

РОЗРОБКА НЕЧІТКОГО КЛАСИФІКАТОРА РІВНЯ САЖІ У ВІДПРАЦЬОВАНИХ ГАЗАХ АВТОМОБІЛЯ

Вінницький національний технічний університет

Анотація

Проведені експерименти із настроювання нечіткого класифікатора для задачі визначення норми концентрації сажі у вихлопі автомобіля з урахуванням нових критеріїв навчання.

Ключові слова: класифікація, нечітка база знань.

Abstract

Experiments with the setting up of a fuzzy classifier for the task of determining the norm of soot concentration in the car exhaust taking into account the new learning criteria.

Keywords: classification, fuzzy knowledge base.

Класифікація є однією з найбільш популярних задач інтелектуального аналізу даних і вирішує поширені проблеми в техніці, екології, медицині, економіці та інших галузях людського життя. Класифікація об'єкту полягає в його віднесенні за атрибутами до певного класу [1].

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 [2]:

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

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:

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

Нечітким рішенням задачі класифікації буде нечітка множина

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

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

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

Для ефективної роботи класифікатора його необхідно настроїти, тобто знайти такі параметри функцій приналежностей термів вхідних змінних і вагових коефіцієнтів, які мінімізують відхилення між бажаними і дійсними результатами нечіткої класифікації на навчальній вибірці.

Серед найбільш популярних критеріїв навчання нечіткого класифікатора можна виділити такі як частота помилок класифікації, відстань між нечіткими результатами класифікації та експериментальними даними, а також їх комбінації з додатковими штрафами. Розглянемо детальніше нові критерії [3]:

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.

Для перевірки критеріїв на найкращу безпомилковість розглядається тестова задача визначення норми концентрації сажі у вихлопі автомобіля. База даних містить результати по 3-м показникам, а саме завантаженість, позиція педалі і циклова подача. Для експериментів визначено 1 з 3 класів нормування концентрації сажі у вихлопі автомобіля, а саме в межах норми, перевищує норму, значно перевищує норму.

Навчальну вибірку сформовано з рядків бази даних з граничними значеннями кожного із 3 атрибутів. Додатково в навчальну вибірку включено всі непарні рядки бази даних. Всі інші дані занесено в тестову вибірку.

Експерименти проведено для нечіткого класифікатора за такою базою знань:

Якщо x_1 (завантаженість) низька і x_2 (позиція педалі) слабка і x_3 (циклова подача) низька, тоді y =Клас 1

Якщо x_1 (завантаженість) висока і x_2 (позиція педалі) сильна і x_3 (циклова подача) низька, тоді y =Клас 1

Якщо x_1 (завантаженість) висока і x_2 (позиція педалі) слабка і x_3 (циклова подача) висока, тоді y =Клас 2

Якщо x_1 (завантаженість) дуже висока і x_2 (позиція педалі) сильна і x_3 (циклова подача) висока, тоді y =Клас 3

Проведені експерименти виявили перевагу нових критеріїв навчання у порівнянні з традиційними. Ідентифікація автомобілів з підвищеною концентрацією сажі у відпрацьованих газах дозволить приймати управлінські рішення, що дадуть змогу поліпшити стан навколишнього природного середовища за рахунок покращення організації перевезень в місті.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Штовба С.Д. Ідентифікація багатофакторних залежностей за допомогою баз знань. Лабораторний практикум : електронний навчальний посібник / С. Д. Штовба, А. В. Галушак – Вінниця : ВНТУ, 2015. – 96 с.

2. Штовба С. Д. Навчання нечіткого класифікатора з урахуванням лише головних конкурентів / С. Д. Штовба, А. В. Галушак // Вісник Вінницького політехнічного інституту. - 2016. - №1. – С.124-132.

3. Галушак А.В. Ефективні критерії навчання нечітких класифікаторів / А.В. Галушак // Матеріали XV Міжнародної конференції «Контроль і управління в складних системах (КУСС-2020)», м. Вінниця 8-9 жовтня 2020 р. – Електрон. текст. дані. – Вінниця : ВНТУ, 2020.

Галушак Анастасія Володимирівна — асистент кафедри комп'ютерних систем управління, Вінницький національний технічний університет, м. Вінниця, e-mail: galushchak.a.v@gmail.com.

Galushchak Anastasiia V. — assistant of Computer Control System Department, Vinnytsia National Technical University, Vinnytsia, e-mail: galushchak.a.v@gmail.com.