

Иерархический кластерный анализ: пример

Введение в кластерный анализ

Этапы иерархического кластерного анализа

Шаг 1. Анализ исходных переменных и подготовка данных.

- Все шкалы должны быть количественными.
- Нужно убрать выбросы.

Шаг 2. Выбор методов кластеризации.

Для нашего примера:

- Метод Уорда как основной.
- Метод межгрупповых связей для сравнения.

Шаг 3. Построение модели. Получение протокола и дендрограммы.

Этапы иерархического кластерного анализа

Шаг 4. Определение количества кластеров. Сравнение результатов кластеризации

Шаг 5. Интерпретация модели.

- Чтобы понять, что характеризует кластеры, надо посмотреть их характеристики.
- Самый простой вариант — «координаты центров кластеров» как отражение «сути» построенных групп.

Преимущества агломеративных методов классификации

1. Виден весь процесс кластеризации.
2. Разнообразие методов кластеризации.
3. Возможность выбора.

Ограничения агломеративных методов классификации

1. Не подходят для больших выборок.
2. Нужны дополнительные действия для того, чтобы проинтерпретировать модель (увидеть характеристики кластеров).