КЛАСТЕРНЫЙ АНАЛИЗ

Автор: словарь
16.07.2008 19:49 -

Кластерный анализ [англ. cluster — гроздь, группа, скопление] — широкий класс процедур многомерного статистического анализа, позволяющих произвести автоматизированную группировку наблюдений в однородные классы — кластеры. Входным материалом для К. а. служат попарные показатели различия или сходства классифицируемых объектов. В частности, это могут быть коэффициенты корреляции между переменными. В результате К. а. исследователь получает разбиение объектов на классы. Среди множества различных алгоритмов К. а. наибольшей популярностью до сих пор пользуются алгоритмы т.н. иерархической кластеризации, предложенные еще в 1960-е гг. Эти алгоритмы дают не одно, а несколько разбиений на разных уровнях сходства (различия) между объектами. При этом строится т.н. дерево кластеризации, позволяющее наглядно увидеть, какие классы объектов являются более внутренне однородными, а какие — более общими, внутренне разнородными, а также увидеть, на каких уровнях более мелкие однородные классы объединяются в более крупные. Применение компьютерной техники расширило возможности прикладного использования К. а. специалистами, далекими от понимания математических тонкостей этого метода. Многие пакеты стандартных статистических программ (такие, как BMDP, SPSS, Systat, SAS, отечественный СТАДИЯ и др.) включают как минимум один алгоритм кластерного анализа. При этом от внимания неискушенных пользователей этих программ нередко ускользает, что результаты К. а. в значительной степени зависят не только от состава классифицируемых объектов, не только от состава признаков (параметров) этих объектов, по которым устанавливается сходство— различие (близость—расстояние) между ними, но и от избранной метрики (формулы) сходства или различия. Особой разновидностью К. а. является "анализ клик", или "клайк-анализ" (от англ. clique — клика), который дает разбиение объектов на пересекающиеся классы. Процедуры К. а. и "анализа клик" близки к таким полуавтоматизированным и в значительной степени визуальным методам группировки, как анализ корреляционных плеяд, граф-анализ, путевой анализ и т. п. В психодиагностике К. а. позволяет выявить группы испытуемых, обладающих близкими психодиагностическими профилями (соотношением выявленных и измеренных психических свойств).

При конструировании тестов К. а. позволяет сгруппировать родственные тестовые пункты (задания) и перепроверить тем самым результаты другого метода построения факторных шкал — результаты факторного анализа. В социальной психологии К. а. позволяет увидеть, как группируются респонденты (например, члены какой-то группы или организации) по сходству их социальных установок или оценок. В экспериментальной психосемантике К. а. позволяет промоделировать категориальные структуры сознания в виде иерархически организованных систем индивидуальных значений.

А.Г. Шмелев