

КЛАСТЕРИЗАЦІЯ ОБ'ЄКТІВ В УМОВАХ НЕЧІТКИХ ВИХІДНИХ ДАНИХ

Карпенко В.В., Іванчихін Ю.В., Андрієнко С.А., Рябоконь Р.Н.

*Національний технічний університет
«Харківський політехнічний інститут», м. Харків*

Канонічна задача кластеризації ставиться і вирішується наступним чином. Нехай розташування n об'єктів задано відповідними наборами координат у m -мірному просторі ознак. Ці об'єкти мають бути розподілені на k груп (кластерів). При вирішенні багатьох завдань діагностики стану (в економіці, техніці, соціології, медицині тощо) розташування центрів групування заздалегідь відоме. У результаті кластеризації об'єкти мають бути розподілені так, щоб сума квадратів відстаней від об'єктів цієї групи до відповідних центрів групування була мінімальною. Процедура кластеризації дуже проста. Для кожного з об'єктів обчислюється відстань до центрів групування та обирається найближчий. У цьому випадку порядок приєднання об'єктів не впливає на результат. Дещо складніша процедура кластеризації у випадку, якщо положення центрів групування заздалегідь не визначено. У цьому випадку використовується матричний метод, який реалізується наступним чином. Спочатку в m -мірному просторі координат встановлюється початкове розташування k центрів групування, максимально віддалених один від одного. Це можна зробити, наприклад, розташувати ці центри рівновіддалено на колі, умовним центром якого є точка з координатами, рівними середнім значенням відповідних координат об'єктів групування, а радіус дорівнює середньому значенню відстаней від об'єктів до умовного центру групування. Тепер формується матриця відстаней від кожного об'єкта кластеризації до кожного з центрів групування. Далі обирається пара "об'єкт групування - центр групування" з мінімальним значенням відстані. Відповідний об'єкт приєднується до відповідного кластера і виконується операція коригування положення центру цього кластера. Далі процедура повторюється необхідну кількість разів.

Процедура кластеризації ускладнюється, якщо координати об'єктів задані нечітко своїми функціями приналежності. При цьому нечітка відстань між об'єктами обчислюється відповідно до правил виконання операцій над нечіткими числами [1], а мінімальна з відстаней визначається за методикою, запропонованою в [2].

Література:

1. Raskin L., Sira O. Performing arithmetic operations over the (L–R)-type fuzzy numbers // Eastern-European Journal of Enterprise Technologies. 2020. Vol. 3(105), No 4. P. 6–11.
2. Raskin L., Sira O., Ivanchykhin Y. Models and methods of regression analysis under conditions of fuzzy initial data // Eastern-European Journal of Enterprise Technologies. 2017. Vol. 4(88). P. 12–19.