

Анастасія Галушак

**ЕФЕКТИВНІ КРИТЕРІЇ НАВЧАННЯ НЕЧІТКИХ КЛАСИФІКАТОРІВ**

Нечіткі класифікатори є корисним інструментом в області інтелектуального аналізу даних. Вони забезпечують моделі точність та інтерпретабельність. Широко та успішно використовуються в екологічних дослідженнях, розпізнаванні обличч, мови та підписів, медичному аналізі зображень, дистанційному зондуванні Землі і т.п.

Нечіткий класифікатор являє собою відображення  $\mathbf{X} = (x_1, x_2, \dots, x_n) \rightarrow y \in \{l_1, l_2, \dots, l_m\}$  на основі бази нечітких правил. Отже, базу правил нечіткого класифікатора запишемо так [1]:

$$\begin{aligned} \text{Якщо } (x_1 = \tilde{a}_{1j} \text{ та } x_2 = \tilde{a}_{2j} \text{ та } \dots \text{ та } x_n = \tilde{a}_{nj}) \\ \text{з вагою } w_j, \text{ тоді } y = d_j, j = \overline{1, k} \end{aligned} \quad (1)$$

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 (1):

$$\mu_j(\mathbf{X}^*) = w_j \cdot (\mu_j(x_1^*) \wedge \mu_j(x_2^*) \wedge \dots \wedge \mu_j(x_n^*)) \quad j = \overline{1, k} \quad (2)$$

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

$$\mu_{l_s}(y^*) = \max_{\forall j: d_j=l_s} (\mu_j(\mathbf{X}^*)), \quad s = \overline{1, m} \quad (3)$$

Нечітким рішенням задачі класифікації буде нечітка множина

$$\tilde{y}^* = \left( \frac{\mu_{l_1}(y^*)}{l_1}, \frac{\mu_{l_2}(y^*)}{l_2}, \dots, \frac{\mu_{l_m}(y^*)}{l_m} \right) \quad (4)$$

Кінцевим результатом виведення оберемо ядро нечіткої множини (4), тобто клас з максимальним ступенем належності:

$$y^* = \arg \max_{\{l_1, l_2, \dots, l_m\}} \max_{s=1, m} (\mu_{l_s}(y^*)).$$

Для підвищення безпомилковості нечіткий класифікатор навчають за експериментальними даними. Для цього змінюють його параметри, щоб мінімізувати відстань між експериментальними даними та результатами нечіткого виведення. Цю відстань, яка називається критерієм навчання, можна визначати у різний спосіб [1].

Серед найбільш популярних критеріїв навчання нечіткого класифікатора можна виділити такі:

частота помилок класифікації, відстань між нечіткими результатами класифікації та експериментальними даними, а також їх комбінації з додатковими штрафами. В даній роботі пропонуються нові критерії навчання нечітких класифікаторів:

1) відстань між головними конкурентами зі штрафом за помилкове рішення.

Ідея цього критерію полягає у врахуванні різниці належностей нечіткого висновку лише до головних конкурентів. За алгоритмом логічного виведення рішенням обирають клас із максимальним ступенем належності.

2) квадратична відстань між головними конкурентами зі штрафом за помилкове рішення.

Цей критерій є модифікацією попереднього. Відмінність полягає у використанні не абсолютних відстаней, а їх квадратів.

Експерименти з виявлення критерію навчання, який забезпечує найкращу безпомилковість проводились для тестової задачі Wine Dataset з UCI Machine Learning Repository. Критерій в основі якого лежить квадратична відстань між головними конкурентами зі штрафом має незначну перевагу для даної задачі.

**Література**

1. Штовба С.Д., Галушак А.В. Критерії навчання нечіткого класифікатора на основі відстані між головними конкурентами // Радіоелектроніка, інформатика, управління. - 2016. - № 2. - С.70-76.