

ПРОГНОЗИРОВАНИЕ РЕЗУЛЬТАТОВ ФУТБОЛЬНЫХ МАТЧЕЙ С ПОМОЩЬЮ НЕЧЕТКИХ МОДЕЛЕЙ С ГЕНЕТИЧЕСКОЙ И НЕЙРОННОЙ НАСТРОЙКАМИ

**Александр Ротштайн, Мортон Познер*

***Анна Ракитянская*

**Кафедра промышленного инжиниринга и менеджмента
Технологический колледж Иерусалима –
Машон Лев
21 Хаваад Халеуми, 91160, Иерусалим, Израиль
rot@mail.jct.ac.il
posner@mail.jct.ac.il*

***Кафедра прикладной математики
Винницкий Государственный технический университет
Хмельницкое шоссе 95, 21021 Винница, Украина
h_rakit@hotmail.com*

Предложена модель прогнозирования результатов футбольных матчей, использующая информацию о предыдущих результатах команды противника. Модель основывается на методе идентификации нелинейных зависимостей, использующем нечеткие базы знаний. Правдоподобные результаты моделирования могут быть получены с помощью настройки нечетких правил с использованием данных из турнирных таблиц. Процедура настройки предполагает выбор параметров членских функций нечетких условий и весовых коэффициентов правил методами комбинации генетической и нейронной оптимизации. Предложенная модель может быть использована в коммерческих компьютерных программах для прогнозирования результатов футбольных матчей в букмекерских конторах.

Ключевые слова: модель прогнозирования спортивной игры, нечеткие правила Если-То (If-Then), настройка нечеткой модели, генетический алгоритм, нейронная сеть.

1. Введение

Футбол – очень популярный вид спорта, привлекающий огромное количество поклонников этой игры. Прогнозирование результатов футбольных матчей вызывает интерес по двум причинам: демонстрации возможностей дифференциальных математических методов [5,7] на его примере и желанию получить выигрыш, делая ставки на правильные результаты матча.

Модели и компьютерные программы предсказания результатов спортивных игр разрабатываются на протяжении многих лет (например, <http://dmiwww.cs.tutfl/riku>). Большинство из них используют стохастические методы описания неопределенности: регрессивный и авторегрессивный анализ [4,8,20], метод Байезиана в комбинации с цепями Маркова и методом Монте Карло [2, 6, 16, 17]. Осо-

бенностями таких моделей являются: достаточно высокая сложность, большое количество допущений, потребность в наличии большого массива статистических данных. Кроме того, эти модели не всегда легко интерпретировать. Через несколько лет появляются модели, использующие нейронные сети для предсказания результатов футбольного матча [1, 9, 19]. Их можно рассматривать как универсальные аппроксиматоры нелинейных зависимостей, опробованные на экспериментальных данных. Для них также необходимо иметь массивы статистических данных, а физическое значение весов между нейронами после обучения определить нельзя.

На практике эксперты и футбольные фанаты обычно достаточно точно предсказывают результаты игры с помощью простого рассуждения, основанного на здравом смысле, например:

IF (если) команда T_1 постоянно выигрывала в предыдущих матчах

AND (и) команда T_2 постоянно проигрывала в предыдущих матчах

AND (и) и в предыдущих матчах между командами T_1 и T_2 команда T_1 выигрывала

THEN (то) следует ожидать выигрыша команды T_1 .

Эти выражения могут рассматриваться как концепция накопленного экспертами опыта и могут быть формализованы с помощью нечеткой логики [21]. По этой причине мы использовали вышеприведенные выражения в качестве поддержки для построения модели прогнозирования.

Метод идентификации нелинейных зависимостей, основанный на нечеткой логике, был предложен в [10,11]. Различные теоретические и практические аспекты этого метода рассматриваются в [12-15]. В данной работе мы описываем применение нечетких баз знаний и метода [10,11] предсказания результатов футбольного матча.

Процесс моделирования имеет две фазы. В первой фазе определяется структура нечеткой модели, которая объединяет искомым результат матча с результатами предыдущих игр для обеих команд. Для моделирования может использоваться обобщенный нечеткий аппроксиматор, предложенный в [10, 11]. Во второй фазе происходит настройка нечеткой модели, т.е. по доступным экспериментальным данным находятся оптимальные параметры. Для настройки используется комбинация генетического алгоритма и нейронной сети. Генетический алгоритм позволяет приблизительно определить область глобального минимального расстояния между моделью и экспериментальными результатами. Для настройки точных параметров модели и их адаптивной коррекции до появления новых экспериментальных данных используется нейронный подход.

2. Нечеткая модель предсказания

2.1 Структура модели

Цель моделирования – расчет результата матча между командами T_1 и T_2 , характеризуемого разницей между забитыми и пропущенными голами y . Допустим $y \in [y, y] = [-5, 5]$. Для построения модели предсказания определим значение y на следующих пяти уровнях:

- d_1 большой проигрыш (BL), $y = -5, -4, -3$;
- d_2 малый проигрыш (SL), $y = -2, -1$;
- d_3 ничья (D), $y = 0$;
- d_4 малый выигрыш (SW), $y = 1, 2$;
- d_5 большой выигрыш (BW), $y = 3, 4, 5$.

Предположим, что на результат матча (y) влияют следующие факторы:

x_1, x_2, \dots, x_5 результаты пяти предыдущих игр для команды T_1 ;

x_6, x_7, \dots, x_{10} результаты пяти предыдущих игр для команды T_2 ;

x_{11}, x_{12} результаты двух предыдущих игр между командами T_1 и T_2 .

Очевидно, что значения факторов x_1, x_2, \dots, x_{12} изменяются в диапазоне от -5 до 5.

Иерархическая взаимосвязь между выходной переменной y и входными переменными x_1, x_2, \dots, x_{12} представлена в виде дерева на рис. 1.

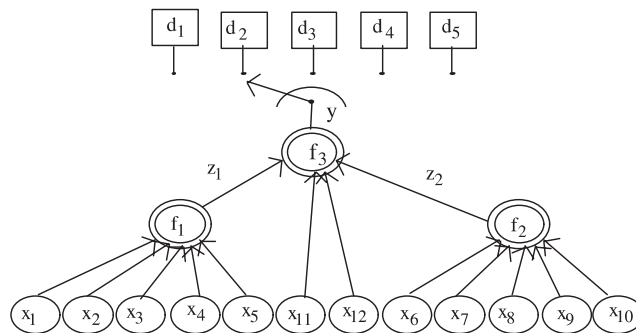


Рисунок 1. Структура модели предсказания

Это дерево можно представить в виде соотношений

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

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

$$z_2 = f_2(x_6, x_7, \dots, x_{10}),$$

где z_1 (z_2) – предсказание результата матча для команды T_1 (T_2) на основании предыдущих результатов x_1, x_2, \dots, x_5 (x_6, x_7, \dots, x_{10}).

Переменные x_1, x_2, \dots, x_{12} , и z_1 (z_2) рассматриваются как лингвистические переменные [21], которые могут быть оценены с помощью вышеупомянутых нечетких условий: BL, SL, D, SW and BW.

Для описания соотношений (1)–(3) используется матрица знаний экспертов (таблицы 1,2). Эти матрицы соответствуют нечетким правилам Если-То (IF-THEN), выведенным в соответствии со здравым смыслом. Пример одного из таких правил представлен ниже в таблице 2:

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 $y = d_5$.

Таблица 1.

Знание о соотношениях (2) и (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	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	

Таблица 2.

Знание о соотношении

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

2.2 Нечеткий аппроксиматор

Для применения баз нечетких правил (таблицы 1,2) используется обобщенный нечеткий аппроксиматор (рис. 2), предложенный в [10,11].

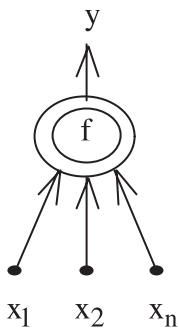


Рисунок 2. Обобщенный нечеткий аппроксиматор

Этот аппроксиматор описывает зависимость $y=f(x_1, x_2, \dots, x_n)$ между входными x_1, x_2, \dots, x_n и выходным y с помощью матрицы знаний экспертов (таблица 3).

Таблица 3.

Матрицы знаний экспертов

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

Нижеприведенная база нечетких знаний соответствует этой матрице:

IF $[(x_1 = a_1^{j1}) \text{ AND } \dots (x_i = a_i^{j1}) \text{ AND } \dots (x_n = a_n^{j1})]$ (с весом w_{j1})...
 ... OR $[(x_1 = a_1^{jk1}) \text{ AND } \dots (x_i = a_i^{jk1}) \text{ AND } \dots (x_n = a_n^{jk1})]$ (с весом w_{jk1}),
 THEN $y = d_j, j = 1, m$ (4)

где a_i^p – лингвистическое условие для оценки переменной x_i в ряду с номером $p = kj$;

kj – номер рядов сочетания, соответствующий классу d_j of выходной переменной y ;

w_{jp} – число в интервале $[0,1]$, характеризующее субъективное измерение уверенности эксперта по отношению к утверждению под номером $p = kj$.

Классы $d_j, j = 1, m$ формируются переводом в однозначные числа диапазона $[y]$ выходной переменной в следующей m уровней:

$$[y, \bar{y}] = \underbrace{[y_1, y_1]}_{d_1} \cup \dots \cup \underbrace{[y_{j-1}, y_j]}_{d_j} \cup \dots \cup \underbrace{[y_{m-1}, \bar{y}]}_{d_m}$$

Как показано в [10-12], нижеследующая аппроксимация объекта соответствует нечеткой базе знаний (4):

$$y = \frac{y_1 \mu^{d_1}(y) + y_2 \mu^{d_2}(y) + \dots + y_m \mu^{d_m}(y)}{\mu^{d_1}(y) + \mu^{d_2}(y) + \dots + \mu^{d_m}(y)}, \quad (5)$$

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

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

где $\mu^{d_j}(y)$ – членская функция выходной переменной y класса $d_j \in [y_{j-1}, y_j]$;

$\mu^{jp}(x_i)$ – членская функция переменной x_i по условию a_i^p ;

b_i^{jp}, c_i^{jp} – параметры членской функции настройки для переменной x_i со следующей интерпретацией: b – координаты максимума, $\mu^{jp}(b_i^{jp}) = 1$; c – параметр концентрации (компрессия – расширение).

Соотношения (5)-(7) определяют общую модель нелинейной функции $y = f(x_1, x_2, \dots, x_n)$ таким образом

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

где $X = (x_1, x_2, \dots, x_n)$ – вектор входных переменных,
 $W = (w_1, w_2, \dots, w_n)$ – весовой вектор нечетких правил,
 $B = (b_1, b_2, \dots, b_n)$ и $C = (c_1, c_2, \dots, c_n)$ – векторы параметров членской функции,

N – общее количество правил,
 q – общее количество нечетких условий,
 F – оператор связи вход-выход, соответствующий формулам (5)-(7).

2.3. Нечеткая модель предсказания

С использованием нечеткого аппроксиматора (8) (рис.2) и дерева предсказаний (рис.1), модель предсказания может быть представлена в следующем виде:

$$Y = F_y(x_1, x_2, \dots, x_{12}, W_p, B_p, C_p, W_2, B_2, C_2, W_3, B_3, C_3) \quad (9)$$

где F_y – оператор связи вход-выход, соответствующий формулам (1)-(3),

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

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

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

векторы весов правил в формулах (2), (3), (1), соответственно;

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

$B_3 = (b_{11-12}^{BL}, b_{11-12}^{SL}, b_{11-12}^D, b_{11-12}^{SW}, b_{11-12}^{BW})$ – векторы центров переменных $x_1, x_2, \dots, x_5, x_6, x_7, \dots, x_{10}$ и x_{11}, x_{12} – членская функция условий BL, SL, \dots, BW .

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

$C_3 = (c_{11-12}^{BL}, c_{11-12}^{SL}, c_{11-12}^D, c_{11-12}^{SW}, c_{11-12}^{BW})$ – векторы параметров концентрации для переменных $x_1, x_2, \dots, x_5, x_6, x_7, \dots, x_{10}$ и x_{11}, x_{12} – членская функция условий BL, SL, \dots, BW .

В модели (9) мы допускаем, что для всех переменных x_1, x_2, \dots, x_5 нечеткие условия BL, SL, \dots, BW имеют одинаковую членскую функцию. Такое же допущение справедливо для переменных x_6, x_7, \dots, x_{10} и переменных x_{11}, x_{12} (см. рис.6).

3. Постановка задачи настройки нечеткой модели

Обучающие данные в форме пар M экспериментальных данных предположительно получены с помощью турнирных таблиц:

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

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

результаты предыдущих матчей команд T_1 и T_2 в опыте под номером l , y_l – результат матча между командами T_1 и T_2 в опыте под номером l .

Сущность настройки модели предсказания состоит в нахождении таких параметров членских функций (b -, c -) и весов нечетких правил (w -), которые обеспечивают минимальное расхождение между теоретическими и экспериментальными результатами:

$$\sum_{l=1}^M (F_y(\hat{x}_1^l, \hat{x}_2^l, \dots, \hat{x}_{12}^l, W_l, B_l, C_l) - \hat{y}_l)^2 = \min_{W_l, B_l, C_l}, \quad i=1,2,3$$

Для решения нелинейной оптимизационной задачи (10) предлагается комбинация генетического алгоритма и нейронной сети. Генетический алгоритм обеспечивает приблизительное офф-лайнное нахождение области глобального минимума; а нейронная сеть используется для он-лайнного улучшения неизвестных значений параметров.

4. Генетическая настройка нечеткой модели

4.1 Структура алгоритма

Для реализации генетического алгоритма решения оптимизационной задачи (10) необходимо определить следующие главные операции и понятия [3, 11]: *хромосома* – кодированная версия решения, *популяция* – первоначальный набор версий решений, *функция соответствия* – критерий выбора версий; *кроссовер* – операция генерирования вариантов-потомков от вариантов-родителей; *мутация* – случайное изменение элементов хромосом.

Если $P(f)$ - хромосомы –родители и $C(t)$ – хромосомы-потомки в t -й итерации, то общая структура генетического алгоритма будет иметь такую форму:

begin

$t:=0$

Определить начальную популяцию как $P(t)$

Оценить $P(t)$ используя функцию соответствия

while (условие завершения не соблюдено) **do**

Генерировать $C(t)$ операций кроссовера с $P(t)$

Выполнить мутацию $C(t)$

Оценить $C(t)$ используя функцию соответствия

Выбрать популяцию $P(t+1)$ из $P(t)$ и $C(t)$

$t:=t+1$

end;

end.

4.2. Кодирование

Определим хромосому как вектор-ряд бинарных кодов параметров членской функции и весов правил (Рис.3).

B_1	C_1	W_1	B_2	C_2	W_2	B_3	C_3	W_3
-------	-------	-------	-------	-------	-------	-------	-------	-------

Рисунок 3. Структура хромосомы

4.3 Кроссовер и мутация

Операция кроссовера определена в рис. 4. Она состоит в обмене частями хромосом в каждом их векторов параметров членских функций ($B_p, C_p, B_2, C_2, B_3, C_3$) и каждом их векторов весов правил (W_1, W_2, W_3). Точки кроссовера, показанные пунктирными линиями, выбираются случайным образом. Верхние индексы (1 и 2) в векторах параметров соответствуют первой и второй хромосоме-родителям.

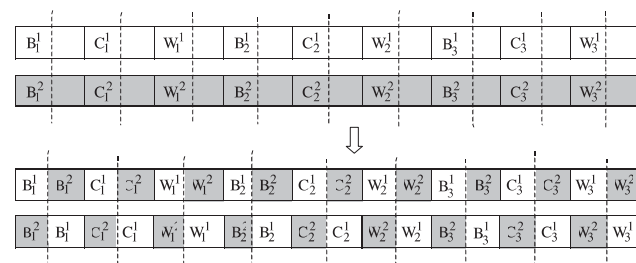


Рисунок 4. Структура операции кроссовера

Мутация (Mu) предполагает случайное изменение (c некоторой вероятностью) элементов хромосом:

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

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

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

где $RANDOM([\underline{x}, x])$ – операция нахождения случайного числа, равномерно распределенного в интервале $[\underline{x}, x]$.

4.4 Выбор

Выбор хромосом-родителей для операции кроссовера не должен выполняться случайным образом. Используем процедуру выбора, отдавая предпочтение лучшим решениям. Чем больше функция пригодности некоторых хромосом, тем больше вероятность получения такой хромосо-

мой потомков [3, 11]. Используем критерий (10) со знаком минус в качестве функции пригодности, т.е., чем выше степень адаптируемости хромосомы к выполнению критерия оптимизации, тем больше функция пригодности. При выполнении генетического алгоритма размеры популяции остаются неизменными. Поэтому важно удалять хромосомы с худшей функцией пригодности их полученной популяции после операций кроссовера и мутации.

5. Нейронная настройка нечеткой модели

5.1. Нейро- нечеткая сеть предсказания

Для он-лайнной настройки нечеткой модели предсказания введем нечеткий правила Если-То (IF-THEN) в специальную нейронную сеть, построенную с использованием элементов таб. 4 [14].

Элементы нейро-нечеткой сети.

Узел сети	Имя узла	Функция узла	Узел сети	Имя узла	Функция узла
	Ввод	$v = u$		Класс правил	$v = \sum_{i=1}^1 u_i$
	Нечетное условие	$v = m^T(u)$		Устранение нечетности (\bar{d}_j центр класса \bar{d}_j)	$v = \frac{\sum_{j=1}^m u_j \bar{d}_j}{\sum_{j=1}^m u_j}$
	Нечетное правило	$v = \prod_{i=1}^1 u_i$			

Полученная таким образом нейро-нечеткая сеть показана на рис.5.

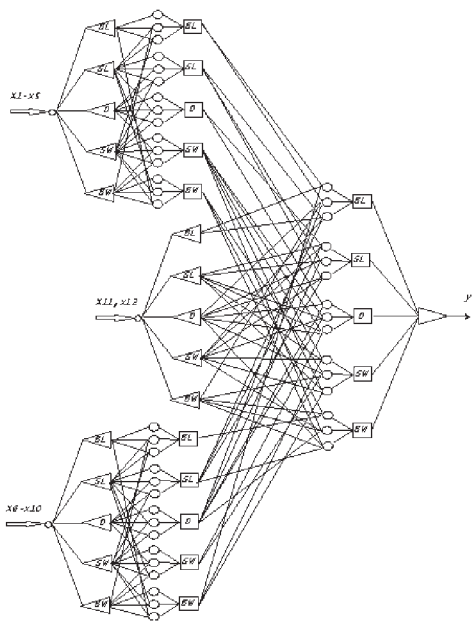


Рисунок 5. Нейро-нечеткая сеть предсказания

5.2 Рекурсивные соотношения для настройки

Следующая система рекурсивных соотношений используется для он-лайнного обучения модели предсказания:

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

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

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

минимизируя критерий

$$E_t = \frac{1}{2} (\hat{y}_t - y_t)^2$$

используемый для теории нейронных сетей, где \hat{y}_t и y_t – теоретическая и экспериментальная разница забитых и пропущенных голов на t -м шаге обучения;

Таблица 4. пропущенных голов на t -м шаге обучения;

$w_{jp}(t)$, $c_i^{jp}(t)$, $b_i^{jp}(t)$ – веса правил и параметры членских функций на t -м шаге обучения; η – параметр обучения, который может быть выбран в соответствии с рекомендациями [18].
 Формулы для расчета частных производных в рекурсивных соотношениях (11-13) представлены в Приложении.

6. Результаты компьютерного эксперимента

Для настройки нечеткой модели использовались результаты турнирных таблиц Финского футбольного чемпионата, принесшего минимальное количество сюрпризов. Данные обучения включают результаты 1056 матчей последних 8 лет с 1994 до 2001. Результаты настройки нечеткой модели представлены в таблицах 5-8 и на рис.6.

Таблица 5. Веса нечетких правил в соотношении (2)

Генетический алгоритм	Нейро-нечеткая модель
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

Таблица 6.
Веса нечетких правил в соотношении (3)

Генетический алгоритм	Нейро-нечеткая модель
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

Таблица 7.
Веса нечетких правил в соотношении (1)

Генетический алгоритм	Нейро-нечеткая модель
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

Таблица 8.
b- и c- параметры членских функций после настройки

Условия	Генетический алгоритм						Нейро-нечеткая сеть					
	x_1, x_3, \dots, x_5		x_6, x_7, \dots, x_{10}		x_{11}, x_{12}		x_1, x_3, \dots, x_5		x_6, x_7, \dots, 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

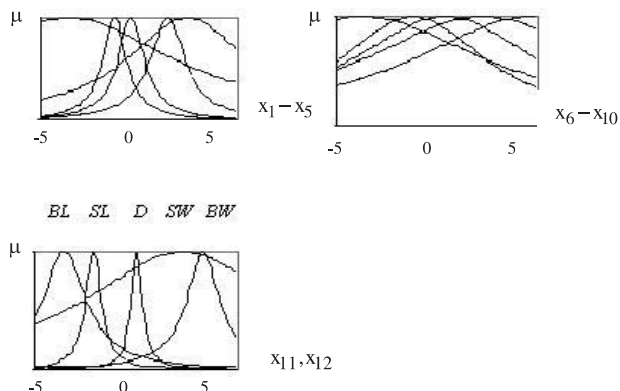


Рисунок 6. Членские функции после настройки

Для тестирования модели предсказания использовались результаты 350 матчей с 1991 до 1993. Фрагмент тестовых данных и результатов предсказания показан в таблице 9, где:
 T_1, T_2 – названия команд,
 y, d – реальные (экспериментальные) результаты,
 y_G, d_G – результаты предсказания после генетической настройки нечеткой модели,
 y_N, d_N – результаты предсказания после нейронной настройки нечеткой модели.
 Символ * указывает на несовпадения теоретических и экспериментальных результатов.

Фрагмент результатов предсказания

Таблица 9.

1	T_1	T_2		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	Результат	\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

Продолжение таблицы 9.

1	T ₁	T ₂		x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	Результат	\hat{y}	\bar{d}	y_G	d_G	y_N	D_N
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
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

Характеристики эффективности алгоритмов настройки нечеткой модели для тестовых данных размещены в таблице 10.

Таблица 10.

Алгоритмы настройки характеристик эффективности

Характеристики эффективности		Генетическая настройка	Нейронная настройка
Время настройки		52 мин	7 мин
Количество итераций		25000	5000
Вероятность правильного предсказания для разных решений	d_1 - большой проигрыш	30/35=0.857	32/35=0.914
	d_2 - малый проигрыш	64/84 = 0.762	70 / 84 = 0.833
	d_3 - ничья	38/49 = 0.775	43/49 = 0.877
	d_4 - малый выигрыш	97 / 126 = 0.770	106/126 = 0.841
	d_5 - большой выигрыш	49/56 = 0.875	53 / 56 = 0.946

Таблица 10 показывает, что лучшие результаты предсказания получаются для предельных классов решения (большой выигрыш и проигрыш d_1 и d_5), а худшие результаты предсказания – для небольшого выигрыша и проигрыша (d_2 и d_3).

Приложение

Частные производные в соотношениях (11) – (13) характеризуют чувствительность ошибки (E_t) к изменению параметров нейро-нечеткой сети и рассчитываются следующим образом:

$$\frac{\partial E_t}{\partial w_3^{jp}} = e_1 e_2 e_3 \frac{\partial m^{dj}(y)}{\partial w_3^{jp}}, \quad \frac{\partial E_t}{\partial c_{11,12}^{jp}} = e_1 e_2 e_3 e_4 \frac{\partial m^{jp}(x_1)}{\partial c_{11,12}^{jp}},$$

$$\frac{\partial E_t}{\partial b_{11,12}^{jp}} = e_1 e_2 e_3 e_4 \frac{\partial m^{jp}(x_1)}{\partial b_{11,12}^{jp}}, \quad \frac{\partial E_t}{\partial w_1^{jp}} = e_1 e_2 e_3 e_5 e_6 \frac{\partial m^{jp}(z_1)}{\partial w_1^{jp}},$$

$$\frac{\partial E_t}{\partial c_{1-5}^{jp}} = e_1 e_2 e_3 e_5 e_6 e_8 \frac{\partial m^{jp}(x_1)}{\partial c_{1-5}^{jp}}, \quad \frac{\partial E_t}{\partial b_{1-5}^{jp}} = e_1 e_2 e_3 e_5 e_6 e_8 \frac{\partial m^{jp}(x_1)}{\partial b_{1-5}^{jp}},$$

$$\frac{\partial E_t}{\partial w_2^{jp}} = e_1 e_2 e_3 e_5 e_7 \frac{\partial m^{jp}(z_2)}{\partial w_2^{jp}}, \quad \frac{\partial E_t}{\partial c_{6-10}^{jp}} = e_1 e_2 e_3 e_5 e_7 e_9 \frac{\partial m^{jp}(x_1)}{\partial c_{6-10}^{jp}},$$

$$\frac{\partial E_t}{\partial b_{6-10}^{jp}} = e_1 e_2 e_3 e_5 e_7 e_9 \frac{\partial m^{jp}(x_1)}{\partial b_{6-10}^{jp}},$$

где

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

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

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

$$= \frac{1}{m^{jp}(x_i)} m^{jp}(z_1)m^{jp}(z_2)m^{jp}(x_{11})m^{jp}(x_{12}), \quad i = 1, 1, 1, 2,$$

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

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

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

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

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

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

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

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

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

Заклучение

Предложенная модель обеспечивает предсказание результата футбольного матча с использованием информации о результатах игр обеих команд. Модель предсказания базируется на методе идентификации нелинейных зависимостей "прошлое – будущее" нечеткими правилами Если-То (IF-THEN).

Правдоподобные результаты моделирования могут быть получены благодаря настройке нечетких правил Если-То (IF-THEN) с использованием данных турнирных таблиц. Процедура настройки заключается в нахождении параметров нечетких условий членских функций и весов нечетких правил при использовании комбинации генетических (офф-лайнных(off-line) и нейронных (онлайнных (on-line) оптимизационных методов.

В будущем нечеткая модели предсказания может быть усовершенствована посредством принятия в расчет при формировании нечетких правил некоторых дополнительных факторов, например таких, как: игра на своем/чужом поле, количество игроков с травмами, различные психологические эффекты.

Представленная в работе модель может быть использована при разработке коммерческих компьютерных программ предсказания результатов матчей для букмекерских контор. Кроме того, методология разработки и настройки данной нечеткой модели может быть использована для разработки нечетких экспертных систем в других областях.

Литература

- [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, in 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, 320 p. (in Russian)
- [12] A. Rotshtein, and D. KatePnikov, 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] TsyarkinYa, Information Theory of Identification, Nauka, 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.

Александр Ротштайн. Доктор технических наук, профессор кафедры промышленного инжиниринга и менеджмента Технологического колледжа Иерусалима – Machon Lev.

21 Havaad Haleumi, 91160,
Jerusalem, Israel

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

Мортон Познер. Профессор кафедры промышленного инжиниринга и менеджмента Технологического колледжа Иерусалима – Machon Lev

21 Havaad Haleumi, 91160,
Jerusalem, Israel

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

Анна Ракитянская. Кандидат технических наук, кафедра прикладной математики Винницкого государственного технического университета

E-mail: h.rakit@hotmail.com