



	Единица наблюдения (о чем делаются выводы)	Есть характеристики ед. наблюдения	Есть целевая переменная	Итог (правило/принадлежность к группе / значение целевой переменной)	Совстречаемость ед. наблюдения?
Кластеризация	item	да	нет	принадлежность к группе	нет
Ассоциативные правила		нет	нет	правило	да
Subgroup Discovery	feature	да	да	правило	нет

У вас есть сервис для покупки, чтения, обсуждения книг.

Задача 1: вы хотите сделать персонализированную рекомендацию для пользователей вашего сервиса. Но вы ограничены в возможностях и не можете сделать отдельное предложение для каждого пользователя. Вы решили выделить группы пользователей, а дальше делать одинаковые предложения каждой группе. У вас есть данные о поведении пользователей (пол, возраст, любимый жанр, регулярность использования сервиса, общие затраты, количество прочитанных книг, количество написанных отзывов).

Задача 2: У вас есть данные о поведении пользователей (пол, возраст, любимый жанр, регулярность использования сервиса, общие затраты, количество прочитанных книг, количество написанных отзывов). Вы хотите понять, кто те пользователи, которые тратят больше

Задача 3: У вас есть данные о поведении пользователей (пол, возраст, любимый жанр, регулярность использования сервиса, общие затраты, количество прочитанных книг, количество написанных отзывов). Вы хотите понять, кто те пользователи, которые тратят мало, но много пишут отзывов

Задача 4: У вас есть данные о том, что именно пользователи покупают по акции «Купи две книги, получи третью бесплатно». Вы хотите предлагать третью книгу в корзину даже за пределами акции (т.е. пользователь выбирает две книги, ему рекомендуем третью)

Задача 5: У вас есть данные о поведении пользователей (пол, возраст, любимый жанр, регулярность использования сервиса, общие затраты, количество прочитанных книг, количество написанных отзывов). Вы хотите понять, связаны ли характеристики между собой (например, «женщины, читающие фантастику, пользуются сервисом не реже одного раза в пару дней»)

Задача 6: вы хотите понять, что влияет на удовлетворенность пользователя. У вас есть данные о поведении пользователей (пол, возраст, любимый жанр, регулярность использования сервиса, общие затраты, количество прочитанных книг, количество написанных отзывов), его оценка удовлетворенности, характеристики книг, которые он прочитал (жанр, рейтинг, тональность отзывов, серия или нет, номер книги в серии)

Задача 7: вы хотите понять, какие характеристики пользователя связаны с тем фактом, что он не читает другие книги из серий. У вас есть характеристики пользователя, оценка предыдущих книг и характеристики книг (жанр, рейтинг, тональность отзывов, номер книги в серии)

Задача 8: вы хотите понять, понравится ли пользователю книга. У вас есть характеристики книги и история просмотров пользователя с характеристиками просмотренных и прочитанных книг

В **Subgroup Discovery** результат это правило, с помощью которого мы разделяем на группы.

Совстречаемость нужна только при использовании ассоциативных правил (здесь нам не нужны характеристики). Меньше нужна в **Subgroup Discovery** и совсем не нужна при кластерном анализе

Кластеризация использует все доступные характеристики, она не выделит подмножества или какие-то узкие правила. Это больше про ассоциативные правила