

¹Д-р техн. наук, професор, професор кафедри комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця, Україна

²Асистент кафедри комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця, Україна

КРИТЕРІЇ НАВЧАННЯ НЕЧІТКОГО КЛАСИФІКАТОРА НА ОСНОВІ ВІДСТАНИ МІЖ ГОЛОВНИМИ КОНКУРЕНТАМИ

Класифікація це віднесення об'єкта за деякими ознаками до одного з класів. До класифікації зводяться різноманітні задачі прийняття рішень в інженерії, економіці, медицині, соціології та в інших областях. В нечітких класифікаторах залежність «входи – вихід» описуються за допомогою лінгвістичних правил <Якщо – тоді>, антецеденти яких містять нечіткі терми «низький», «середній», «високий» тощо. Для підвищення безпомилковості нечіткий класифікатор навчають за експериментальними даними. В даній роботі запропоновано нові критерії навчання нечіткого класифікатора, які враховують різницю належностей нечіткого висновку лише до головних конкурентів. За правильної класифікації головним конкурентом прийнятого рішення є клас, що має другий за величиною ступінь належності. У випадку неправильної класифікації помилково прийняте рішення є головним конкурентом правильного класу. Проведені комп'ютерні експерименти із навчання нечіткого класифікатора для розпізнавання трьох сортів італійських вин засвідчили суттєву перевагу нових критеріїв. Серед нових критеріїв помірну перевагу має критерій на основі квадратичної відстані між головними конкурентами з штрафом за помилкове рішення. Нові критерії можуть застосовуватися не лише для навчання нечітких класифікаторів, але і для навчання деяких інших моделей, наприклад, нейронних мереж.

Ключові слова: класифікація, нечітка база знань, навчання, критерії навчання, головні конкуренти.

НОМЕНКЛАТУРА

\tilde{a}_{ij} – нечіткий терм, яким оцінюється ознака x_i в j -му правилі, $i = \overline{1, n}$, $j = \overline{1, k}$;

$\Delta_r(\mathbf{K})$ – мітка коректності класифікації r -го об'єкту за нечіткою моделлю з параметрами \mathbf{K} ;

$\mu_j(X^*)$ – ступінь виконання j -го правила з бази знань;

$\mu_{l_s}(y^*)$ – ступінь належності вхідного вектору \mathbf{X}^* до класу l_s ;

$\mu_j(x_i^*)$ – ступінь належності значення x_i^* нечіткому терму \tilde{a}_{ij} ;

$\mu_{l_s}(y_r)$ – ступінь належності r -го об'єкту навчальної вибірки до класу l_s ;

$\mu_{l_s}(\mathbf{K}, \mathbf{X}_r)$ – ступінь належності висновку до класу l_s , який розраховано за нечіткою моделлю з параметрами \mathbf{K} для вхідного вектора \mathbf{X}_r ;

$\mu_{win}(\mathbf{X}_r)$ – ступінь належності вхідного вектора \mathbf{X}_r класу переможцю;

$\mu_{vicewin}(\mathbf{X}_r)$ – ступінь належності вхідного вектора \mathbf{X}_r до класу з другим рангом;

$\mu(x)$ – функція належності для змінної x ;

\wedge – t -норма;

b – координата максимуму гаусової функції належності;

c – коефіцієнт концентрації гаусової функції належності.

$Crit_1$ – перший критерій навчання;

$Crit_2$ – другий критерій навчання;

$Crit_3$ – третій критерій навчання;

$Crit_4$ – четвертий критерій навчання;

$Crit_5$ – п'ятий критерій навчання;

$D_r(\mathbf{K})$ – відстань між бажаною та дійсною вихідними нечіткими множинами при класифікації r -го об'єкту на основі нечіткої бази знань з параметрами;

d_j – категоріальне значення консеквента j -го правила;

F – нечітка модель;

k – кількість правил;

\mathbf{K} – вектор параметрів нечіткої бази знань, які налаштовуються;

l_m – мітка класу з номером m ;

M – обсяг навчальної вибірки;

p – штрафний коефіцієнт;

\mathbf{P} – вектор параметрів функцій належності термів бази знань;

smax – операція знаходження другого за величиною елемента множини;

T – обсяг тестової вибірки;

$Vicewin$ – клас з другим за величиною ступенем належності;

\mathbf{W} – вектор вагових коефіцієнтів правил нечіткої бази знань;

win – клас-переможець з максимальним ступенем належності;

w_j – ваговий коефіцієнт j -го правила;

\mathbf{X} – вектор вхідних атрибутів;

\mathbf{X}_r – вхідні атрибути r -го об'єкту;

y – результат класифікації;

y_r – клас r -го об'єкту.

ВСТУП

Класифікація це віднесення об'єкта за деякими ознаками до одного з класів. До класифікації зводяться різноманітні задачі прийняття рішень в інженерії, економіці, медицині, соціології, політиці, спорті та в інших областях.

Останнім часом все більш популярними стають нечіткі класифікатори, тобто класифікатори, в процесі функціонування або навчання яких використовуються нечіткі множини [1]. Сьогодні переважно застосовуються нечіткі класифікатори на основі логічного виведення по базі продукційних правил, антецеденти яких містять нечіткі терми «низький», «середній», «високий» тощо. Кожне правило задає область вхідних атрибутів, в межах якої об'єкти належать одному класу. Границі цих областей нечіткі, тому один і той же об'єкт може одночасно належати декільком класам, але з різним ступенем.

Для підвищення безпомилковості нечітких класифікатор навчають по експериментальним даним. Для цього змінюють його параметри, щоб мінімізувати відстань між експериментальними даними та результатами нечіткого виведення. Цю відстань, яку назвемо критерієм навчання, можна визначити у різний спосіб [2]. Критерії навчання впливають на зміну параметрів нечіткого класифікатора на кожній ітерації алгоритму налаштування і, відповідно призводять до різних результатів. Тому, метою статті є знаходження такого критерія, використання якого забезпечує найкращу результативність навчання нечіткого класифікатора.

1 ПОСТАНОВКА ЗАДАЧІ

Нечіткий класифікатор являє собою відображення $\mathbf{X} = (x_1, x_2, \dots, x_n) \xrightarrow{F} y \in \{l_1, l_2, \dots, l_m\}$, яке реалізується логічним виведенням по базі нечітких правил. Навчання нечіткої бази знань здійснимо за експериментальними даними. Відповідно до принципу зовнішнього доповнення [3] експериментальні дані розіб'ємо на навчальну вибірку з M рядків «входи – вихід» та тестову вибірку з T таких рядків:

$$(\mathbf{X}_r, y_r), r = \overline{1, M}, \quad (1)$$

$$(\mathbf{X}_r, y_r), r = \overline{1, T} \quad r = \overline{1, T}, \quad (2)$$

де $\mathbf{X}_r = (x_{r1}, x_{r2}, \dots, x_{rn})$; $y_r \in \{l_1, l_2, \dots, l_m\}$.

Задача полягає у знаходженні такого критерію навчання *Crit*, застосування якого під час налаштування нечіткої бази знань F на вибірці (1) забезпечує мінімальну частоту помилок класифікації на тестовій вибірці (2).

2 ОГЛЯД ЛІТЕРАТУРИ

Класифікація на основі нечітких множин була запропонована 50 років тому в статті [4]. Перші роботи із ідентифікації залежностей за допомогою нечітких класифікаторів з'явилися в середині 90-х років. Вперше задача структурної ідентифікації шляхом відбору правил нечіткого класифікатора за показниками безпомилковості та складності бази знань розглянута в роботі [5]. Пізніше в [6] на основі критерію безпомилковості здійснили і параметричну ідентифікацію шляхом налаштування ваг правил. При зміні лише ваги правил можна отримати швидкі алгоритми навчання [7], але вони не гарантують високої безпомилковості нечіткого класифікатора. Для підвищення безпомилковості нечіткого класифікатора на етапі параметричної ідентифікації налаштовують не лише ваги правил, а й функції належності. Першими роботами з налашту-

вання функцій належності нечіткого класифікатора є статті [8, 9, 10]. В них критерієм навчання виступає квадратична нев'язка між двома нечіткими множинами – бажаними та реальними результатами класифікації. В сучасних роботах навчання нечітких класифікаторів здійснюють саме за цим критерієм (дивись, наприклад, [11]).

В роботі [7] запропоновано новий критерій навчання, який поєднує частоту помилок та квадратичну нев'язку між нечіткими бажаними та дійсними результатами логічного виведення. У випадку помилкової класифікації квадратична нев'язка зважується штрафним коефіцієнтом. Комп'ютерні експерименти в роботах [2, 7] показали перевагу цього критерію навчання над частотою помилок та над квадратичною нев'язкою. Але ця перевага є незначною і навчання нечітких класифікаторів не завжди є результативним. Тому виникає зацікавленість у нових критеріях, навчання за якими забезпечує кращу безпомилковість нечіткого класифікатора.

Зміна параметрів нечіткого класифікатора може відбуватися за логічним висновком для одного об'єкта навчальної вибірки або за результатами логічного виведення за усією навчальною вибіркою [6]. Ми розглядаємо пакетний режим навчання, коли параметри класифікатора модифікуються за логічними висновками за усією вибіркою (1).

Грунтуючись на [2] базу правил нечіткого класифікатора запишемо так:

$$\text{Якщо } (x_1 = \tilde{a}_{1j} \text{ та } x_2 = \tilde{a}_{2j} \text{ та } \dots \text{ та } x_n = \tilde{a}_{nj} \text{ з вагою } w_j),$$

$$\text{тоді } y = d_j, \quad j = \overline{1, k}, \quad (3)$$

де $w_j \in [0, 1]$, $j = \overline{1, k}$; $d_j \in \{l_1, l_2, \dots, l_m\}$.

Класифікація об'єкта з атрибутами $\mathbf{X}^* = (x_1^*, x_2^*, \dots, x_n^*)$ здійснюється таким чином. Спочатку розраховується ступінь виконання j -го правила з бази (3):

$$\mu_j(\mathbf{X}^*) = w_j \cdot (\mu_j(x_1^*) \wedge \mu_j(x_2^*) \wedge \dots \wedge \mu_j(x_n^*)),$$

$$j = \overline{1, k}. \quad (4)$$

Ступінь належності \mathbf{X}^* до класів l_1, l_2, \dots, l_m розраховується так:

$$\mu_s(y^*) = \max_{\forall j: d_j=l_s} (\mu_j(\mathbf{X}^*)), \quad s = \overline{1, m}. \quad (5)$$

Нечітким рішенням задачі класифікації буде нечітка множина

$$\tilde{y}^* = \left(\frac{\mu_{l_1}(y^*)}{l_1}, \frac{\mu_{l_2}(y^*)}{l_2}, \dots, \frac{\mu_{l_m}(y^*)}{l_m} \right). \quad (6)$$

Кінцевим результатом виведення оберемо ядро нечіткої множини (6), тобто клас з максимальним ступенем належності:

$$y^* = \arg \max_{\{l_1, l_2, \dots, l_m\}} (\mu_s(y^*)).$$

Навчання нечіткого класифікатора полягає в знаходженні такого вектора $\mathbf{K} = (\mathbf{P}, \mathbf{W})$, який мінімізує частоту помилок класифікації на тестовій вибірці. При цьому для зміни координат вектора \mathbf{K} використовується лише навчальна вибірка (1). Координати вектора \mathbf{K} змінюємо на кожній ітерації алгоритму оптимізації за відстанню між результатами логічного виведення та експериментальними даними з вибірки (1). Цю відстань, яку назвемо критерієм навчання, можна визначити у різний спосіб.

Сьогодні найвідомішими є 3 критерії навчання нечіткого класифікатора: на основі частоти помилок [6], на основі відстані між нечіткими результатами класифікації і експериментальними даними [8–11], та їх комбінації [2, 7].

Критерій 1 – частота помилок класифікації:

$$Crit_1 = \frac{1}{M} \sum_{r=1, M} \Delta_r(\mathbf{K}), \quad (7)$$

$$\text{де } \Delta_r(\mathbf{K}) = \begin{cases} 1, & \text{якщо } y_r \neq F(\mathbf{K}, \mathbf{X}_r) \\ 0, & \text{якщо } y_r = F(\mathbf{K}, \mathbf{X}_r) \end{cases}$$

Переваги критерію полягають в його простоті та ясній змістовній інтерпретації. Але цільова функція в задачі оптимізації за цим критерієм приймає дискретні значення, що ускладнює застосування швидких градієнтних методів оптимізації, особливо за малих вибірок даних.

Критерій 2 – квадратична нев'язка між двома нечіткими множинами – бажаними та реальними результатами класифікації. Для її розрахунку значення вихідної змінної y в навчальній вибірці фаззифікують таким чином:

$$\left. \begin{aligned} \tilde{y} &= \left(\frac{1}{l_1}, \frac{0}{l_2}, \dots, \frac{0}{l_m} \right), & \text{якщо } y = l_1 \\ \tilde{y} &= \left(\frac{0}{l_1}, \frac{1}{l_2}, \dots, \frac{0}{l_m} \right), & \text{якщо } y = l_2 \\ &\vdots \\ \tilde{y} &= \left(\frac{0}{l_1}, \frac{0}{l_2}, \dots, \frac{1}{l_m} \right), & \text{якщо } y = l_m \end{aligned} \right\} \quad (8)$$

Критерій навчання враховує відстань між логічним висновком у формі нечіткої множини (6) та бажаним нечітким значенням вихідної змінної (8):

$$Crit_2 = \sum_{r=1, M} D_r(\mathbf{K}). \quad (9)$$

Для розрахунку $D_r(\mathbf{K})$ використовується евклідова метрика:

$$D_r(\mathbf{K}) = \sum_{s=1, m} (\mu_{l_s}(y_r) - \mu_{l_s}(\mathbf{K}, \mathbf{X}_r))^2, \quad (10)$$

Перевага критерію $Crit_2$ полягає в урахуванні міри впевненості в прийнятому рішенні на основі ступенів належності об'єкту різним класам. В критерії $Crit_1$ вважається, що результат класифікації об'єкту є абсолютно достовірним, тобто неважливо наскільки ступінь належ-

ності у рішення більший, ніж у інших альтернатив – на 0,0001 чи на 1. Крім того, цільова функція в задачі навчання за критерієм (9) не має довгих плато, тому вона придатна до оптимізації градієнтними методами. Але близькі до границь класів об'єкти вносять майже однаковий вклад в критерій навчання (9) як за правильної, так і за помилкової класифікації, тому навчання може бути нерезультативним.

Критерій 3 – квадратична нев'язка між нечіткими бажаними та реальними результатами класифікації з додатковим штрафом за помилкове рішення. Ідея полягає в збільшенні відстані D для помилково класифікованих об'єктів:

$$Crit_3 = \sum_{r=1, M} (\Delta_r(\mathbf{K}) \cdot p + 1) \cdot D_r(\mathbf{K}), \quad (11)$$

де $p > 0$.

3 МАТЕРІАЛИ І МЕТОДИ

Нижче пропонуються нові критерії навчання нечіткого класифікатора, які враховують різницю належностей нечіткого висновку лише до головних конкурентів.

Критерій 4 – відстань між головними конкурентами з штрафом за помилкове рішення. Ідея цього нового критерію полягає у врахуванні різниці належностей нечіткого висновку лише до головних конкурентів. За алгоритмом виведення рішенням обирається клас з максимальним ступенем належності. Позначимо цей клас-переможець через win . У випадку правильної класифікації головним конкурентом прийнятого рішення є $vicewin$ – клас з другим за величиною ступенем належності. Чим більша різниця між належністю до класів win та $vicewin$, тим більша впевненість у логічному висновку, і тим далі об'єкт знаходиться від границі розділу класів.

Для r -го об'єкту з вибірки (1) $\mu_{win}(\mathbf{X}_r) = \max_{s=1, m} (\mu_{l_s}(\mathbf{X}_r))$ та $\mu_{vicewin}(\mathbf{X}_r) = \text{smax}_{s=1, m} (\mu_{l_s}(\mathbf{X}_r))$. Відповідно, різниця між головними конкурентами дорівнює $\mu_{win}(\mathbf{X}_r) - \mu_{vicewin}(\mathbf{X}_r)$.

В критерії навчання враховуватимемо відносні показники, розділивши різницю на ступінь належності класу-переможцю. За правильної класифікації відносна різниця дорівнює $D_r^1 = \frac{\mu_{win}(\mathbf{X}_r) - \mu_{vicewin}(\mathbf{X}_r)}{\mu_{win}(\mathbf{X}_r)}$, а за непра-

вильної – $D_r^0 = \frac{\mu_{win}(\mathbf{X}_r) - \mu_{y_r}(\mathbf{X}_r)}{\mu_{win}(\mathbf{X}_r)}$. Крім того,

аналогічно до (11), за помилкової класифікації зважимо різницю штрафним коефіцієнтом $p \geq 1$. Математично критерій навчання запишемо таким чином:

$$Crit_4 = p \cdot \sum_{\substack{y_r \neq F(\mathbf{K}, \mathbf{X}_r) \\ r=1, M}} D_r^0(\mathbf{K}) - \sum_{\substack{y_r = F(\mathbf{K}, \mathbf{X}_r) \\ r=1, M}} D_r^1(\mathbf{K}). \quad (12)$$

Критерій 5 – квадратична відстань між головними конкурентами з штрафом за помилкове рішення. Цей

критерій є модифікацією попереднього. Відмінність полягає у використанні не відносних різниць, а їх квадратів:

$$Crit_5 = p \cdot \sum_{\substack{y_r \neq F(\mathbf{K}, \mathbf{X}_r) \\ r=1, M}} D_r^0(\mathbf{K})^2 - \sum_{\substack{y_r = F(\mathbf{K}, \mathbf{X}_r) \\ r=1, M}} D_r^1(\mathbf{K})^2.$$

Піднесення до квадрату в $Crit_5$ дозволяє, як і в методі найменших квадратів, збільшити важливість великих різниць – тобто грубих промахів.

4 ЕКСПЕРИМЕНТИ

Метою експериментів є виявлення критерію навчання, який забезпечує найкращу безпомилковість. Розглядається тестова задача Wine Dataset з UCI Machine Learning Repository. Вона полягає у виявленні сорту винограду (y), з якого виготовлено вино. База даних містить результати лабораторних аналізів по 13-ти показникам 178 зразків італійських вин, виготовлених в одному регіоні. Для кожного зразка вказано, з якого із з трьох сортів винограду виготовлено вино.

Навчальну вибірку сформуємо з рядків бази даних з граничними значеннями кожного із 13 атрибутів. Додатково в навчальну вибірку включимо всі непарні рядки бази даних. Всі інші дані занесемо в тестову вибірку. В результаті отримаємо навчальну вибірку з 100 рядків і тестову – з 78. Побудуємо нечіткий класифікатор вин за трьома ознаками: x_7 – flavanoids, x_{10} – color intensity та x_{13} – proline. Експерименти проведемо для нечіткого класифікатора з базою знань з табл. 1. Нечіткі терми задамо гаусовою функцією належності:

$$\mu(x) = \exp\left(-\frac{(x-b)^2}{2c^2}\right),$$

де $c > 0$.

Параметри функцій належності початкового нечіткого класифікатора наведено в табл. 2.

Таблиця 1 – Нечітка база знань класифікатора вин [6]

№	x_7	x_{10}	x_{13}	y
1	–	–	Високий	Сорт 1
2	Високий	Високий	Середній	Сорт 1
3	–	Низький	Низький	Сорт 2
4	Низький	Низький	Середній	Сорт 2
5	Низький	Високий	–	Сорт 3

Таблиця 3 – Вплив штрафного коефіцієнта в критерії навчання на безпомилковість класифікатора на тестовій вибірці (статистика 200 експериментів)

t -норма	Критерій	Середня безпомилковість				
		$p = 1$	$p = 3$	$p = 5$	$p = 7$	$p = 9$
min	$Crit_3$	0,0856	0,0915	0,0912	0,0910	0,0960
	$Crit_4$	0,0732	0,0640	0,0662	0,0736	0,0770
	$Crit_5$	0,0690	0,0604	0,0602	0,0688	0,0657
prod	$Crit_3$	0,0727	0,0656	0,0665	0,0681	0,0649
	$Crit_4$	0,0565	0,0512	0,0493	0,0499	0,0527
	$Crit_5$	0,0448	0,0443	0,0456	0,0486	0,0446

Таблиця 2 – Початкові параметри функцій належності термів нечіткого класифікатора

Вхідна змінна	Терм	Параметри	
		b	c
x_7	Низький	2	0,34
	Високий	2	5,08
x_{10}	Низький	6	1,28
	Високий	6	13
x_{13}	Низький	3	2,78
	Середній	3	10
	Високий	3	16,8

Для кожного критерію проведемо 1000 експериментів із навчання нечіткої бази знань на основі квазіньютонівського алгоритму. Після навчання кожний класифікатор перевіримо на тестовій вибірці за частотою помилок (критерій $Crit_1$). Під час навчання налаштуємо вагові коефіцієнти перших чотирьох правил. Достовірність п'ятого правила не викликає сумнівів, тому його ваговий коефіцієнт залишимо рівним 1. Налаштуємо також коефіцієнти концентрації (c) функції належності кожного нечіткого терма. Для збереження інтерпретованості бази знань згідно до [11] налаштуємо координати максимумів (b) функцій належності лише не крайніх термів. В базі знань 1 некрайній терм – «Середній», координату максимуму якого і будемо змінювати. Таким чином, загальна кількість параметрів для налаштування становить $4 + 7 + 1 = 12$. Початкові точки для навчання оберемо випадково – для вагових коефіцієнтів правил з діапазону $[0, 1]$, а для параметрів функцій належності в межах $\pm 30\%$ від значень з табл. 2.

Проведемо 2 серії експериментів. Першу серію для нечіткого класифікатора з реалізацією t -норми мінімумом (min), а другу – з реалізацією t -норми добутком (prod). В експериментах з критеріями $Crit_3$, $Crit_4$ та $Crit_5$ спочатку визначимо прийнятний рівень штрафного коефіцієнту. Для цього проведемо по 200 оціночних експериментів для $p = 1, 3, \dots, 9$. Результати цих експериментів наведено в табл. 3. В ній напівжирним шрифтом виділено за яких значень штрафного коефіцієнту навчання відбувається краще. Саме за таких значень штрафного коефіцієнту проведемо решту експериментів.

5 РЕЗУЛЬТАТИ

Результати експериментів наведено на рис. 1–3 та в табл. 4.

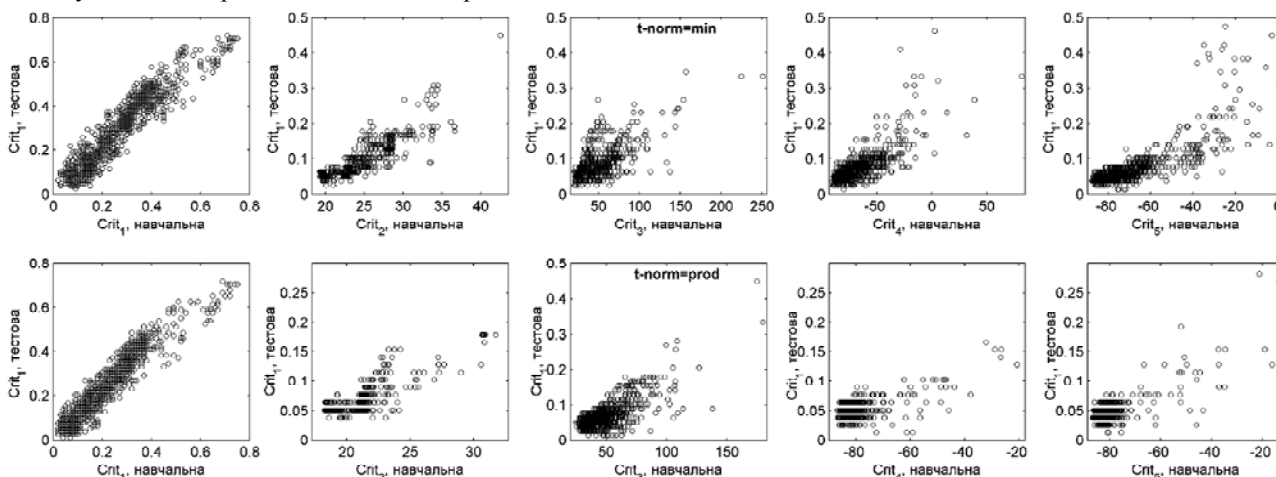


Рисунок 1 – Розподіл результатів навчання нечіткого класифікатора (статистика 1000 експериментів)

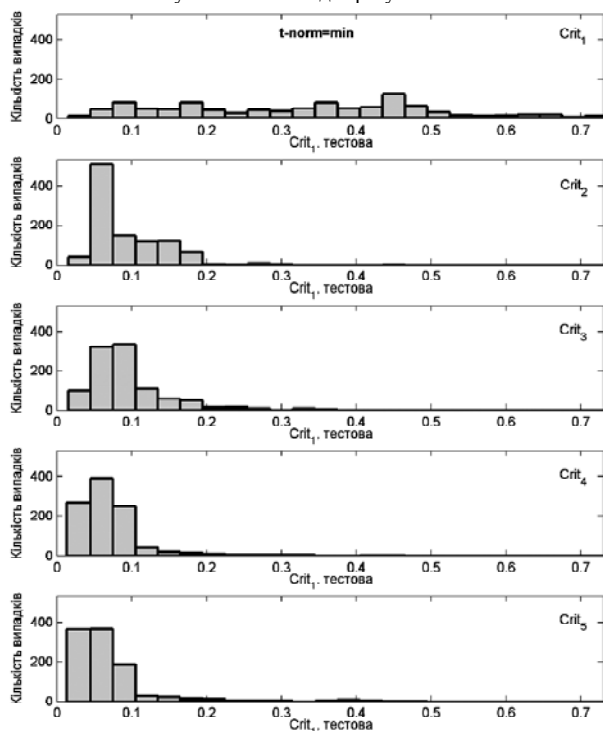


Рисунок 2 – Розподіл якості навчання класифікатора, *t*-норму якого реалізовано мінімумом

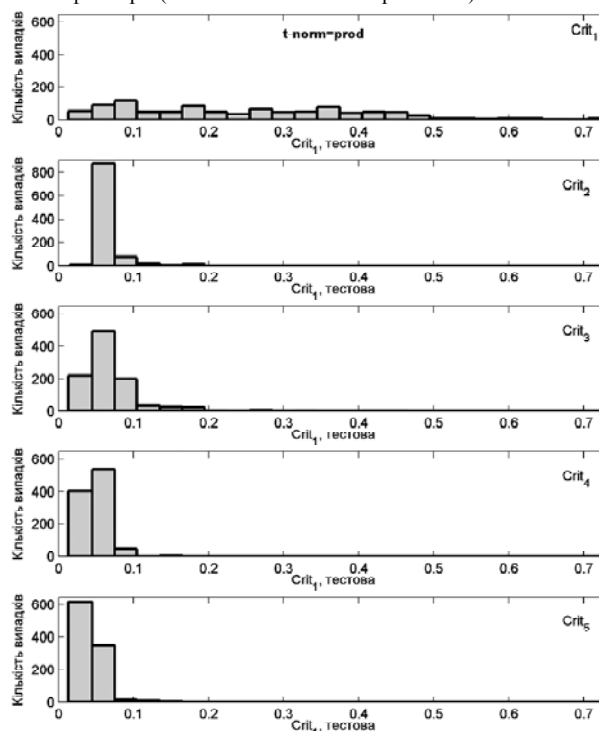


Рисунок 3 – Розподіл якості навчання класифікатора, *t*-норму якого реалізовано добутком

Таблиця 4 – Статистика навчання нечітких класифікаторів (напівжирним виділено найкращі результати)

<i>t</i> -норма	Критерій навчання	Частота помилок (<i>Crit_i</i>) на тестовій вибірці			
		мінімальне	середнє	медіанне	максимальне
min	<i>Crit₁</i>	0,0256	0,3194	0,3333	0,7179
	<i>Crit₂</i>	0,0256	0,0900	0,0641	0,4487
	<i>Crit₃</i>	0,0256	0,0921	0,0769	0,3462
	<i>Crit₄</i>	0,0128	0,0680	0,0513	0,4615
	<i>Crit₅</i>	0,0128	0,0680	0,0513	0,4744
prod	<i>Crit₁</i>	0,0128	0,2459	0,2308	0,7179
	<i>Crit₂</i>	0,0385	0,0631	0,0641	0,1795
	<i>Crit₃</i>	0,0128	0,0663	0,0641	0,4487
	<i>Crit₄</i>	0,0128	0,0503	0,0513	0,1667
	<i>Crit₅</i>	0,0128	0,0454	0,0385	0,2821

6 ОБГОВОРЕННЯ

Результати експериментів на рис. 1 вказують на корельованість значень критеріїв $Crit_1 - Crit_5$ на навчальній вибірці з частотою помилок на тестовій вибірці. Відповідно ці критерії можна застосовувати для навчання нечіткого класифікатора. Щодо результативності навчання (див. табл. 4 та рис. 2, 3), то вона суттєво краща при використанні нових критеріїв $Crit_4$ та $Crit_5$. Нові критерії забезпечують кращу безпомилковість як в середньому, так і за кількістю найкращих випадків навчання. Серед нових критеріїв помірну перевагу має $Crit_5$.

ВИСНОВКИ

У роботі запропоновано нові критерії навчання нечіткого класифікатора, які враховують різницю належностей нечіткого висновку лише до головних конкурентів. За правильної класифікації головним конкурентом прийнятого рішення є клас з другим за величиною ступенем належності. У випадку неправильної класифікації помилково прийняте рішення є головним конкурентом правильного класу.

Проведені нами експерименти із навчання нечіткого класифікатора для UCI-задачі із розпізнавання італійських вин засвідчили суттєву перевагу нових критеріїв. Серед нових критеріїв помірну перевагу має критерій на основі квадратичної відстані між головними конкурентами з штрафом за помилкове рішення. Відповідно, мету дослідження досягнуто – встановлено, що для налаштування нечітких класифікаторів доцільно використовувати нові критерії навчання на основі різниці належності рішення до головних конкурентів. Нові критерії можуть застосовуватися не лише для настроювання нечітких класифікаторів, але і для навчання деяких інших моделей, наприклад, нейронних мереж.

ПОДЯКИ

Публікація містить результати досліджень, проведених при грантовій підтримці Державного фонду фундаментальних досліджень за конкурсним проектом №62/201–2015.

Штовба С. Д.¹, Галушак А. В.²

¹Д-р техн. наук, професор, професор кафедри комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця, Україна

²Асистент кафедри комп'ютерних систем управління, Вінницький національний технічний університет, Вінниця, Україна

ОБУЧЕНИЕ НЕЧЕТКОГО КЛАССИФИКАТОРА НА ОСНОВЕ РАССТОЯНИЯ МЕЖДУ ГЛАВНЫМИ КОНКУРЕНТАМИ

Классификация – это отнесение объекта по некоторым признакам к одному из классов. К классификации сводятся разнообразные задачи принятия решений в инженерии, экономике, медицине, социологии и других областях. В нечетких классификаторах зависимость «входы – выход» описываются с помощью лингвистических правил «Если – то», antecedentes которых содержат нечеткие термины «низкий», «средний», «высокий» и т. п. Для повышения безошибочности нечеткий классификатор обучают по экспериментальным данным. В данной работе предложены новые критерии обучения нечеткого классификатора, которые учитывают разницы принадлежностей нечеткого вывода только к главным конкурентам. При правильной классификации главным конкурентом принятого решения является класс со второй по величине степенью принадлежности. В случае неправильной классификации ошибочно принятое решение является главным конкурентом правильного класса. Компьютерные эксперименты по обучению нечеткого классификатора для распознавания трех сортов итальянских вин показали существенное преимущество новых критериев. Среди новых критериев обучения небольшое преимущество имеет критерий в форме квадратичного расстояния между главными конкурентами со штрафом за ошибочное решение. Новые критерии могут применяться не только для обучения нечетких классификаторов, но и для обучения других моделей, например, нейронных сетей.

Ключевые слова: классификация, нечеткая база знаний, обучение, критерии обучения, главные конкуренты.

СПИСОК ЛІТЕРАТУРИ

1. Kuncheva L. I. Fuzzy classifier design / L. I. Kuncheva // Studies in Fuzziness and Soft Computing. – Berlin – Heidelberg: Springer-Verlag, 2000. – Vol. 49. – 314 p.
2. Shtovba S. Analyzing the criteria for fuzzy classifier learning / S. Shtovba, O. Pankevich, A. Nagorna // Automatic Control and Computer Sciences. – 2015. – Vol. 49, № 3. – P. 123–132.
3. Madala H. R. Inductive learning algorithms for complex systems modeling / H. R. Madala, A. G. Ivakhnenko. – Boca Raton : CRC Press, 1994. – 368 p.
4. Bellman R. Abstraction and pattern classification / R. Bellman, R. Kalaba, L. Zadeh // Journal of Mathematical Analysis and Applications. – 1966. – Vol. 13, № 1. – P. 1–7.
5. Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms / [Ishibuchi H., Nozaki K., Yamamoto N., Tanaka H.] // Fuzzy sets and systems. – 1994. – Vol. 65, № 2. – P. 237–253.
6. Ishibuchi H. Classification and modeling with linguistic information granules: advanced approaches / H. Ishibuchi, T. Nakashima, M. Nii. – Berlin-Heidelberg : Springer-Verlag, 2005. – 307 p.
7. Штовба С. Д. Порівняння критеріїв навчання нечіткого класифікатора / С. Д. Штовба // Вісник Вінницького політехнічного інституту. – 2007. – № 6. – С. 84–91.
8. Abe S. Tuning of a fuzzy classifier derived from data / S. Abe, M. S. Lan, R. Thawonmas // International Journal of Approximate Reasoning. – 1996. – Vol. 14. – P. 1–24.
9. Nauck D. A neuro-fuzzy method to learn fuzzy classification rules from data / D. Nauck, R. Kruse // Fuzzy Sets and Systems. – 1997. – Vol. 89, № 3. – P. 277–288.
10. Rotshtein A. P. Design and Tuning of Fuzzy If – Then Rules for Automatic Classification / A. P. Rotshtein, D. I. Katelnikov // Proc. of NAFIPS'98 – International Conf. «Annual Meeting of North American Fuzzy Information Processing Society», Tampa, USA, 1998. – P. 50–55.
11. Rudzinski F. A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers / F. Rudzinski // Applied Soft Computing. – 2016. – Vol. 38. – P. 118–133.
12. Shtovba S. Ensuring accuracy and transparency of Mamdani fuzzy model in learning by experimental data / S. Shtovba // Journal of Automation and Information Sciences. – 2007. – Vol. 39, № 8. – P. 39–52.

Стаття надійшла до редакції 21.12.2015.

Після доробки 26.01.2016.

Shtovba S. D.¹, Galushchak A. V.²

¹Prof., Dr.Sc., Vinnytsia National Technical University, Vinnytsia, Ukraine

²Assistant, Vinnytsia National Technical University, Vinnytsia, Ukraine

FUZZY CLASSIFIER LEARNING BASED ON DISTANCE BETWEEN THE MAIN COMPETITORS

The classification problem is the assignment an object with certain features to one of classes. Various engineering, management, economic, political, medical, sport, and other problems are reduced to classification. In fuzzy classifiers «inputs – output» relation is described by linguistic <If – then> rules. Antecedents of these rules contain fuzzy terms «low», «average», «high» etc. To increase the correctness it is necessary to tune the fuzzy classifier on experimental data. The new criteria for fuzzy classifier learning that take into account the difference of membership degrees to the main competitors only are proposed. When the classification is correct, the main competitor of the decision is the class with the second largest membership degree. In cases of misclassification the wrong decision is the main competitor to the correct class. Computer experiments with learning the fuzzy classifier of 3 kinds of Italian wines recognition showed a significant advantage of the new criteria. Among new learning criteria the criterion in the form of squared distance between main competitors with the penalty for wrong decision has minor advantage. New criteria can be used not only for tuning fuzzy classifiers but for tuning some other models, such as neural networks.

Keywords: classification, fuzzy knowledge base, tuning, learning criteria, main competitors.

REFERENCES

1. Kuncheva L. I. Fuzzy classifier design, *Studies in Fuzziness and Soft Computing*. Berlin, Heidelberg, Springer-Verlag, 2000, Vol. 49, 314 p.
2. Shtovba S. Pankevich O., Nagorna A. Analyzing the criteria for fuzzy classifier learning, *Automatic Control and Computer Sciences*, 2015, Vol. 49, No. 3, pp. 123–132.
3. Madala H. R., Ivakhnenko A. G. Inductive learning algorithms for complex systems modeling. Boca Raton, CRC Press, 1994, 368 p.
4. Bellman R., Kalaba R., Zadeh L. Abstraction and pattern classification, *Journal of Mathematical Analysis and Applications*, 1966, Vol. 13, No. 1, pp. 1–7.
5. Ishibuchi H., Nozaki K., Yamamoto N., Tanaka H. Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms, *Fuzzy sets and systems*, 1994, Vol. 65, No. 2, pp. 237–253.
6. Ishibuchi H., Nakashima T., Nii M. Classification and modeling with linguistic information granules: advanced approaches advanced approaches to linguistic data mining. Ishibuchi, Berlin-Heidelberg, Springer-Verlag, 2005, 307 p.
7. Shtovba S. D. Porivnjannja kriteriiv navchannja nechitkogo klasifikatora, *Visnik Vinnic'kogo politehnicznego institutu*, 2007, No. 6, pp. 84–91.
8. Abe S., Lan M. S., Thawonmas R. Tuning of a fuzzy classifier derived from data, *International Journal of Approximate Reasoning*, 1996, Vol. 14, pp. 1–24.
9. Nauck D., Kruse R. A neuro-fuzzy method to learn fuzzy classification rules from data, *Fuzzy Sets and Systems*, 1997, Vol. 89, No. 3, pp. 277–288.
10. Rotshtein A. P., Katelnikov D. I. Design and Tuning of Fuzzy If–Then Rules for Automatic Classification, *Proc. of NAFIPS'98 – International Conf. «Annual Meeting of North American Fuzzy Information Processing Society»*. Tampa, USA, 1998, pp. 50–55.
11. Rudzicki F. A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers, *Applied Soft Computing*, 2016, Vol. 38, pp. 118–133.
12. Shtovba S. Ensuring accuracy and transparency of Mamdani fuzzy model in learning by experimental data, *Journal of Automation and Information Sciences*, 2007, Vol. 39, No. 8, pp. 39–52.