**Aula Prática 13**

**Regras de Associação**

**Murielly Oliveira Nascimento – 11921BSI222**

1. Dado a base de dados de transações, calcule

|  |  |
| --- | --- |
| **Id Transação** | **Itens comprados** |
| 1 | {a,d,e} |
| 24 | {a,b,c,e} |
| 12 | {a,b,d,e} |
| 31 | {a,c,d,e} |
| 15 | {b,c,e} |
| 22 | {b,d,e} |
| 29 | {c,d} |
| 40 | {a,b,c} |
| 33 | {a,d,e} |
| 38 | {a,b,e} |

1. Qual o número máximo de itemsets que podem ser extraídos desse conjunto de transações (incluindo aqueles que podem ter suporte zero)?

Como há 5 itens (a,b,c,d,e) são ou 35 itemsets candidatos.

1. Suponha que um 4-itemset é frequente, de acordo com um valor de suporte. Quantas regras podem ser obtidas a partir de um 4-itemsets?

Dado um 4-itemset frequente podem ser geradas ou 14 regras de associação candidatas.

1. Considere o conjunto de 3-itemsets frequentes {1,2,3}, {1,2,4}, {1,2,5}, {1,3,4}, {1,3,5}, {2,3,4}, {2,3,5}, {3,4,5}
2. Liste todos os 4-itemsets candidatos obtidos pelo processo de geração de candidatos usando a estratégia Fk-1 x F1.

Unir {1,2,3} {1,2,4} = {1,2,3,4}

Unir {1,3,4}, {1,3,5} = {1,3,4,5}

Unir {2,3,4}, {2,3,5} = {2,3,4,5}

1. Liste todos os 4-itemsets candidatos obtidos pelo processo de geração de candidatos usando a estratégia do algoritmo Apriori.

Frequência:

1 ->5

2 -> 5

3 -> 6

4 -> 4

5 -> 4

2-itemsets gerados:

{1,2}, frequência 3

{1,3}, frequência 2

{1,4}, frequência 0

{1,5}, frequência 0

{2,3}, frequência 3

{2,4}, frequência 1

{2,5}, frequência 1

{3,4}, frequência 3

{3,5}, frequência 2

{4,5}, frequência 1

Após a poda: {1,2}, {1,3}, {2,3}, {2,4}, {2,5} {3,4}, {3,5}, {4,5}

3-itemsets gerados:

{1,2,3} frequência 1

{2,3,4} frequência 1

{2,3,5} frequência 1

{2,4,5} frequência 0

{3,4,5} frequência 1

Após a poda: {1,2,3}, {2,3,4}, {2,3,5}, {3,4,5}

4-itemsets gerados:

{2,3,4,5}

1. Liste todos os 4-itemsets que irão sobreviver após a etapa de poda dos candidatos pelo algoritmo Apriori.

4-itemsets gerados:

{2,3,4,5}

1. Execute o algoritmo Apriori na base de dados mostrada na tabela a seguir. Mostre o passo-passo da execução e quais são os itemsets gerados ao fim da execução do algoritmo. Considere o suporte de 0.3. Quantas regras válidas podem ser regras podem ser geradas a partir dos itemsets frequentes obtidos pelo algoritmo Apriori. Não é necessário calcular a confiança das regras, basta mostrar quantas regras podem ser geradas e como você chegou nesse cálculo.

Tabela

Descrição gerada automaticamente

Se cada subconjunto for considerado, são gerados = 5 + 10 + 10 = 25.

Calculando a frequência dos itens:

A = 5

B = 7

C = 5

D = 9

E = 6

Note que todos os itens têm frequência superior ao suporte.

2-Itemsets gerados: {a,b}, {a,c}, {a,d}, {a,e}, {b,c}, {b,d}, {b,e}, {c,d}, {d,e}

Calculando a frequência dos itens:

{a,b} -> 3

{a,c} -> 1

{a,d} -> 1

{a,e} -> 0

{b,c} -> 2

{b,d} -> 4

{b,e} -> 0

{c,d} -> 4

{d,e} -> 6

Após a poda, permanecem os seguintes itens:

{a,b} -> 3

{b,d} -> 4

{c,d} -> 4

{d,e} -> 6

3-Itemsets gerados: {a,b,d}, {c,d,e}, {b,c,d}