**Сравнительный анализ Apriori, ECLAT, FP-Growth и FPMax.**

1. **Apriori**

Алгоритм Apriori работает в 2 этапа:

1. Генерация частых наборов
2. k = 1, выбрать частые одноэлементные наборы, где поддержка будет больше минимальной.
3. k = k + 1. Если возможно сгенерировать k-элементный набор, то перейти к 3-му шагу, иначе, к 5-му.
4. Генерация k-элементного набора на основе (k-1)-элементного.
5. Посчитать поддержку для каждого кандидата и убрать те, чей порог ниже минимального. Вернуться на шаг 2.
6. Вернуть все частые k-элементные наборы для всех k.
7. Генерация правил
8. Сгенерировать высоко-уверенные правила для каждого набора

Особенности:

1. Если k-элементный набор часто встречается, то и (k-1)-элементный набор будет встречаться часто.
2. Принцип анти-монотонности.

Недостатки:

1. Трудоемкая и по времени и по памяти генерация новых кандидатов, несмотря на принцип антимонотонности.
2. Многократное повторное сканирование БД на предмет соотнесения сгенерированных кандидатов с реальными транзакциями.

**2. FP-Growth**

Frequent pattern growth

Особенности алгоритма:

* Используется FP-tree, где каждый узел содержит элемент набора, рассчитанную поддержку и ссылки на дочерние элементы набора.
* Нет многократной генерации кандидатов.
* Нет многократного сканирования БД. БД сканируется только 2 раза.

Генерация правила достигается построением базиса, а не перебором.

**3. ECLAT(Equivalence CLAss Transformation)**

Работа алгоритма:

1. Для каждого элемента сохранить список id транзакций (TID).
2. Рассчитать поддержку для каждого k-элементного набора путем пересечения двух (k-1)-элементных наборов.

Особенности алгоритма:

Преимущества:

* Сканирование БД происходит ровно 1 раз.
* Очень быстрый подсчет поддержки.

Недостатки:

* Постоянный TID список может занимать слишком много памяти.

**4.FP-MAX**

FP-Max - это вариант FP-Growth, который фокусируется на получении максимальных наборов предметов. Набор элементов X называется максимальным, если X является частым и не существует частого супер-шаблона, содержащего X. Другими словами, частый шаблон X не может быть под-шаблоном более частого шаблона, чтобы соответствовать определению максимального набора элементов