	a_1^{11}	a_2^{11}	…	a_n^{11}		w_{11}
12	a_1^{12}	a_2^{12}	…	a_n^{12}	d_1	w_{12}
…	…	…	…	…		…
$1k_1$	$a_1^{1k_1}$	$a_2^{1k_1}$	…	$a_n^{1k_1}$		w_{1k_1}
…	…	…	…	…	…	…
m1	a_1^{m1}	a_2^{m1}	…	a_n^{m1}		w_{m1}
m2	a_1^{m2}	a_2^{m2}	…	a_n^{m2}	d_m	w_{m2}
…	…	…	…	…		…
mk_m	$a_1^{mk_m}$	$a_2^{mk_m}$	…	$a_n^{mk_m}$		w_{mk_m}

The fuzzy knowledge base below corresponds to this matrix:

IF $[(x_1=a_1^{j1})$ AND ...$(x_i=a_i^{j1})$ AND ...$(x_n=a_n^{j1})]$(with weight w_{j1})...

... OR $[(x_1=a_1^{jk1})$ AND ...$(x_i=a_i^{jki})$ AND ...$(x_n=a_n^{jkn})]$ (with weight w_{jkj}),

THEN $y=dj, j=\overline{1,m}$ (4)

where a_i^p is linguistic term for variable x_i – evaluation in the row with number $p=kj$;

kj is number of conjunction rows corresponding to class dj of *output* variable y;

w_{ip} is a number in the range [0,1] characterizing the subjective measure of confidence of an expert relative to the statement with number $p=kj$.

Classes $dj, j=\overline{1,m}$, are formed by digitizing the range $[\underline{Y},\overline{Y}]$ of the output variable into the following m levels:

$$[\underline{y},\overline{y}]=\underbrace{[\underline{y},y_1)}_{d_1}\cup\ldots\cup\underbrace{[y_{j-1},y_j)}_{d_j}\cup\ldots\cup\underbrace{[y_{m-1},\overline{y}]}_{d_m}.$$

As is shown in [10-12], the following approximation of object corresponds to fuzzy knowledge base (4):

$$y=\frac{\underline{y}\mu^{d_1}(y)+y_1\mu^{d_2}(y)+...+y_{m-1}\mu^{d_m}(y)}{\mu^{d_1}(y)+\mu^{d_2}(y)+...+\mu^{d_m}(y)}, \quad (5)$$

$$\mu^{d_j}(y)=\max_{p=1,k_j}\left\{w_{jp}\min_{i=1,n}\left[\mu^{jp}(x_i)\right]\right\}, \quad (6)$$

$$\mu^{jp}(x_i)=\frac{1}{1+\left(\frac{x_i-b_i^{jp}}{c_i^{jp}}\right)^2},\quad i=\overline{1,n},\quad j=\overline{1,m}\quad p=k_j, \quad (7)$$

where $\mu^{dj}(y)$ is amembership function of Ihe output y to class $dj\in[y_{j-1},y_j]$;

$\mu^{jp}(x_i)$ is a membersMp function of a variable x_i,- to term a_i^p;

b_i^{jp}, c_i^{jp} are membership function parameters of tuning for variable x_f with next interpretation: b is coordinate of maximum, $\mu^{jp}(b_i^{jp})=1$; c is parameter of concentration (compression-extension).

Correlations (5)-(7) define the general model of non-linear function y =f($x_1,x_2,...,x_n$) as follows:

$$y=F(X,W,B,C), \quad (8)$$

where $X=(x_1,x_2,...,x_n)$ is vector of input variables,

$W=(w_1,w_2,...,w_n)$ is fuzzy rules weights vector,

$B=(b_1,b_2,...,b_n)$ and $C=(c_1,c_2,...,c_n)$ are vectors of membership functions parameters,

N is total number of rules,

q is total number of fuzzy terms,

F is operator of inputs-output connection, corresponding to formulae (5)-(7).

2.3 Fuzzy Model of Prediction

Using fuzzy approximator: (8) (Fig.2) and tree of evidence (Fig.l), the prediction model can be described in the following form:

$$Y=F_y(x_1,x_2,...,x_{12}, W_1, B_1, C_1, W_2, B_2, C_2, W_3, B_3, C_3,) \quad (9)$$

where Fy is operator of inputs- output connection, corresponding to correlations (1)- (3),

$$W_1=((w_1^{11},...,w_1^{13}),...,(w_1^{51},...,w_1^{53})),$$

$$W_2=((w_2^{11},...,w_2^{13}),...,(w_2^{51},...,w_2^{53})),$$

$$W_3=((w_3^{11},...,w_3^{13}),...,(w_3^{51},...,w_3^{53})),$$

are vectors of rales weights in the correlations (2), (3), (1), respectively,

$B_1=(b_{1-5}^{BL},b_{1-5}^{SL},b_{1-5}^{D},b_{1-5}^{SW},b_{1-5}^{BW})$, $B_2=(b_{6-10}^{BL},b_{6-10}^{SL},b_{6-10}^{D},b_{6-10}^{SW},b_{6-10}^{BW})$,

$B_3=(b_{11-12}^{BL},b_{11-12}^{SL},b_{11-12}^{D},b_{11-12}^{SW},b_{11-12}^{BW})$ are vectors of centers for variables $x_1,x_2,...,x_5$, $x_6,x_7,...,x_{10}$, and x_{11},x_{12} membersMp functions to terms $BL, SL,.,., BW$;

$C_1=(c_{1-5}^{BL},c_{1-5}^{SL},c_{1-5}^{D},c_{1-5}^{SW},c_{1-5}^{BW})$, $C_2=(c_{6-10}^{BL},c_{6-10}^{SL},c_{6-10}^{D},c_{6-10}^{SW},c_{6-10}^{BW})$,

$C_3=(c_{11-12}^{BL},c_{11-12}^{SL},c_{11-12}^{D},c_{11-12}^{SW},c_{11-12}^{BW})$ are vectors of concentration parameters for variables $x_1,x_2,...,x_5$, $x_6,x_7,...,x_{10}$ and x_{11},x_{12} a membersMp functions to terms $BL, SL,...,BW$.

In model (9) we assume that for all of variables $x_1,x_2,...,x_5$ fazzy terms $BL, SL,.,.,BW$ have the same membership functions. Same assumption we made for variables $x_6,x_7,...,x_{10}$ and variables x_{11},x_{12} (see Fig.6).

3. Problem statement of the fuzzy model tuning

Training data in the form of *M* pairs of experimental data assumed to be obtained with use of tournament tables:

$$\left\langle \hat{X}_l , \hat{y}_l \right\rangle, \; l=\overline{1,M},$$

where $$\hat{X}_l = \left\{ (\hat{x}_1^l, \hat{x}_2^l, ..., \hat{x}_5^l), (\hat{x}_6^l, \hat{x}_7^l, ..., \hat{x}_{10}^l), (\hat{x}_{11}^l, \hat{x}_{12}^l) \right\}$$

are the previous matches results for teams T_1 and T_2 in the experiment number *l*, y_1 is the game result between teams T_1 and T_2 in experiment number *l*.

The essence of the prediction model tuning consists in such membership functions parameters *(b-, c-)* and ftizzy rules weights (w-) finding, which provide for the minimum distance between theoretical and experimental results:

$$\sum_{l=1}^{M} \left(F_y(\hat{x}_1^l, \hat{x}_2^l, ..., \hat{x}_{12}^l, W_i, B_i, C_i) - \hat{y}_l \right)^2 = \min_{W_i, B_i, C_i}, \quad i=1,2,3$$

To solve the non- linear optimization problem (10) we propose a combination of genetic algorithm and neural network. The genetic algorithm provides for rough off-line finding the area of global minimum; and neural network is used for on-line improvement of unknown parameters values.

4. Genetic tuning of the fuzzy model

4.1 Structure of the Algorithm

To realize the genetic algorithm of optimization problem (10) solution it is necessary to define the following main notions and operations [3, 11]: *chromosome* – coded version of solution, *population* – initial set of solutions versions, *fitness function*- criterion of versions selection; *cross over* -operation of variants- offspring generation from variants-parents; *mutation* – random change of chromosome elements.

If *P(f)* are chromosomes-parents and *C(t)* are chromosomes- offsprings on a *t* -th iteration, then the general structure of the genetic algorithm will have Ihe form of: **begin**

begin
t:=0
Set the initial population *P(t)*
Evaluate the *P(t)* using fitness function
while (no condition of completion) do
Generate the *C(t)* by operation of cross- over with *P(t)*
Perform mutation of *C(t)* Evaluate C(/) using fitness function
Select the population *P(t+l)* from *P(t)* and *C(t)*
t:=t+l
end;
end.

4.2. Coding

We define the chromosome as the vector-row of binary codes of membership functions parameters and rules weights (Fig.3).

B_1	C_1	W_1	B_2	C_2	W_2	B_3	C_3	W_3

Figure 3. Structure of the Chromosome

4.3 Cross- over and Mutation

Operation of cross-over is defined in Fig.4. It consists of the chromosomes parts exchange in each of tiie vectors of membership iunctions parameters *(B_1, C_1, B_2, C_2, B_3,C_3)* and each of the rules weights vectors *(W_1, W_2, W_3)*. The points of cross-over shown in dotted lines are selected randomly. Upper symbols (1 and 2) fa the vectors of parameters correspond to the first and second chromosomes-parents.

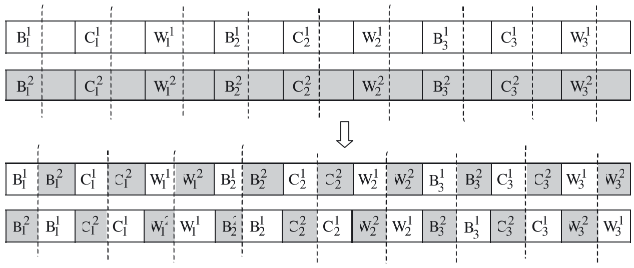


Figure 4. Structure of the Cross-over Operation

Mutation (*Mu*) implies random change (with some probability) of chromosome elements:

$$Mu\left(w_{jp}\right) = RANDOM\left([0,1]\right),$$

$$Mu\left(b_i^{jp}\right) = RANDOM\left(\left[\underline{y}, \overline{y}\right]\right),$$

$$Mu\left(c_i^{jp}\right) = RANDOM\left(\left[\underline{c}_i^{jp}, \overline{c}_i^{jp}\right]\right),$$

where *RANDOM*($[\underline{x}, \overline{x}]$) is the operation of random number finding which is uniformly distributed in interval $[\underline{x}, \overline{x}]$.

4.4 Selection

Selection of chromosomes-parents for cross-over operation should not be performed randomly. We used the selection procedure giving priority to the best solutions. The greater the fitness function of some chromosome the greater is the probability for the given chromosome to yield offsprings [3, 11]. We choose criterion (10) with sign minus as the fitness function, that is, the higher the degree of adaptability of the chromosome to perform the criterion of optimization the greater is the fitness function. While performing genetic algorithm the dimension of the population stays constant That is why after cross-over and mutation operations it is necessary to remove chromosomes having the fitness function of the worst significance from the obtained population.

5. Neural tuning of the fuzzy model

5.1 Neuro-Fuzzy Network of Prediction

For on-line tuning of the fuzzy prediction model we implanted fuzzy IF-THEN rules in a special neural network built with use of elements from Table 4 [14].

Table 4.

Elements of a Neuro-Fuzzy Network

Network Node	Node Name	Node Function	Network Node	Node Name	Node Function
	Input	$v=u$		Class of Rules	$v=\sum_{i=1}^{l} u_i$
	Fuzzy Term	$v=m^T(u)$			
	Fuzzy Rule	$v=\prod_{i=1}^{l} u_i$		Defuzzification ($\bar{d}_j$ is center of class d_j)	$v=\frac{\sum_{j=1}^{m} u_j \bar{d}_j}{\sum_{j=1}^{m} u_j}$

Thus, obtained neuro-fuzzy network is shown in Fig.5.

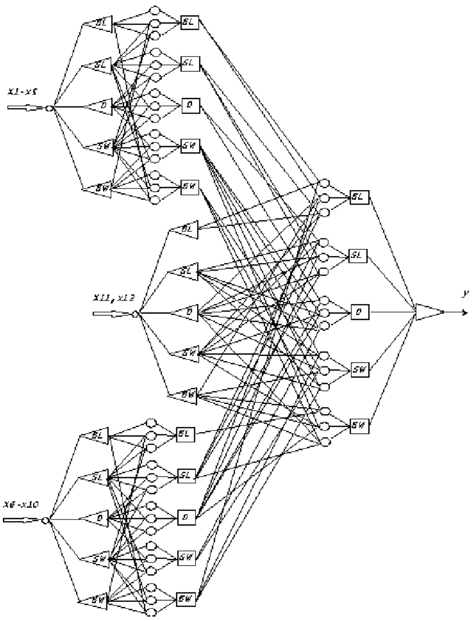


Figure 5. Neuro-Fuzzy Network of Prediction Making

5.2 Recursive Relations for Tuning

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

$$w_{jp}(t+1)=w_{jp}(t)-\eta\frac{\partial E_t}{\partial w_{jp}(t)}, \quad (11)$$

$$c_i^{jp}(t+1)=c_i^{jp}(t)-\eta\frac{\partial E_t}{\partial c_i^{jp}(t)}, \quad (12)$$

$$b_i^{jp}(t+1)=b_i^{jp}(t)-\eta\frac{\partial E_t}{\partial b_i^{jp}(t)}, \quad (13)$$

minimizing the criterion

$E_t=\frac{1}{2}(\hat{y}_t-y_t)^2$, which is used in the theory of neural networks, where y_t and y_t are theoretical and experimental difference of scored and lost goals at the fth step of training;

$w_{jp}(t)$, $c_i^{jp}(t)$, $b_i^{jp}(t)$ are rules weights and membership functions parameters at the t^th step of training;

η is parameter of training which can be chosen in accordance with the recommendations of [18].

The formulae for computation of partial derivatives in recursive relations (11-13) are presented in the Appendix.

6. Computer experiment results

For the fuzzy model timing we used the results from tournament tables of Finland Football Championship characterized by minimal number of sensations. Our training data included results of 1056 matches for the last 8 years from 1994 to 2001. The results of fuzzy model timing are given in Tables 5-8 and Fig.6.

Table 5.

Fuzzy rules weights in correlation (2)

Genethic algorythm	Neuro-fuzzy network
1.0	0.989
1.0	1.000
1.0	1.000
0.8	0.902
0.5	0.561
0.8	0.505
0.6	0.580
1.0	0.613
0.5	0.948
1.0	0.793
0.9	0.868
0.6	0.510
0.6	0.752
0.5	0.500
0.5	0.500

Table 6.

Fuzzy rules weights in correlation (3)

Genethic algorythm	Neuro-fuzzy network
0.7	0.926
0.9	0.900
0.7	0.700
0.9	0.954
0.7	0.700
1.0	1.000
0.9	0.900
1.0	1.000
0.6	0.600
1.0	1.000
0.7	0.700
1.0	1.000
0.8	0.990
0.5	0.500
0.6	0.600

Table 7.

Fuzzy rules weights in correlation (1)

Genethic algorythm	Neuro-fuzzy network
0.7	0.713
0.8	0.782
1.0	0.996
0.5	0.500
0.5	0.541
0.5	0.500
0.5	0.500
0.5	0.522
0.6	0.814
1.0	0.903
0.6	0.503
1.0	0.677
1.0	0.515
0.5	0.514
1.0	0.999

Table 8.

b- and c- parameters of membership functions after tuning

Terms	Genethic algorythm						Neuro-fuzzy network					
	$x_1, x_2, \ldots, x_5$		$x_6, x_7, \ldots, x_{10}$		x_{11}, x_{12}		$x_1, x_2, \ldots, x_5$		$x_6, x_7, \ldots, x_{10}$		x_{11}, x_{12}	
	b-	*c-*	*b-*	*c-*	*b-*	*c-*	*b-*	c-	*b-*	c-	*b-*	*c-*
BL	-4.160	9	-5.153	9	-5.037	3	-4.244	7.772	-4.524	9.303	-4.306	1.593
SL	-2.503	1	-2.212	5	-3.405	1	-1.468	0.911	-1.450	5.467	-2.563	0.555
D	-0.817	1	0.487	1	0.807	1	-0.331	0.434	0.488	7.000	0.050	0.399
SW	2.471	3	2.781	9	2.749	7	1.790	1.300	2.781	9.000	2.750	7.000
BW	4.069	5	5.749	9	5.238	3	3.000	4.511	5.750	9.000	3.992	1.234

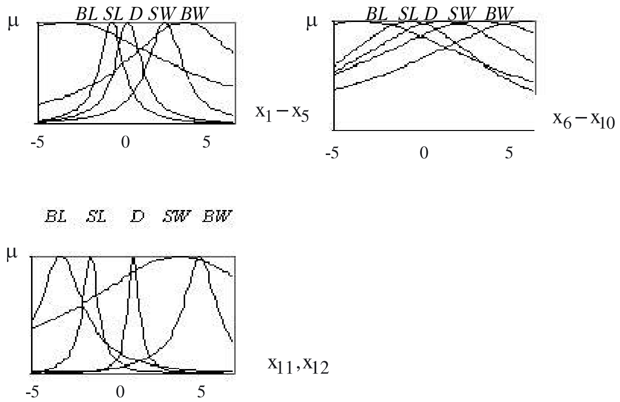


Figure 6. Membership Functions after Tuning

To test the prediction model we used the results of 350 matches from 1991 to 1993. The fragment of testing data and prediction results are shown in Table 9, where:

T_1, T_2 are teams names,

y, d are real (experimental) results,

y_G, d_G are results of prediction after genetic tuning of the fuzzy model,

y_N, d_N results of prediction after neural tuning of the fuzzy model.

Symbol * shows noncoincidences of theoretical and experimental results.

Table 9.

Fragment of the prediction results

1	T_1	T_2		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	Резуль-тат	$\hat{y}$	$\hat{d}$	y_G	d_G	y_N	D_N
1	Kuusysi	Reipas	1991	2	1	2	0	1	-1	0	1	-2	-3	2	1	2-0	2	d4	1	d4	1	d4
2	lives	PPT	1991	1	3	-1	1	0	0	2	-1	-2	0	0	0	2-1	1	d4	0	d3*	0	d3*
3	Haka	Jaro	1991	-1	2	0	-1	1	1	0	-2	-1	-2	-1	1	1-1	0	d3	0	d3	0	d3
4	MP	OTP	1991	3	1	2	0	2	-1	-2	1	-2	-3	1	3	4-0	4	d5	3	d5	3	d5
5	KuPS	HJK	1991	-1	-3	-4	1	-3	1	0	2	0	0	-2	0	1-3	-2	d2	-1	d2	-1	d2
6	IPS	RoPS	1991	3	1	2	-2	0	2	0	1	-1	1	0	-1	1-0	1	d4	0	d3*	0	d3*
7	PPT	Jaro	1991	0	-5	-1	0	1	1	2	-2	-1	1	1	-3	0-1	-1	d2	-1	d2	-1	d2
8	Haka	Reipas	1991	2	-1	3	1	4	2	-2	0	-1	0	-1	2	3-0	3	d5	2	d4*	2	d4*
9	OTP	Kuusysi	1991	-1	-2	-3	-2	0	1	3	4	-1	2	-2	-1	1-4	-3	d1	-3	d1	-3	d1
10	HJK	TPS	1991	1	1	1	0	2	0	1	-1	2	-3	0	2	2-0	2	d4	2	d4	2	d4
11	MyPa	Jaro	1992	-3	1	2	1	0	2	1	-2	-1	0	-2	0	0-0	0	d3	0	d3	0	d3
12	Jazz	lives	1992	2	2	1	-1	0	3	4	-1	0	1	1	-1	2-1	1	d4	0	d3*	1	d4
13	Haka	RoPS	1992	-2	-2	0	1	1	-1	1	1	1	0	1	3	1-1	0	d3	1	d4*	1	d4*
14	HJK	Oulu	1992	2	3	0	0	1	0	-5	1	-2	-1	-1	2	4-0	4	d5	2	d4*	3	d5
15	MP	Kuusysi	1992	0	1	-2	-1	-1	3	1	2	0	1	0	-2	0-3	-3	d1	-3	d1	-3	d1
16	KuPS	HJK	1992	-2	-1	-3	1	-2	4	2	1	2	1	-2	-3	0-5	-5	d1	-4	d1	-4	d1
17	Kuusysi	MP	1992	0	-1	3	2	-1	-3	2	-1	-2	0	1	0	3-1	2	d4	1	d4	1	d4
18	IPS	Haka	1992	-1	2	3	-1	-2	0	-1	0	3	1	-1	1	2-2	0	d3	0	d3	0	d3
19	RoPS	MyPa	1992	-2	-1	2	0	-1	1	-1	1	1	-2	1	-1	1-2	-1	d2	0	d3*	0	d3*
20	Jazz	lives	1992	-2	1	-3	5	-1	1	1	-2	0	-1	2	0	1-0	1	d4	1	d4	1	d4
21	TPS	Jaro	1992	-2	-1	2	-1	-3	1	0	2	-1	3	1	-2	0-2	-2	d2	-1	d2	-1	d2
22	Haka	MyPa	1992	1	1	-1	0	1	0	3	2	1	-1	-1	-3	0-1	-1	d2	-2	d2	-2	d2
23	HJK	RoPS	1992	1	2	0	-1	1	-1	2	2	-1	1	0	0	2-1	1	d4	0	d3*	0	d3*
24	MP	Kuusysi	1992	1	-1	-2	-3	1	1	-1	-2	2	3	-2	1	0-2	-2	d2	-1	d2	-1	d2
25	lives	Kups	1992	3	0	-2	2	-2	1	1	-1	0	-2	1	0	1-0	1	d4	1	d4	1	d4
26	Haka	HJK	1992	0	-2	-1	-1	0	2	3	-1	0	3	-1	-2	0-3	-3	d1	-3	d1	-3	d1
27	Jaro	MyPa	1992	-1	-1	1	2	1	-3	1	2	1	0	1	1	1-1	0	d3	1	d4*	0	d3
28	RoPS	TPS	1992	-1	1	-1	1	4	-5	-2	3	-1	-2	5	1	2-0	2	d4	2	d4	1	d4

Table 9. Continued.

1	T_1	T_2		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	Резуль-тат	$\hat{y}$	$\hat{d}$	y_G	d_G	y_N	D_N
29	MP	lives	1992	1	2	-1	1	0	0	1	0	0	-1	1	-2	2-3	-1	d2	-1	d2	-1	d2
30	Kuusysi	KuPS	1992	2	2	0	3	1	-1	-1	1	-3	0	2	3	4-1	3	d5	3	d5	3	d5
31	Jazz	MP	1993	2	2	2	0	3	-2	-1	0	-1	-3	4	3	5-0	5	d5	4	d5	4	d5
32	Kuusysi	TPS	1993	1	-1	0	-1	1	-2	2	0	-1	1	0	1	0-0	0	d3	0	d3	0	d3
33	MyPa	RoPS	1993	-1	-1	2	2	3	2	-1	1	2	-2	3	-1	2-0	2	d4	1	d4	1	d4
34	Haka	HJK	1993	-3	-1	-2	1	0	1	4	1	2	0	-1	-2	1-3	-2	d2	-1	d2	-1	d2
35	Jaro	lives	1993	2	0	-1	0	-1	-2	-1	-2	2	1	2	0	2-1	1	d4	1	d4	1	d4
36	lives	HJK	1993	1	-2	-1	-1	1	3	1	2	0	1	-1	-1	0-2	-2	d2	-1	d2	-1	d2
37	Jazz	Jaro	1993	2	1	0	1	5	-1	-2	-2	1	-1	2	1	3-0	3	d5	2	d4*	2	d4*
38	MyPa	MP	1993	1	3	1	-1	1	-1	0	2	-1	1	1	0	1-0	1	d4	1	d4	1	d4
39	Kuusysi	Haka	1993	-1	-2	1	1	2	-1	-3	1	-5	2	3	-1	3-1	2	d4	1	d4	1	d4
40	TPS	RoPS	1993	-1	1	-2	1	2	1	2	-1	1	-2	1	1	1-0	1	d4	1	d4	1	d4
41	MP	HJK	1993	-1	-1	0	2	-1	2	3	1	-1	1	-2	1	1-2	-1	d2	0	d3*	0	d3*
42	Kuusysi	Jaro	1993	2	2	-2	1	2	0	-1	2	-2	0	1	2	2-1	1	d4	1	d4	1	d4
43	Jazz	Haka	1993	2	3	2	-1	1	-1	-3	-4	-2	0	2	2	4-0	4	d5	3	d5	3	d5
44	FinnPa	MyPa	1993	-1	1	-2	-1	2	1	-2	-1	1	0	-1	-1	1-2	-1	d2	-1	d2	-1	d2
45	TPS	lives	1993	2	1	2	1	-1	2	2	-2	1	-3	0	2	2-0	2	d4	1	d4	1	d4
46	RoPS	Jazz	1993	-1	-1	2	-2	-1	4	1	5	0	2	1	-3	2-5	-3	d1	-3	d1	-3	d1
47	MyPa	lives	1993	5	0	2	1	1	-3	-1	-2	1	-2	3	0	5-1	4	d5	3	d5	3	d5
48	TPV	Kuusysi	1993	-2	-1	0	1	0	-1	0	2	-1	0	0	1	0-0	0	d3	0	d3	0	d3
49	RoPS	HJK	1993	-1	-1	1	-2	0	3	1	-2	1	1	-2	1	0-2	-2	d2	0	d3*	-1	d2
50	TPS	Jaro	1993	-1	-1	1	2	2	-2	-1	1	-2	1	3	1	1-0	1	d4	1	d4	1	d4

The efficiency characteristics of fuzzy model tuning algorithms for the testing data are shown in Table 10.

Table 10.

Tuning algorithms efficiency characteristics

Efficiency characteristics		Genetic Tuning	Neural Tuning
Tuning Time		52 min	7 min
Number of iterations		25000	5000
Probability of correct prediction for different decisions	d_1 – big loss	30 / 35 = 0.857	32 / 35 =0.914
	d_2 – small loss	64 / 84 = 0.762	70 / 84 = 0.833
	d_3 – draw	38 / 49 = 0.775	43 / 49 = 0.877
	d_4 - small win	97 / 126 = 0.770	106 / 126 = 0.841
	d_5 – big win	49 / 56 = 0.875	53 / 56 = 0.946

Table 10 shows, that the best prediction results we can receive for the marginal decision classes (the loss and win with big score d_1 and d_2), and the worst results of prediction we can receive for the small loss and small win (d_2 and d_4).

Appendix

The partial derivatives in relations (1 1) – (13) characterize the sensitiveness of the error (J,) to a change in parameters of a neuro-fuzzy network and are computed as follows:

$$\frac{\partial E_t}{\partial w_3^{jp}} = e_1 e_2 e_3 \frac{\partial \mu^{dj}(y)}{\partial w_3^{jp}}, \quad \frac{\partial E_t}{\partial c_{11,12}^{jp}} = e_1 e_2 e_3 e_4 \frac{\partial \mu^{jp}(x_i)}{\partial c_{11,12}^{jp}},$$

$$\frac{\partial E_t}{\partial b_{11,12}^{jp}} = e_1 e_2 e_3 e_4 \frac{\partial \mu^{jp}(x_i)}{\partial b_{11,12}^{jp}}, \quad \frac{\partial E_t}{\partial w_1^{jp}} = e_1 e_2 e_3 e_5 e_6 \frac{\partial \mu^{jp}(z_1)}{\partial w_1^{jp}},$$

$$\frac{\partial E_t}{\partial c_{1-5}^{jp}} = e_1 e_2 e_3 e_5 e_6 e_8 \frac{\partial \mu^{jp}(x_i)}{\partial c_{1-5}^{jp}}, \quad \frac{\partial E_t}{\partial b_{1-5}^{jp}} = e_1 e_2 e_3 e_5 e_6 e_8 \frac{\partial \mu^{jp}(x_i}{\partial b_{1-5}^{jp}}$$

$$\frac{\partial E_t}{\partial w_2^{jp}} = e_1 e_2 e_3 e_5 e_7 \frac{\partial \mu^{jp}(z_2)}{\partial w_2^{jp}}, \quad \frac{\partial E_t}{\partial c_{6-10}^{jp}} = e_1 e_2 e_3 e_5 e_7 e_9 \frac{\partial \mu^{jp}(x_i)}{\partial c_{6-10}^{jp}},$$

$$\frac{\partial E_t}{\partial b_{6-10}^{jp}} = e_1 e_2 e_3 e_5 e_7 e_9 \frac{\partial \mu^{jp}(x_i)}{\partial b_{6-10}^{jp}},$$

where

$$e_1 = \frac{\partial E_t}{\partial y} = y_t - \hat{y}_t, \quad e_2 = \frac{\partial y}{\partial \mu^{dj}(y)} = \frac{\bar{d}_j \sum_{j=1}^{m} \mu^{dj}(y) - \sum_{j=1}^{m} \bar{d}_j \mu^{dj}(y)}{\left(\sum_{j=1}^{m} \mu^{dj}(y)\right)^2}$$

$$e_3 = \frac{\partial \mu^{dj}(y)}{\partial\left(\mu^{jp}(z_1)\mu^{jp}(z_2)\mu^{jp}(x_{11})\mu^{jp}(x_{12})\right)} = w_3^{jp},$$

$$e_4 = \frac{\partial\left(\mu^{jp}(z_1)\mu^{jp}(z_2)\mu^{jp}(x_{11})\mu^{jp}(x_{12})\right)}{\partial \mu^{jp}(x_i)} =$$

$$= \frac{1}{\mu^{jp}(x_i)} \mu^{jp}(z_1)\mu^{jp}(z_2)\mu^{jp}(x_{11})\mu^{jp}(x_{12}), \ i = 11,12,$$

$$e_5 = \frac{\partial\left(\mu^{jp}(z_1)\mu^{jp}(z_2)\mu^{jp}(x_{11})\mu^{jp}(x_{12})\right)}{\partial \mu^{jp}(z_i)} =$$

$$= \frac{1}{\mu^{jp}(z_i)} \mu^{jp}(z_1)\mu^{jp}(z_2)\mu^{jp}(x_{11})\mu^{jp}(x_{12}), \ i = 1,2,$$

$$e_6 = \frac{\partial \mu^{jp}(z_1)}{\partial\left(\prod_{i=1}^{5} \mu^{jp}(x_i)\right)} = w_1^{jp}, \qquad e_7 = \frac{\partial \mu^{jp}(z_2)}{\partial\left(\prod_{i=6}^{10} \mu^{jp}(x_i)\right)} = w_2^{jp},$$

$$e_8 = \frac{\partial\left(\prod_{i=1}^{5} \mu^{jp}(x_i)\right)}{\partial \mu^{jp}(x_i)} = \frac{1}{\mu^{jp}(x_i)} \prod_{i=1}^{5} \mu^{jp}(x_i), \quad i = 1,2,\dots,5,$$

$$e_9 = \frac{\partial\left(\prod_{i=6}^{10} \mu^{jp}(x_i)\right)}{\partial \mu^{jp}(x_i)} = \frac{1}{\mu^{jp}(x_i)} \prod_{i=6}^{10} \mu^{jp}(x_i), \quad i = 6,7,\dots,10,$$

$$\frac{\partial \mu^{d_j}(y)}{\partial w_3^{jp}} = \mu^{jp}(z_1)\mu^{jp}(z_2)\mu^{jp}(x_{11})\mu^{jp}(x_{12}),$$

$$\frac{\partial \mu^{jp}(z_1)}{\partial w_1^{jp}} = \prod_{i=1}^{5} \mu^{jp}(x_i), \qquad \frac{\partial \mu^{jp}(z_2)}{\partial w_2^{jp}} = \prod_{i=6}^{10} \mu^{jp}(x_i),$$

$$\frac{\partial \mu^{jp}(x_i)}{\partial c_i^{jp}} = \frac{2c_i^{jp}\left(x_i - b_i^{jp}\right)^2}{\left(\left(c_i^{jp}\right)^2 + \left(x_i - b_i^{jp}\right)^2\right)^2},$$

$$\frac{\partial \mu^{jp}(x_i)}{\partial b_i^{jp}} = \frac{2\left(c_i^{jp}\right)^2\left(x_i - b_i^{jp}\right)}{\left(\left(c_i^{jp}\right)^2 + \left(x_i - b_i^{jp}\right)^2\right)^2}, \quad i = 1,2,\dots,12..$$

Conclusion

The proposed model provides football match result prediction using information about both teams previous games results. The prediction model is based on the method of identification of the nonlinear dependencies "past – future" by fuzzy IF-THEN rules.

The reasonable results of modeling can be achieved by fuzzy IF-THEN rules tuning with use of data from tournament tables. The tuning procedure consists of fuzzy terms membership functions parameters and fuzzy rules weights finding using combination of genetic (off-line) and neural (on-line) optimization techniques.

The future improvement of fuzzy prediction model can be done by taking into account some additional factors in fuzzy rules such as: the game on host/guest field, number of injured players, different psychological effects.

Presented here model can be used for creation of commercial PC-programs of football games results prediction for bookmaker's offices. Besides that the methodology of design and tuning of fuzzy model proposed in this paper can be used for fuzzy expert systems design in another areas.

REFERENCES:

[1] E. Condon, B. Golden, and E. Wasil, Predicting the Success of Nations at the Summer Olympic using Neural Networks, Computers & Operations Research 26 (13) (1999), 1243-65.

[2] M. Dixan, and S. Coles, Modeling Association Football Scores and Inefficiencies in the Football Betting Market, Applied Statistics 46 (2) (1997), 265-280.

[3] M. Gen, and R. Cheng, Genetic Algorithms and Engineering Design, John Wiley & Sons, 1997,352 p.

[4] M. Glickman, and H. Stem, A State-space Model for National Football League Scores, Journal of the American Statistical Association 93 (1998), 25-35.

[5] A. Ivachnenko, and V. Lapa, Forecasting of random processes, Naukova dumka, Kiev, 1971, 416 p. (In Russian)

[6] R Koning, Balance in Competition in Dutch Soccer, The Statistician 49 (2000), 419-431. [7] S. Markidakis, S. Wheelwright, and R. Hindman, Forecasting: Methods and Applications, 3 Ed., John Wiley & Sons, USA, 1998,386 p.

[8] J. Mingers, Rule Induction with Statistical Data – A Comparison with Multiple Regression, J. Operation Research Society 38 (1987), 347-351.

[9] M. Purucker, Neural Network Quarterbacking, IEEE Potentials 15 (3) (1996), 9-15.

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

[11] A. Rotshtein, Intellectual Technologies of Identification: Fuzzy Sets, Genetic Algorithms, Neural Nets, UNTVERSUM, Vinnitsa, 1999,320p. (in Russian)

[12] A. Rotshtein, and D. KateFnikov, Identification of Non-linear Objects by Fuzzy Knowledge Bases, Cybernetics and Systems Analysis 34 (5) (1998), 676-683.

[13] A. Rotshtein, E. Loiko, and D. KatePnikov, Prediction of the Number of Diseases on the basis of Expert-linguistic Information, Cybernetics and Systems Analysis 35 (2) (1999), 335-342.

[14] A. Rotshtein, and Y. Mityushkin, Neurolinguistic Identification of Nonlinear Dependencies, Cybernetics and Systems Analysis 36 (2) (2000), 179-187.

[15] A. Rotshtein, and S. Shtovba, Control of the Dynamic System based on Fuzzy Knowledge Base, Automatic and Calculating Technique 2 (2001), 23-30. (In Russian)

[16] H. Rue, and O. Salvensen, Prediction and Retrospective Analysis of Soccer Matches in a League, The Statistician 49 (2000), 399-418.

[17] H. Stern, On Probability of Winning a Football Game, The American Statistician 45 (1991), 179-183.

[18] TsypkinYa, Information Theory of Identification, Naaka, Moscow, 1984,320 p. (in Russian)

[19] S. Walczar, and J. Krause, Chaos, Neural Networks and Gaming, Proc. Of Third Golden West International Conference, Intelligent Systems, Las Vegas, NV, USA, Eds: E.A. Dordrecht, Kluwer, 1995, pp. 457-66.

[20] K. Willoughby, Determinants of Success in the CFL: a Logistic Regression Analysis, Proc. Of National Annual Meeting to the Decision Sciences, Atlanta, USA, Decision Sci. hist Vol.2, 1997, pp. 1026-1028.

[21] Zadeh LA. Outline of a New Approach to the Analysis of Complex Systems and Decision Process. IEEE Transactions SMC – 3,1 (1973), 28-44.

Alexander Rotshtein.
Dept. of Industrial Engineering and Management Dept. of Applied Mathematics Jerusalem College of Technology - Machon Lev 21 Havaad Haleumi, 91160, Jerusalem, Israel

E-mail: rot@mail.jct.ac.il

Morton Posner.
Dept. of Industrial Engineering and Management Dept. of Applied Mathematics Jerusalem College of Technology - Machon Lev 21 Havaad Haleumi, 91160, Jerusalem, Israel

E-mail: posner@mail.jct.ac.il

Hanna Rakytyanska.
Vinnitsa State Technical University
95 Khmelnitske sh., 21021 Vinnitsa, Ukraine

E-mail: h_rakit@hotmail.com