Тестовые вопросы по "Ассоциативный анализ"

1. Что является основной целью ассоциативного анализа?

Целью ассоциативного анализа, является нахождения всех правил XY(зависимость между связанными событиями или элементами), причем поддержка и достоверность этих правил должны быть выше некоторых наперед определенных порогов, называемых соответственно минимальной поддержкой (minsupport) и минимальной достоверностью (minconfidence).

2. Какой тип данных чаще всего используется в ассоциативном анализе?

Для реализации работы с алгоритмами выделения ассоциаций чаще всего используется специальные типы данных, относящиеся к объектам трех классов: входной массив транзакций (transactions) и на выходе - часто встречающиеся фрагменты данных (itemsets) и правила (rules).

3. Что такое "поддержка" (support) правила в ассоциативном анализе?

"Поддержка" (support) – это одна из метрик, которая используется для измерения распространенности или частоты набора элементов в базе данных.

4. Что такое "достоверность" (confidence) правила в ассоциативном анализе?

"Достоверность" (confidence) - это одна из метрик, которая измеряет вероятность того, что если в транзакции присутствует антецедент (условие), то в этой же транзакции будет присутствовать и консеквент (следствие).

5. Что такое "подъем" (lift) правила в ассоциативном анализе?

"Подъем" (lift) - это одна из метрик, которая измеряет, насколько более вероятно появление консеквента (следствия) в транзакции, содержащей антецедент (условие), по сравнению с вероятностью появления консеквента в произвольной транзакции.

6. Какой алгоритм часто используется для поиска ассоциативных правил?

Наиболее часто используемым алгоритмом для поиска ассоциативных правил является алгоритм Apriori.

7. Что такое "антецедент" в ассоциативном правиле?

В ассоциативных правилах антецедент указывает условия, которые, если выполняются, могут привести к выполнению другого условия.

8. Что такое "консеквент" в ассоциативном правиле?

Консеквент представляет собой результатом или событие, которое вероятно произойдет, если выполнены условия, заданные антецедентом (предпосылкой).

9. Какое значение "подъема" (lift) указывает на независимость антецедента и консеквента?

Значение подъема (lift), равное 1, указывает на независимость антецедента и консеквента в ассоциативном правиле.

10. Ассоциативный анализ чаще всего применяется в: Анализ корзины покупок, формирование товарных рекомендации, планирование и размещение товаров в магазине.

Тестовые вопросы по mlxtend.frequent\_patterns.apriori

1. Какой тип данных принимает функция ***apriori*** из библиотеки ***mlxtend.frequent\_patterns*** ?

DataFrame

2. Какой параметр в функции ***apriori*** отвечает за минимальную поддержку?

min\_support

3. Какой параметр в функции ***apriori*** позволяет использовать названия столбцов DataFrame вместо индексов?

use\_colnames

4. Какой тип данных возвращает функция ***apriori*** ?

DataFrame

5. Какой столбец в возвращаемом DataFrame содержит найденные часто встречающиеся наборы элементов?

itemsets

6. Какой столбец в возвращаемом DataFrame содержит значение поддержки для каждого часто встречающегося набора?

support

7. Как можно отфильтровать результаты, полученные с помощью ***apriori*** , по длине наборов элементов?

Использовать фильтрацию по столбцу itemsets, извлекая длину набора

8. Для чего используется параметр ***low\_memory*** в функции ***apriori*** ?

Для снижения потребления памяти при обработке больших наборов данных.

9. Какой параметр в функции ***apriori*** позволяет задать максимальную длину наборов

элементов?

max\_len

10. Что нужно сделать перед использованием функции ***apriori*** , если данные представлены в виде списка транзакций с текстовыми элементами?

Данные надо закодировать так, чтобы их можно было представить в виде матрицы. Для кодированния данных используем TransactionEncode.

Тестовые вопросы по FPGrowth

1. Что является основной задачей алгоритма FPGrowth?

Основной задачей алгоритма FPGrowth - поиск часто встречающихся наборов элементов.

2. Какое основное преимущество FPGrowth по сравнению с Apriori?

Основное преимущество FPGrowth по сравнению с Apriori - позволяет избежать затратной процедуры генерации кандидатов, характерной для алгоритма Apriori.

3. Какая структура данных используется в алгоритме FPGrowth для хранения информации о транзакциях?

Структура данных FP-дерево (Frequent Pattern Tree).

4. Что такое "условная база паттернов" (conditional pattern base) в алгоритме FPGrowth?

Набор транзакций, содержащих определенный элемент, с удалением элементов, предшествующих ему в дереве.

5. Что такое "условное FP-дерево" (conditional FP-tree) в алгоритме FPGrowth?

FP-дерево, построенное на основе условной базы паттернов.

6. Как влияет порядок элементов в транзакциях на построение FP-дерева?

Порядок влияет на структуру дерева и его эффективность.

7. Какой параметр определяет минимальную поддержку в алгоритме FPGrowth?

min\_support

8. FPGrowth, как и Apriori, используется на первом этапе поиска:

Частых наборов элементов.

9. Какой из следующих недостатков может быть у алгоритма FPGrowth?

Может быть ресурсозатратным при очень больших данных или высокой частоте элементов

10. После нахождения часто встречающихся наборов с помощью FPGrowth, что обычно делается для получения ассоциативных правил?

Генерация ассоциативных правил на основе найденных частых наборов.