



УДК 519.7

МЕТОД КОГНІТИВНОЇ КЛАСТЕРИЗАЦІЇ

Студ. В.В. Паливода, гр. БМР 1-17

Науковий керівник доц. О.Л. Блохін

Київський національний університет технологій та дизайну

Мета і завдання. Мета - знаходження груп схожих об'єктів у вибірці для подальшого їх використання в археології, антропології, медицині, психології, хімії, біології, державному управлінні, філології, маркетингу, соціології та інших дисциплінах.

Завдання - навчитися розбивати вибірку на непересічні підмножини, що називаються кластерами, так, щоб кожен кластер складався з об'єктів, близьких по матриці, а об'єкти різних кластерів істотно відрізнялися. При цьому кожному об'єкту навчитися приписувати номер кластера.

Об'єкт та предмет дослідження. Об'єкт - безліч змінних, за якими будуть оцінюватися об'єкти у вибірці. Предметом є закони обробки, зберігання і відтворення інформації, зв'язок і можливість їх реалізації на інших носіях, дослідження в області штучного інтелекту.

Методи та засоби дослідження.

Об'єднання схожих об'єктів у групи може бути здійснене різними способами. Саме для цього етапу існує цілий ряд методів:

- К-середніх (K-means)
- Нечітка кластеризація C-середніх (C-means)
- Графові алгоритми кластеризації
- Статистичні алгоритми кластеризації
- Алгоритми сімейства FOREL
- Ієрархічна кластеризація або таксономія
- Нейронна мережа Кохонена
- Ансамбль кластеризаторів
- Алгоритми сімейства KRAV
- EM-алгоритм
- Метод просіювання

Відомі широкі застосування кластерного аналізу в маркетингових дослідженнях, а також в області нечіткої математики, коли необхідно створити нечітку базу знань (вхідні величини пов'язати з вихідними за допомогою нечітких правил).

Наукова новизна та практичне значення отриманих результатів.

Кластерний аналіз дуже широко застосовується як у науці, так і в різних напрямках практичної діяльності, тому завжди буде актуальним і носитиме в собі велике значення. Різноманіття розроблених методів вказує на те, що їх можна розглядати як різні більш чи менш успішні варіанти рішення або спроби вирішення проблем, існуючих в області кластерного аналізу.

Результати дослідження.

Кластерний аналіз пред'являє наступні вимоги до даних: по-перше, показники не повинні корелювати між собою; по-друге, показники повинні бути безрозмірними; по-третє, їх розподіл має бути близьким до нормального; по-четверте, показники повинні відповідати вимогу «стійкості», під якою розуміється відсутність впливу на їх значення випадкових факторів; по-п'яте, вибірка повинна бути однорідна, не містити «викидів». Якщо кластерному аналізу передують факторний аналіз, то при вибірці викладені вимоги



виконуються автоматично самою процедурою факторного моделювання. В іншому випадку вибірку потрібно коригувати.

Робота кластерного аналізу спирається на два припущення: перше припущення – розглянуті ознаки об'єкта в принципі допускають бажане розбиття сукупності об'єктів на кластери, друге припущення- правильність вибору масштабу або одиниць вимірювання ознак.

Степінь істинності знань, отриманих в результаті кластерного аналізу, тобто їх адекватність та відповідність дійсності, повністю визначаються ступенем істинності методу кластерного аналізу, з допомогою якого вони отримані.

Висновки

Кластерний аналіз, за оцінкою науковців, має велике значення в проведенні аналітичних досліджень завдяки можливості перетворити великий обсяг різнобічної інформації в упорядкований, компактний вигляд. Це сприяє підвищенню рівня наочності, зрозумілості та сприйняття результатів аналізу, а також створює підґрунтя для прогнозування.

Ключові слова: кластеризація, алгоритм, метод, аналіз, корелювання.

ЛІТЕРАТУРА

1. Барсегян А.А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP – СПб. : BHV, 2007. – 384 с
2. . Гайдишев І. Аналіз та обробка даних: спеціальний довідник - СПб: Питер, 2001.
3. Шуметов В. Г. Шуметова Л. В. Кластерный анализ: подход с применением ЭВМ. — Орел: ОрелГТУ, 2000. — 118 с.
4. Мандель И. Д. Кластерный анализ. — М.: Финансы и статистика, 1988. М
5. Олдендерфер М. С., Блэшфилд Р. К. Кластерный анализ / Факторный, дискриминантный и кластерный анализ: пер. с англ.; Под. ред. И. С. Енюкова. — М.: Финансы и статистика, 1989. — 215 с.