ПРОГНОЗИРОВАНИЕ РЕЗУЛЬТАТОВ ФУТБОЛЬНЫХ МАТЧЕЙ НА ОСНОВЕ НЕЧЕТКИХ ПРАВИЛ

В работе предложена модель прогнозирования результатов футбольных матчей, основанная на нечетком логическом выводе по базе лингвистических правил типа «ЕСЛИ — ТО». Тестирование модели показывает, что полученные прогнозы хорошо согласуются с действительными исходами футбольных матчей. Предложенный подход может использоваться для предсказания результатов других спортивных соревнований, а также для экстраполяции и прогнозирования процессов в условиях ограниченного статистического материала в таких экономических задачах, как прогнозирование курсов валют и акций, прогнозирование объемов спроса и предложения товаров и услуг, оценка качества инновационных проектов и др. ■

Prediction of Football Match Results Based on Fuzzy Rules

A model for football match results prediction based on linguistic “IF — THEN” rules and fuzzy logic inference procedures is proposed in the article. The testing shows that the proposed model achieves satisfactory estimation of actual results. The given approach can be used both to predict the results of other sport events, and to gain good extrapolation in case of insufficient data in such economic tasks as prediction of exchange and stock rates, prediction of demand and supply of goods and services, assessment of innovation project quality etc. ■

С. Д. Штовба
к. т. н., доц.
Винницкого государственного технического университета
(Украина)

Serhiy D. Shtovba

В. В. Вивдюк
студент
Винницкого государственного технического университета
(Украина)

Viktor V. Vivdyuk

Научный консультант
д. т. н., проф. А. П. Ротштейн

Введение

В последнее время спорт привлекает все большее количество людей, аккумулирует значительные финансовые, материальные и интеллектуальные ресурсы, и постепенно превращается в важный элемент экономики. Прогнозирование результатов спортивных соревнований само по себе является важной задачей, составляющей основу букмекерского бизнеса. Кроме того, эта задача может служить хорошим полигоном для тестирования различных методов экстраполяции и прогнозирования результатов процессов в условиях ограниченного статистического материала при большом количестве влияющих факторов, некоторые из которых заранее неизвестны.

В настоящей работе предложена модель прогнозирования результатов футбольных матчей, в основу которой положены формализованные в виде нечеткой базы знаний лингвистические высказывания-правила типа «ЕСЛИ — ТО». Для настройки и тестирования модели использовались данные о матчах чемпионатов Украины по футболу 2000–2002 гг.

1. Постановка задачи

Украинский футбольный чемпионат представляет собой турнир четырнадцати команд, победитель среди которых определяется количеством набранных очков в очных встречах со всеми соперниками. Чемпионат проходит в два круга, то есть каждая команда играет 26 поединков.

С кибернетической точки зрения задача построения модели прогнозирования результата футбольного матча сводится к поиску функционального отображения вида:

$$X = \{x_1, x_2, \ldots, x_n\} \rightarrow D \in \{d_1, d_2, d_3\}, \tag{1.1}$$

где X — вектор влияющих факторов, которыми могут быть уровень мастерства команд, погодные условия, результаты прошлых встреч команд, место проведения игры (дома или в гостях) и т. д.; D — результат встречи, который оценивается термами: d_1 — выигрыш первой команды (хозяина поля), d_2 — ничья, d_3 — выигрыш второй (гостевой) команды.

Рис. 1. Типовая структура нечеткой модели.
Обозначения: X — входной четкий вектор; $\tilde{X}$ — вектор нечетких множеств, соответствующий входному вектору X; $\tilde{Y}$ — результат логического вывода в виде нечеткого множества; Y — выходной результат в виде четкого числа.

2. Селекция влияющих факторов

Нами отобраны следующие факторы, которые оказывают наибольшее влияние на результат поединка:

x_1 — разница потерь ведущих игроков (разница между количеством травмированных и дисквалифицированных футболистов в первой команде (хозяине поля) и количеством травмированных и дисквалифицированных футболистов в гостевой команде);

x_2 — разница игровых динамик (разница очков, набранных командой-хозяином поля и гостевой командой в последних пяти турах);

x_3 — разница в классе команд (разница мест, которые занимают команда-хозяин и гостевая команда в текущем чемпионате);

x_4 — «фактор поля» (рассчитываться как $HP/HG - GP/GG$, где HP — общее количество очков, набранное командой-хозяином поля в домашних играх текущего чемпионата; HG — общее количество домашних матчей, проведенных командой-хозяином в текущем чемпионате; GP — общее количество очков, набранное гостевой командой в текущем чемпионате на выезде; GG — общее количество выездных игр, проведенных гостевой командой в текущем чемпионате);

x_5 — «встречи команд» (разница забитых и пропущенных мячей двух команд во всех чемпионатах Украины).

Заметим, что значения этих факторов не являются конфиденциальной информацией — их можно легко определить до начала футбольного матча.

3. Нечеткая модель прогнозирования

Нечеткая модель представляет собой аппроксимацию зависимости «входы — выход» на основе лингвистических высказываний типа «ЕСЛИ — ТО» и операций нечеткого логического вывода [1, 2, 5]. Основные положения теории нечетких множеств, используемые в настоящем разделе, приведены в Приложении. Типовая структура нечеткой модели показана на рис. 1. Нечеткая модель содержит следующие блоки:

— фаззификатор преобразует фиксированный вектор влияющих факторов (X) в вектор нечетких множеств $\tilde{X}$, необходимых для выполнения нечеткого логического вывода;

— нечеткая база знаний содержит информацию о зависимости $Y = f(X)$ в виде лингвистических правил типа «ЕСЛИ — ТО»;

— машина нечеткого логического вывода на основе правил базы знаний определяет значение выходной переменной в виде нечеткого множества $\tilde{Y}$, соответствующего нечетким значениям входных переменных ($\tilde{X}$);

— дефаззификатор преобразует выходное нечеткое множество $\tilde{Y}$ в четкое число.

Для нечеткого моделирования зависимости (1.1) необходимо:

— представить входные ($x_1 - x_5$) и выходную (y) переменные в виде лингвистических переменных;

— формализовать в виде нечеткой базы знаний экспертные лингвистические высказывания о взаимосвязи входов и выхода;

— обучить нечеткую модель путем настройки функций принадлежности и весов правил с целью минимизации отклонения между результатами моделирования и экспериментальными данными.

В соответствии с [3, 4] синтез нечеткой базы знаний может трактоваться как этап структурной идентификации зависимости (1.1) по экспертно-экспериментальным данным. В этом случае база знаний представляет собой грубую модель зависимости (1.1), которую затем необходимо настроить на этапе параметрической идентификации.

3.1. Лингвистические переменные

Для лингвистической оценки входных и выходной переменных будем использовать терм-множества, приведенные в табл. 1. Для повышения точности прогнозирования при оценке выходной переменной используется пять лингвистических термов: «Крупный проигрыш», «Проигрыш», «Ничья», «Выигрыш», «Крупный выигрыш», которые зададим на универсальном множестве разницы голов, забитых командой-хозяином поля и гостевой командой. В этом случае:

$$d_1 = \text{Выигрыш} \cup \text{Крупный выигрыш};$$
$$d_2 = \text{Ничья};$$
$$d_3 = \text{Проигрыш} \cup \text{Крупный проигрыш}.$$

Таблица 1. Терм-множества входных и выходной переменных

Переменная	Терм-множества	Параметры функции принадлежности	
		c	b
x_1	Большая скамейка (БС)	2.55	-6
	Одинаковая скамейка (ОС)	2.55	0
	Короткая скамейка (КС)	2.55	6
x_2	Существенный проигрыш (СП)	4.25	-15
	Проигрыш (П)	4.25	-5
	Выигрыш (В)	4.25	5
	Существенный выигрыш (СВ)	4.25	15
x_3	Лидер (Л)	2.76	-13
	Верхняя половина (ВП)	2.76	-6.5
	Середина (С)	2.76	0
	Нижняя половина (НП)	2.76	6.5
	Аутсайдер (А)	2.76	13
x_4	Абсолютная неудача (АНд)	0.7	-2
	Неудача (Нд)	0.7	-0.33
	Преимущество (Пр)	0.7	1.33
	Абсолютное преимущество (АПр)	0.7	3
x_5	Позорные встречи (Пз)	8.5	-20
	Равные встречи (Р)	8.5	0
	Разгромные встречи (Рз)	8.5	20
y	Крупный проигрыш (КП)	0.64	-3
	Проигрыш (П)	0.44	-0.9
	Ничья (Н)	0.44	0
	Выигрыш (В)	0.44	0.9
	Крупный выигрыш (КВ)	0.64	3

Формализацию лингвистических термов осуществим с помощью гауссовской функции принадлежности:

$$\mu^t(x) = e^{\frac{-(x-b)^2}{2c^2}}, \qquad (3.1)$$

где $\mu^t(x)$ — функция принадлежности переменной x к терму t; b — параметр функции принадлежности, соответствующий координате максимума ($\mu^t(x) = 1$); c — параметр сжатия-растяжения функции принадлежности.

Параметры функций принадлежности для каждого лингвистического терма приведены в табл. 1. На этом этапе параметры функций принадлежности выбирались экспертно, впоследствии на этапе настройки они будут оптимизированы.

3.2. Нечеткая база знаний

Экспертные лингвистические высказывания, отражающие взаимосвязь между факторами $x_1 - x_5$ и результатом футбольного матча (y), представлены в табл. 2.

Каждая строка таблицы соответствует одному правилу, например, для первой строки это правило:

ЕСЛИ $x_1 =$ «Большая скамейка» И
$x_2 =$ «Существенный выигрыш» И
$x_3 =$ «Лидер» И
$x_4 =$ «Абсолютное преимущество» И
$x_5 =$ «Разгромные встречи»,
ТО $y =$ «Крупный выигрыш».

Нечеткий логический вывод результатов футбольного матча осуществляется по следующей системе нечетких логических уравнений:

$$\mu_{\text{КВ}}(X) =$$

$$\begin{gathered}\mu_{\text{БС}}(x_1) \wedge \mu_{\text{СВ}}(x_2) \wedge \mu_{\text{Л}}(x_3) \wedge \mu_{\text{АП}_\text{р}}(x_4) \wedge \mu_{\text{Р}_\text{з}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{ВП}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{Р}_\text{з}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{Л}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{Р}_\text{з}}(x_5) \vee \\ \mu_{\text{БС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{ВП}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{Р}}(x_5);\end{gathered}$$

$$\mu_{\text{В}}(X) =$$

$$\begin{gathered}\mu_{\text{БС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{ВП}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{Р}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}_\text{з}}(x_5) \vee \\ \mu_{\text{КС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{ВП}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{Р}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}_\text{з}}(x_5);\end{gathered}$$

$$\mu_{\text{Н}}(X) =$$

$$\begin{gathered}\mu_{\text{ОС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5) \vee \\ \mu_{\text{КС}}(x_1) \wedge \mu_{\text{СП}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{НП}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{П}_\text{з}}(x_5) \vee \\ \mu_{\text{БС}}(x_1) \wedge \mu_{\text{СП}}(x_2) \wedge \mu_{\text{ВП}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5);\end{gathered}$$

$$\mu_{\text{П}}(X) =$$

$$\begin{gathered}\mu_{\text{БС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{АН}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{НП}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{П}_\text{з}}(x_5) \vee \\ \mu_{\text{КС}}(x_1) \wedge \mu_{\text{СП}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{П}_\text{з}}(x_5) \vee \\ \mu_{\text{ОС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{А}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5);\end{gathered}$$

$$\mu_{\text{КП}}(X) =$$

Таблица 2. Нечеткая база знаний

№	x_1	x_2	x_3	x_4	x_5	y
1	БС	СВ	Л	АПр	Рз	КВ
2	ОС	В	ВП	Пр	Рз	КВ
3	ОС	П	Л	Пр	Рз	КВ
4	БС	В	ВП	Пр	Р	КВ
5	ОС	В	С	Нд	Рз	В
6	КС	П	ВП	Пр	Р	В
7	ОС	В	С	Нд	Рз	В
8	БС	СВ	НП	Пр	Р	В
9	ОС	В	С	Нд	Р	Н
10	КС	СП	С	Нд	Р	Н
11	ОС	П	НП	Пр	Пз	Н
12	БС	СП	ВП	Нд	Р	Н
13	БС	П	С	АНд	Р	П
14	ОС	В	НП	Нд	Пз	П
15	КС	СП	С	Пр	Пз	П
16	ОС	П	А	Нд	Р	П
17	КС	СП	А	АНд	Р	КП
18	ОС	СП	НП	Нд	Пз	КП
19	КС	П	НП	АНд	Р	КП
20	БС	СП	НП	Нд	Пз	КП

$$\begin{aligned}
&\mu_{\text{БС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{АН}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5) \vee \\
&\mu_{\text{ОС}}(x_1) \wedge \mu_{\text{В}}(x_2) \wedge \mu_{\text{НП}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{П}_\text{з}}(x_5) \vee \\
&\mu_{\text{КС}}(x_1) \wedge \mu_{\text{СП}}(x_2) \wedge \mu_{\text{С}}(x_3) \wedge \mu_{\text{П}_\text{р}}(x_4) \wedge \mu_{\text{П}_\text{з}}(x_5) \vee \\
&\mu_{\text{ОС}}(x_1) \wedge \mu_{\text{П}}(x_2) \wedge \mu_{\text{А}}(x_3) \wedge \mu_{\text{Н}_\text{д}}(x_4) \wedge \mu_{\text{Р}}(x_5),
\end{aligned} \tag{3.2}$$

где $\mu_{\text{А}}(x)$ — степень принадлежности значения переменной к лингвистическому терму А.

Приведенная система логических уравнений получена из базы знаний (табл. 2) путем замены термов на функции принадлежности и логических операций И и ИЛИ на операции минимума ($\wedge$) и максимума ($\vee$), соответственно.

3.3. Алгоритм прогнозирования

Прогнозирование результатов конкретного футбольного матча происходит по следующему алгоритму:

Шаг 1. Определить значения влияющих факторов $X^* = (x_1^*, x_2^*, x_3^*, x_4^*, x_5^*)$ для данного матча.

Шаг 2. Вычислить степени принадлежности значений влияющих факторов к нечетким термам из базы знаний, приведенной в табл. 2.

Шаг 3. Подставить найденные на шаге 2 значения в систему нечетких логических уравнений (3.2) и определить степени принадлежности решения к термам «Крупный проигрыш», «Проигрыш», «Ничья», «Выигрыш», «Крупный выигрыш».

Шаг 4. Определить результирующее нечеткое множество как объединение соответствующих нечетких множеств:

$$\tilde{y} = \bigcup_{q \in (\text{КП, П, Н, В, КВ})} \int_{-3}^{3} \min(\mu_q(X^*), \mu_q(y))/y.$$

Шаг 5. Провести дефаззификацию полученного на предыдущем шаге нечеткого множества с использованием метода центра тяжести [5]:

$$y = \frac{\int_{-3}^{3} y \mu_{\tilde{y}}(y) dy}{\int_{-3}^{3} \mu_{\tilde{y}}(y) dy}.$$

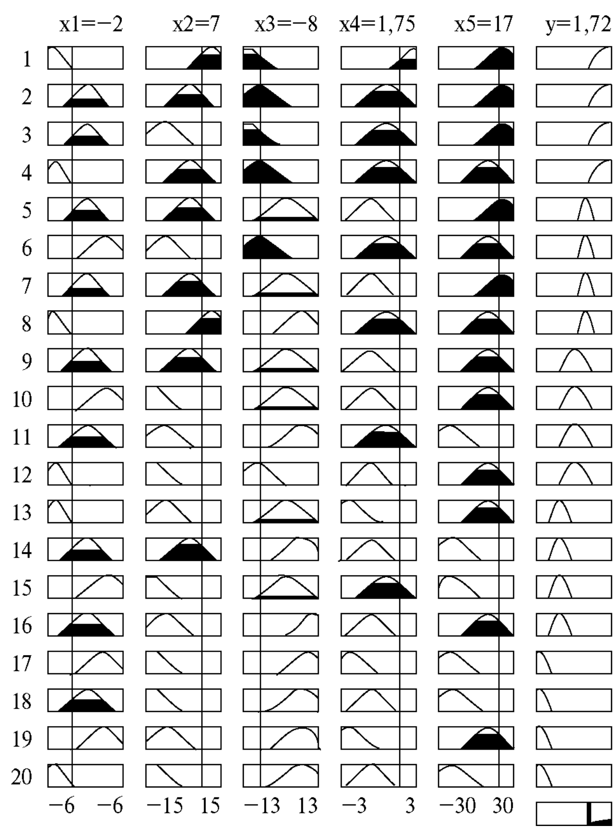

Рис. 2. Процесс прогнозирования результата матча «Шахтер — Металлург Д».

Шаг 6. По найденной разнице голов определить результат футбольного матча:

$$D = \begin{cases} d_1, \text{if} y \in (0.5, 3] \\ d_2, \text{if} y \in [-0.5, 0.5] \\ d_3, \text{if} y \in [-3, -0.5) \end{cases} .$$

В качестве примера рассмотрим матч «Шахтер — Металлург (Донецк)» (15 сентября 2001 г.), который закончился со счетом 3:1. Матчу соответствуют следующие значения влияющих факторов: $x_1 = -2$; $x_2 = 7$; $x_3 = -8$; $x_4 = 1.75$; $x_5 = 17$. Предложенная нечеткая модель прогнозирует разницу голов $y = 1.72$, что соответствует решению d_1 — домашняя победа. Рис. 2 иллюстрирует процедуру нечеткого логического вывода для этого матча с использованием инструментов **Fuzzy Logic Toolbox** программной среды **MatLab**.

3.4. Настройка нечеткой модели

В соответствии с [3, 4] настройка нечеткой модели состоит в нахождении таких ее параметров, которые минимизируют среднеквадратическое отклонение между экспериментальными данными из обучающей выборки и результатами моделирования. Настраиваемыми параметрами модели являются параметры функции принадлежности (3.1) для каждого лингвистического терма из базы знаний и веса правил $w_1 - w_{20}$. Общее количество настраиваемых параметров равно 64 (20 правил и 22 терма по два параметра функции принадлежности для каждого).

Настройка нечеткой модели была проведена на выборке из 105 матчей, сыгранных в двух чемпионатах. Оптимальные параметры нечеткой модели — функции принадлежности и веса правил приведены на рис. 3 и в табл. 3, соответственно.

4. Тестирование модели

Тестирование модели на выборке из 175 матчей показывает, что среднеквадратическое отклонение составляет 1.60, при этом статистическая оценка вероятности правильного предсказания исхода поединка (выигрыш, ничья, проигрыш) составляет 0.64, что в два раза лучше, чем при случайном угадывании.

Проиллюстрируем возможности предложенных моделей на примере прогнозирования результатов 9-го чемпионата Украины. Исходной информацией, необходимой для прогнозирования, служат результаты игр за 1-й — 8-й чемпионаты и результаты первых пяти туров 9-го чемпионата. На основе этой информации необходимо спрогнозировать результаты оставшихся 6, 7,. . . , 26 туров. В каждом туре происходит 7 матчей, поэтому необходимо спрогнозировать результаты 21×7=147 игр.

Прогнозирование будем осуществлять по следующей методике. Вначале спрогнозируем результаты 6-го тура. Затем с учетом полученных данных пересчитаем значение факторов $x_2 - x_5$ и спрогнозируем результаты 7-го тура, потом 8-го и т. д. по 26-й тур. Заметим, что значение фактора x_5 («встречи команд») необходимо пересчитывать начиная только со второго круга чемпионата, т. е. с 14-го тура. Значение фактора x_1 в i-м туре определим как $x_1(i) = x_1(6)/(i-5)$, где $x_1(6)$ — значение фактора x_1 в 6-м туре. Это позволит снизить влияние данного фактора при моделировании последних туров, так как на момент прогнозирования количество травмированных и дисквалифицированных игроков является величиной неизвестной.

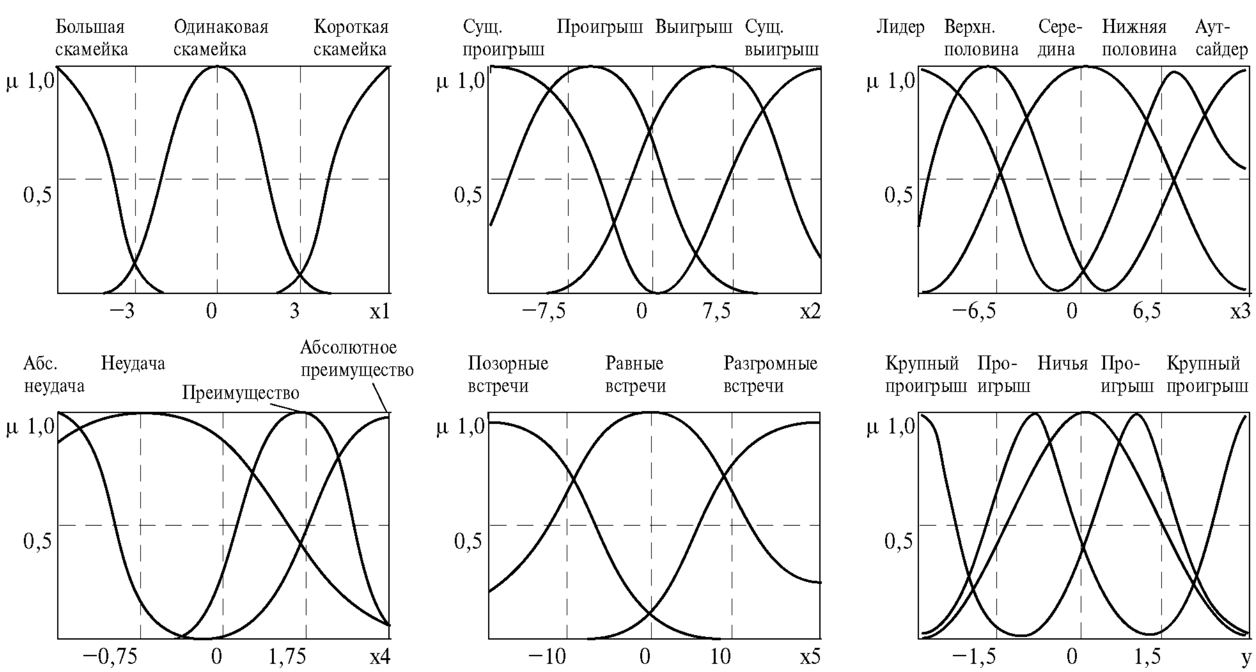

Рис. 3. Оптимальные функции принадлежности.

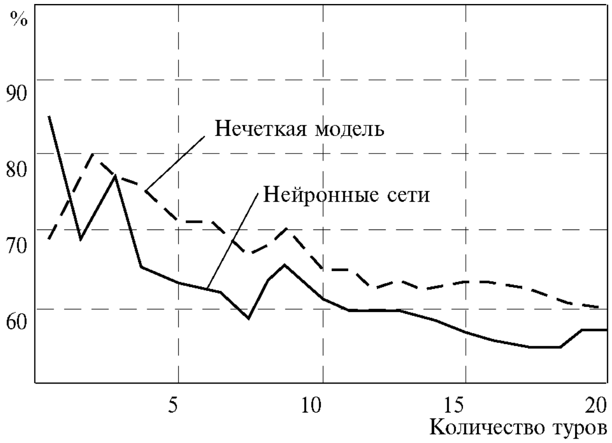

Рис. 4. Зависимость точности прогнозирования от количества туров.

Результаты прогнозирования в виде зависимости точности прогнозирования от количества туров и в виде итоговой таблицы чемпионата показаны на рис. 4 и в табл. 4, соответственно. Для сравнения здесь же приведены результаты прогнозирования, полученные с помощью нейронной сети, настроенной по той же обучающей выборке. В качестве нейронной сети использовался многослойный персептрон [2, 4] с двумя промежуточными слоями, состоящими из пяти нейронов. Как видно из таблицы, нечеткая модель достаточно точно определяет лидеров и аутсайдеров, т. е. наиболее важные результаты чемпионата — команды, занимающие призовые места и команды, покидающие высшую лигу.

Выводы

В работе предложен подход к прогнозированию результатов спортивных соревнований. В основу метода положена формализация экспертных лингвистических высказываний-правил в виде нечеткой базы знаний с последующей настройкой по экспериментальным данным. Приведенная в статье нечеткая база знаний позволяет прогнозировать исход футбольного матча на основе доступной информации о текущем состоянии команд и результатах предыдущих поединков.

Тестирование нечеткой базы знаний показывает, что результаты прогнозирования не хуже, чем при использовании нейронных сетей. При этом, в отличие от нейронных сетей, прогнозирующая модель является прозрачной с возможностью содержательного объяснения принятого решения. Предложенный подход может использоваться для предсказания результатов других спортивных турниров, а также для экстраполяции и прогнозирования результатов процессов в условиях ограниченного статистического материала в таких экономических задачах, как прогнозирование курсов валют и акций, прогнозирование объемов спроса и предложения товаров и услуг, оценка качества инновационных проектов и др.

Приложение. Основные положения теории нечетких множеств

Определение 1. *Нечетким множеством* $\tilde{A}$ на универсальном множестве U называется совокупность пар $(\mu_A(u), u)$, где $\mu_A(u)$ — степень принадлежности элемента $u \in U$ к нечеткому множеству $\tilde{A}$. Степень принадлежности — это число из диапазона [0, 1]. Чем выше степень принадлежности, тем в большей мере элемент универсального множества соответствует свойствам нечеткого множества.

Определение 2. *Функцией принадлежности* называется функция, которая позволяет вычислить степень принадлежности произвольного элемента универсального множества к нечеткому множеству.

Если универсальное множество состоит из конечного количества элементов $U = \{u_1, u_2, \dots, u_k\}$, тогда нечеткое множество $\tilde{A}$ записывается в виде:

$$\tilde{A} = \sum_{i=1}^{k} \mu_A(u_i)/u_i.$$

В случае непрерывного множества U используют следующее обозначение:

$$\tilde{A} = \int_U \mu_A(u)/u.$$

Примечание: знаки $\sum$ и $\int$ в этих формулах означают совокупность пар $\mu_A(u)$ и u.

Определение 3. *Лингвистической переменной* называется переменная, значениями которой могут быть слова или словосочетания некоторого естественного или искусственного языка.

Определение 4. *Терм-множеством* называется множество всех возможных значений лингвистической переменной.

Определение 5. *Термом* называется любой элемент терм-множества. В теории нечетких множеств терм формализуется нечетким множеством с помощью функции принадлежности.

Определение 6. *Дефаззификацией* называется процедура преобразования нечеткого множества в четкое число.

В теории нечетких множеств процедура дефаззификации аналогична нахождению характеристик положения (математического ожидания, моды, медианы) случайных величин в теории вероятностей. Простейшим способом выполнения процедуры дефаззификации является выбор четкого числа, соответствующего максимуму функции принадлежности. Однако пригодность этого способа ограничивается лишь одноэкстремальными функциями принадлежности. Для многоэкстремальных функций принадлежности наиболее часто используется дефаззификация путем нахождения центра тяжести плоской фигуры, ограниченной осями координат и функцией принадлежности.

Определение 7. *Нечеткой базой знаний* о влиянии факторов $X = \{x_1, x_2, \dots, x_n\}$ на значение параметра y называется совокупность логических высказываний типа:

$$\text{ЕСЛИ} \left(x_1 = a_1^{j1}\right) \text{ И } \left(x_2 = a_2^{j1}\right) \text{ И} \dots \text{И } \left(x_n = a_n^{j1}\right)$$

$$\text{ИЛИ} \left(x_1{=}a_1^{j2}\right) \text{ И } \left(x_2{=}a_2^{j2}\right) \text{ И} \dots \text{И } \left(x_n{=}a_n^{j2}\right) \dots$$

$$\text{ИЛИ} \left(x_1{=}a_1^{jk_j}\right) \text{ И } \left(x_2{=}a_2^{jk_j}\right) \text{ И} \dots \text{И } \left(x_n{=}a_n^{jk_j}\right),$$

$$\text{ТО} \quad y = d_j, \quad \text{для всех} \quad j = \overline{1, m},$$

Таблица 3. Оптимальные веса правил

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Вес	0.5	0.95	0.5	0.63	0.5	0.75	0.5	1	0	0.22	0.1	0.08	0.01	0.29	0.3	1	1	1	1	1

Таблица 4. Результирующая турнирная таблица

Команда	Результаты чемпионата		Прогноз нечеткой модели		Прогноз нейронной сети	
	очки	место	очки	место	очки	место
Динамо	58	1	59	1–2	63	1–2
Шахтер	57	2	59	1–2	54	4
Днепр	52	3	57	3	57	3
Металлург Д	51	4	53	4	63	1–2
Металлург М	37	5	33	8	35	8
ЦСКА	36	6	26	10	39	6
Металлург Зп	32	7	41	7	27	10
Таврия	30	8	49	5	37	7
Карпаты	27	9	47	6	44	5
Металлист	25	10	22	11	34	9
Кривбасс	24	11	27	9	22	11
Ворскла	20	12	3	14	3	14
Сталь	11	13	21	12	10	13
Нива	9	14	11	13	13	12

где a_i^{jp} — нечеткий терм, которым оценивается переменная x_i в строчке с номером jp $(p=\overline{1,k_j})$; k_j — количество строчек-конъюнкций, в которых выход y оценивается нечетким термом d_j, $j=\overline{1,m}$; m — количество термов, используемых для лингвистической оценки выходного параметра y.

С помощью операций $\cup$ (ИЛИ) и $\cap$ (И) нечеткую базу знаний из определения 7 перепишем в более компактном виде:

$$\cup_{p=1}^{k_j}\left[\cap_{i=1}^{n}\left(x_i=a_i^{jp}\right)\right]\to y=d_j,\quad j=\overline{1,m}. \qquad (\text{П}1)$$

Определение 8. *Нечетким логическим выводом* называется аппроксимация зависимости $y = f(x_1, x_2, \ldots, x_n)$ с помощью нечеткой базы знаний и операций над нечеткими множествами.

Пусть $\mu^{jp}(x_i)$ — функция принадлежности входа x_i нечеткому терму a_i^{jp}, $i=\overline{1,n}$, $j=\overline{1,m}$, $p=\overline{1,k_j}$, т. е. $a_i^{jp}=\int\limits_{\underline{x}_i}^{\overline{x}_i}\mu^{jp}(x_i)/x_i$; $\mu^{d_j}(y)$ — функция принадлежности выхода y нечеткому терму d_j, $j=\overline{1,m}$, т. е. $d_j=\int\limits_{\underline{y}}^{\overline{y}}\mu^{d_j}(y)/y$. Тогда степень принадлежности конкретного входного вектора $X^*=\{x_1^*,x_2^*,\ldots,x_n^*\}$ нечетким термам d_j из базы знаний (П1) определяется следующей системой нечетких логических уравнений:

$$\mu^{d_j}(X^*)=\underset{p=\overline{1,k_j}}{\vee}\ \underset{i=\overline{1,n}}{\wedge}\left[\mu^{jp}(x_i^*)\right],\quad j=\overline{1,m},$$

где $\vee(\wedge)$ — операция максимума (минимума).

Нечеткое множество $\tilde{y}$, соответствующее входному вектору X^*, определяется следующим образом:

$$\tilde{y}=\bigcup_{j=\overline{1,m}}\int\limits_{\underline{y}}^{\overline{y}}\min\left(\mu^{d_j}(X^*),\mu^{d_j}(y)\right)/y,$$

где $\bigcup$ — операция объединения нечетких множеств.

Четкое значение выхода y, соответствующее входному вектору X^*, определяется в результате дефаззификации нечеткого $\tilde{y}$.

Литература

1. Прикладные нечеткие системы: Пер. с япон. / К. Асан, Д. Ватада, С. Иван и др. Под ред. Т. Тэрано, К. Асан, М. Сугэно. М.: Мир, 1993.
2. *Ротштейн А. П.* Интеллектуальные технологии идентификации. Винница: Вiнниця-УНIВЕРСУМ, 1999.
3. *Ротштейн А. П., Кательников Д. И.* Идентификация нелинейных зависимостей нечеткими базами знаний // Кибернетика и системный анализ. 1998. № 5. С. 53–61.
4. *Looney C.* Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists. Oxford University Press, 1997.
5. *Zimmerman H.-J.* Fuzzy Set Theory and Its Applications / 3rd ed. Kluwer Academic Publisher, 1996.

■ Поступила в редакцию 9 мая 2002 года; в окончательном варианте 9 июня 2002 года.