

КЛАСТЕРИЗАЦІЯ ВХІДНОЇ ІНФОРМАЦІЇ ДЛЯ ДІАГНОСТИКИ НЕВРОЛОГІЧНИХ ЗАХВОРЮВАНЬ

Телішевська А.В.

*Чернівецький факультет Національного технічного університету
«Харківський політехнічний інститут», м. Чернівці*

При діагностиці неврологічних захворювань значна кількість діагностичних ознак мають описовий характер [1], тому виникає необхідність застосування методів стохастичного факторного аналізу, які дають змогу врахувати вплив сукупності факторів, що носять імовірнісний, невизначений характер.

Основою кластеризації об'єктів є матриця відстаней D . Використовуючи матрицю відстаней D , можна реалізувати агломеративну ієрархічну процедуру кластерного аналізу (КА). Принцип роботи ієрархічних агломеративних процедур полягає в послідовному об'єднанні груп елементів спочатку найближчих, а потім все більш віддалених один від одного [2].

На першому кроці алгоритму кожне спостереження розглядається як окремий кластер. Надалі на кожному кроці роботи алгоритму відбувається об'єднання двох найближчих кластерів, тобто якщо

$$\min_{i,j} d(K_i, K_j) = d(K_s, K_q) = d_{sq},$$

то створюється новий кластер $K_{sp} = K_s \mathbf{U} K_p$, і знову будується матриця відстаней, розмірність якої знижується на одиницю. Робота алгоритму закінчується, коли всі спостереження об'єднані в один кластер.

Ієрархічні методи кластеризації, хоча й точні, але трудомісткі: на кожному кроці необхідно вибудовувати матрицю відстаней для всіх поточних кластерів. Тому при наявності великої кількості об'єктів застосовують інші методи, загальним недоліком яких є апріорне знання числа кластерів. В ієрархічних же процедурах кластеризації оптимальне число кластерів – один з результатів роботи алгоритму.

1. Телішевська А.В. Формалізація вхідної інформації для діагностики неврологічних захворювань. / А.В. Телішевська, А.І. Поворознюк // Матеріали науково-практичної конференції «MicroCad 2011». – Харків. – 2011. – С. 162 – 167.
2. Ким Дж.-О. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка. – М: Финансы и статистика, 1989. – 215 с.