1. (Cơ bản) Cho bảng dữ liệu ở một cửa hàng tạp hóa có 6 giao dịch như sau:

|  |  |
| --- | --- |
| Transaction ID | Items |
| T1 | HotDogs, Buns, Ketchup |
| T2 | HotDogs, Buns |
| T3 | HotDogs, Coke, Chips |
| T4 | Chips, Coke |
| T5 | Chips, Ketchup |
| T6 | HotDogs, Coke, Chips |

* 1. Tìm tập phổ biến bằng giải thuật Apriori
* Với min\_sup = 33.33% và tổng số giao dịch là 6 => Tần số xuất hiện tối thiểu của phần tử để thỏa min\_sup là 2.
* 1 phần tử:

C1 = {

HotDogs: 4,

Buns: 2,

Ketchup: 2,

Coke: 3,

Chips: 4

}

* + - * Với min\_sup 2 ta được các tập phổ biến 1 phần tử:

L1 = {{HotDogs}, {Buns}, {Ketchup}, {Coke}, {Chips}}

* 2 phần tử:

C2 = {

{HotDogs, Buns}: 2,

{HotDogs, Ketchup}: 1,

{HotDogs, Coke}: 2,

{HotDogs, Chips}: 2,

{Buns, Ketchup}: 1,

{Buns, Coke}: 0,

{Buns, Chips}: 0,

{Ketchup, Coke}: 0,

{Ketchup, Chips}: 1,

{Coke,Chips}: 3

}

* + - * Với min\_sup 2 ta được các tập phổ biến 2 phần tử (tập các ứng viên thoải min\_support\_count):

L2 = {{HotDogs, Buns}, {HotDogs, Coke}, {HotDogs, Chips}, {Coke, Chips}}

* 3 phần tử:

C3 = {

{HotDogs, Buns, Coke}:0,

{HotDogs, Buns, Chips}:0,

{HotDogs, Coke, Chips}:2

}

* + - * Với min\_sup 2 ta được các tập phổ biến 3 phần tử (tập các ứng viên thoải min\_support\_count):

L3 = {{HotDogs, Coke, Chips}}

* 4 phần tử: Vì L3 chỉ có 1 tập con nên không tiến hành tìm tập ứng cử viên 4 phần tử.
  1. Xây dựng cây FP
     + Xây dựng cây FP
* Tập phổ biến 1 phần tử và tần suất xuất hiện:

|  |  |
| --- | --- |
| item | supp |
| HotDogs | 4 |
| Buns | 2 |
| Ketchup | 2 |
| Coke | 3 |
| Chips | 4 |

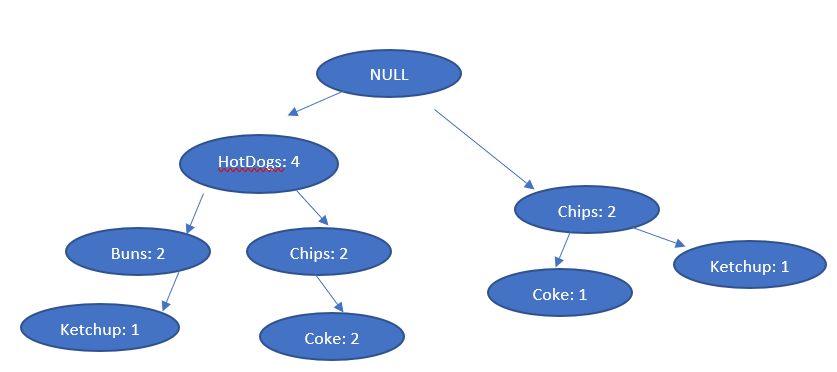
* Tập phổ biến 1 phần tử và tần xuất thỏa điều kiện ≥ min\_supp và đã được sắp xếp giảm dần.

|  |  |
| --- | --- |
| item | supp |
| HotDogs | 4 |
| Chips | 4 |
| Coke | 3 |
| Buns | 2 |
| Ketchup | 2 |

* Điền cột Order Frequent Items vào:

|  |  |  |
| --- | --- | --- |
| TID | items | Order Frequent Items |
| T1 | HotDogs, Buns, Ketchup | HotDogs, Buns, Ketchup |
| T2 | HotDogs, Buns | HotDogs, Buns |
| T3 | HotDogs, Coke, Chips | HotDogs, Coke, Chips |
| T4 | Chips, Coke | Chips, Coke |
| T5 | Chips, Ketchup | Chips, Ketchup |
| T6 | HotDogs, Coke, Chips | HotDogs, Coke, Chips |

* Cây FP
* Header Tablle



* + - Xây dựng cơ sở điều kiện cho mỗi mục phổ biến

|  |  |
| --- | --- |
| item | Conditional Pattern Base Chips |
| Chips | {HotDogs}: 2 |
| Coke | {HotDogs, Chips,}: 2, {Chips}: 1 |
| Buns | {HotDogs}: 2 |
| Ketchup | {HotDogs, Buns}: 1, {Chips}: 1 |

* + - Xây dựng cây FP điều kiện

|  |  |  |
| --- | --- | --- |
| item | Conditional Pattern Base Chips | Conditional FP-Tree (supp ≥2) |
| Buns | {HotDogs}: 2 | {{HotDogs}:2} |Buns |
| Chips | {HotDogs}: 2 | {{HotDogs}:2} |Chips |
| Coke | {HotDogs, Chips,}: 2, {Chips}: 1 | {{HotDogs, Chips}:2, {Chip}:1} |Coke |
| Ketchup | {HotDogs, Buns}: 1, {Chips}: 1 | {} |
| HotDogs | {} | {} |

* + - Kết quả các mẫu thường xuyên

|  |  |  |
| --- | --- | --- |
| item | Conditional FP-Tree | Frequent Pattern Buns |
| Buns | {HotDogs}: 2 | Buns | {Buns:2}, {{HotDogs, Buns}:2} |
| Chips | {HotDogs}: 2 | Chips | {Chips:2}, {{HotDogs, Chips}:2} |
| Coke | {{HotDogs, Chips}:2, {Chips}:1} |Coke | {Coke: 3}, {HotDogs, Coke: 2}, {Coke, Chip:3}, {HotDogs, Chips, Coke:2} |
| Ketchup | {} | {Ketchup:2} |
| HotDogs | {} | {HotDogs:4} |

Lấy những tập có 2 phần tử trở lên và thỏa min\_supp ta được các tập hợp sau đây:

L2 = {

{HotDogs, Buns}: 2,

{HotDogs, Chips}: 2,

{HotDogs, Coke}:2,

{Chips, Coke}:3

}

L3 = {{HotDogs, Chips, Coke}:2}

* 1. Tìm các luật kết hợp dựa trên các tập phổ biến
* Xét tập phổ biến {HotDogs, Buns}: 2 có min\_conf = 60%

{HotDogs}: 4, {Buns}: 2 {HotDogs, Buns}: 2

Ta có

|  |  |
| --- | --- |
| HotDogs 🡪 Buns | 2/4 = 50% |
| Buns 🡪 HotDogs | 2/2 = 1 |

Với min\_conf = 60% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

**Buns 🡪 HotDogs**

* Xét tập phổ biến {HotDogs, Chips}: 2 có min\_conf = 60%

{HotDogs}: 4, {Chips}: 4, {HotDogs, Chips}: 2

Ta có

|  |  |
| --- | --- |
| HotDogs 🡪 Chips | 2/4 = 50% |
| Chips 🡪 HotDogs | 2/4 = 50% |

Với min\_conf = 60% dựa vào bảng trên ta có luật kết hợp ta không tìm được luật thỏa yêu cầu min\_conf = 60%

* Xét tập phổ biến {HotDogs, Coke}: 2 có min\_conf = 60%

{HotDogs}: 4, {Coke}: 3, {HotDogs, Coke}: 2

Ta có

|  |  |
| --- | --- |
| HotDogs 🡪 Coke | 2/4 = 50% |
| Coke 🡪 HotDogs | 2/3 = 6.67% |

Với min\_conf = 60% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

**Coke 🡪 HotDogs**

* Xét tập phổ biến {Chips, Coke}: 2 có min\_conf = 60%

{Chips}:4, {Coke}: 3, {Chips, Coke}:3

Ta có

|  |  |
| --- | --- |
| Chips 🡪 Coke | 3/4 =75% |
| Coke 🡪 Chips | 3/3 = 100% |

Với min\_conf = 60% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

**Coke 🡪 Chips, Chips 🡪 Coke**

* Xét tập phổ biến {HotDogs, Chips, Coke}: 2 có min\_conf = 60%

{HotDogs}: 4, {Coke}: 3, {Chips}: 4, {HotDogs, Coke}: 2, {HotDogs, Chips}: 2, {Chips, Coke}: 3, {HotDogs, Chips, Coke}: 2

Ta có

|  |  |
| --- | --- |
| HotDogs 🡪{Chips, Coke} | 2/4 = 50% |
| Coke 🡪 {HotDogs, Chips} | 2/3 = 66,7% |
| Chips 🡪 {HotDogs, Coke} | 2/4 = 50% |
| {Chips, Coke} 🡪 HotDogs | 2/3 =66,7% |
| {HotDogs, Chips} 🡪 Coke | 2/2 = 100% |
| {HotDogs, Coke} 🡪 Chips | 2/2 = 100% |

Với min\_conf = 60% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

Coke 🡪 {HotDogs, Chips}, {Chips, Coke} 🡪 HotDogs, {HotDogs, Chips} 🡪 Coke, {HotDogs, Coke} 🡪 Chips

1. (Cơ bản) Cho bảng dữ liệu ở một cửa hàng văn phòng phẩm như sau:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| TID | KÉO | COMPA | THƯỚC | TẬP TRẮNG | BÚT BI | BÚT MÀU | TẨY |
| T1 |  | x |  | x | x |  |  |
| T2 | x |  | x | x | x |  |  |
| T3 |  | x |  | x | x |  |  |
| T4 | x | x |  | x | x |  |  |
| T5 |  |  | x |  |  |  |  |
| T6 |  |  |  |  | x |  |  |
| T7 |  |  |  | x |  |  |  |
| T8 |  |  |  |  |  |  | x |
| T9 |  |  |  |  |  | x | x |
| T10 |  |  |  |  |  | x |  |

* 1. Tìm tập phổ biến bằng giải thuật Apriori
* Với min\_sup = 30% và tổng số giao dịch là 10 => Tần số xuất hiện tối thiểu của phần tử để thỏa min\_sup là 3.
* Tập các ứng viên 1 phần tử và tần số xuất hiện của nó (support count) tương ứng như sau:

C1 = {

Kéo: 2,

Compa: 3,

Thước: 2,

Tập trắng: 5,

Bút bi: 5,

Bút màu: 2,

Tẩy:2

}

* + - * Các tập phổ biến thỏa min\_sup 3 là: Compa, Tập trắng, Bút bi
      * Các tập phổ biến 1 phần tử (tập các ứng viên thoải min\_support\_count):

L1 = {

{Compa},

{Tập trắng},

{Bút bi}

}

Tập các ứng viên 2 phần tử và tần số xuất hiện của nó

C2 = {

{Compa; Tập trắng}: 3,

{Compa; Bút bi}:3,

{Bút bi, Tập trắng}:4

}

* + - * Các tập phổ biến thỏa min\_sup 3 là: {Compa; Tập trắng}, {Compa; Bút bi}, {Bút bi, Tập trắng}
      * Các tập phổ biến 2 phần tử (tập các ứng viên thoải min\_support\_count):

L2 = {

{Compa; Tập trắng},

{Compa; Bút bi},

{Bút bi, Tập trắng}

}

* Tập các ứng viên 3 phần tử và tần số xuất hiện của nó

C3 ={{Compa; Tập trắng; Bút bi}:3}

* + - * Các tập phổ biến thỏa min\_sup 3 là: {Compa; Tập trắng; Bút bi}
      * Các tập phổ biến 3 phần tử (tập các ứng viên thoải min\_support\_count):

L3 = {{Compa; Tập trắng; Bút bi}}

* Vì L3 chỉ còn 1 tập hợp nên không tiến hành tìm tập ứng cử viên 4 phần tử.
  1. Xây dựng cây FP
     + Xây dựng cây FP
* Tập phổ biến 1 phần tử và tần suất xuất hiện của nó

|  |  |
| --- | --- |
| item | supp |
| Kéo | 2 |
| Compa | 3 |
| Thước | 2 |
| Tập trắng | 5 |
| Bút bi | 5 |
| Bút màu | 2 |
| Tẩy | 2 |

* Tập phổ biến 1 phần tử và tần xuất thỏa điều kiện ≥ min\_supp và đã được sắp xếp giảm dần

|  |  |
| --- | --- |
| item | supp |
| Bút bi | 5 |
| Tập trắng | 5 |
| Compa | 3 |

* Điền cột Order Frequent Items vào

|  |  |  |
| --- | --- | --- |
| TID | items | Order Frequent Items |
| T1 | Compa, Tập trắng, Bút bi | Bút bi, Tập trắng, Compa |
| T2 | Kéo, Thước, Tập trắng, Bút bi | Bút bi, Tập trắng |
| T3 | Compa, Tập trắng, Bút bi | Bút bi, Tập trắng, Compa |
| T4 | Kéo, Compa, Tập trắng, Bút bi | Bút bi, Tập trắng, Compa |
| T5 | Thước |  |
| T6 | Bút bi | Bút bi |
| T7 | Tập trắng | Tập trắng |
| T8 | Tẩy |  |
| T9 | Bút màu, Tẩy |  |
| T10 | Bút màu |  |

* Vẽ cây FP
* Header Table

|  |  |
| --- | --- |
| item | supp |
| Bút bi | 5 |
| Tập trắng | 5 |
| Compa | 3 |

* + - Xây dựng cơ sở điều kiện cho mỗi mục phổ biến

|  |  |
| --- | --- |
| item | Conditional Pattern Base |
| Compa | {Bút bi; Tập trắng}: 3 |
| Tập trắng | {Bút bi}: 4 |
| Bút bi | {} |

* + - Xây dựng cây FP điều kiện

|  |  |  |
| --- | --- | --- |
| item | Conditional Pattern Base | Conditional FP-Tree (supp ≥2) |
| Compa | {Bút bi; Tập trắng}: 3 | {{Bút bi}:3, {Tập trắng}:3} |Compa |
| Tập trắng | {Bút bi}:4 | {Bút bi}:4 | Tập trắng |

* + - Kết quả các mẫu thường xuyên

|  |  |  |
| --- | --- | --- |
| item | Conditional FP-Tree | Frequent Pattern |
| Compa | {{Bút bi}:3, {Tập trắng}:3} |Compa | {Compa:3}, {Compa; Bút bi}:3, {Compa: Tập trắng}:3, {Compa, Bút bi, Tập trắng}: 3 |
| Tập trắng | {Bút bi}:4 | Tập trắng | {Tập trắng: 4}, {Bút bi; Tập trắng}: 4 |

Lấy những phần tử có 2 phần tử trở lên và thỏa min\_supp ta được các tập hợp sau đây:

L2 = {

{Compa; Bút bi}:3,

{Compa; Tập trắng}: 3,

{Bút bi; Tập trắng}:4

}

L3 = {{Compa; Bút bi; Tập trắng}:3}

* 1. Tìm các luật kết hợp dựa trên các tập phổ biến
* Xét tập phổ biến {Compa; Bút bi}: 3 có min\_conf = 80%

{Compa}:3, {Bút bi}: 5, {Compa; Bút bi}:3

Ta có

|  |  |
| --- | --- |
| Compa 🡪 Bút bi | 3/3 = 100% |
| Bút bi 🡪 Compa | 3/5 = 60% |

Với min\_conf = 80% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

**Compa 🡪 Bút bi**

* Xét tập phổ biến {Compa; Tập trắng}: 3 có min\_conf = 80%

{Compa}:3, {Tập trắng}: 5, {Compa; Tập trắng}:3

Ta có

|  |  |
| --- | --- |
| Compa 🡪 Tập trắng | 3/3 = 100% |
| Tập trắng 🡪 Compa | 3/5 = 60% |

Với min\_conf = 80% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

**Compa 🡪 Tập trắng**

* Xét tập phổ biến {Bút bi; Tập trắng}:4 có min\_conf = 80%

{Bút bi}: 5, {Tập trắng}: 5, {Bút bi; Tập trắng}: 4

Ta có

|  |  |
| --- | --- |
| Bút bi🡪 Tập trắng | 4/5 = 80% |
| Tập trắng 🡪 Bút bi | 4/5 = 80% |

Với min\_conf = 80% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

**Bút bi 🡪 Tập trắng, Tập trắng 🡪 Bút bi**

* Xét tập phổ biến {Compa; Bút bi; Tập trắng}: 3 có min\_conf = 80%

{Compa}: 3, {Bút bi}: 5, {Tập trắng}: 5, {Compa; Bút bi}: 3, {Compa; Tập trắng}: 3; {Bút bi; Tập trắng}: 4, {Compa; Bút bi; Tập trắng}: 3

Ta có

|  |  |
| --- | --- |
| Compa 🡪 {Bút bi; Tập trắng} | 3/3 = 100% |
| Bút bi 🡪 {Tập trắng; Compa} | 3/5 = 60% |
| Tập trắng 🡪 {Bút bi; Compa} | 3/5 = 60% |
| {Bút bi; Tập trắng} 🡪 Compa | 3/4 =75% |
| {Tập trắng; Compa} 🡪 Bút bi | 3/3 = 100% |
| {Bút bi; Compa} 🡪 Tập trắng | 3/3 = 100% |

Với min\_conf = 80% dựa vào bảng trên ta có luật kết hợp thỏa yêu cầu là:

Compa 🡪 {Bút bi; Tập trắng},

{Tập trắng; Compa} 🡪 Bút bi, {Bút bi; Compa} 🡪 Tập trắng

1. (Cơ bản) CSDL về Nhân viên được cho trong bảng sau:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Giới tính (GT) | Tuổi (T) | Năng lực làm việc (NL) | Đã lập gia đình (LGD) | Thu nhập (TN) | Thăng chức (TC) |
| 1 | Nữ | 20..25 | Giỏi | Rồi | Rất cao | Có |
| 2 | Nam | 20..25 | Khá | Chưa | Khá | Không |
| 3 | Nữ | 26..30 | Giỏi | Chưa | Khá | Có |
| 4 | Nữ | 31..40 | T.Bình | Chưa | T.Bình | Có |
| 5 | Nam | 26..30 | T.Bình | Rồi | Rất cao | Không |
| 6 | Nữ | 26..30 | Khá | Chưa | Cao | Không |
| 7 | Nữ | 31..40 | Khá | Chưa | T.Bình | Không |
| 8 | Nam | 26..30 | Khá | Rồi | Cao | Có |
| 9 | Nữ | >40 | Giỏi | Rồi | T.Bình | Không |
| 10 | Nữ | 26..30 | Giỏi | Chưa | Khá | Có |

* 1. Tìm tập phổ biến bằng giải thuật Apriori
* Với min\_sup = 30% và tổng số giao dịch là 10 => Tần số xuất hiện tối thiểu của phần tử để thỏa min\_sup là 3.
* Tập các ứng viên 1 phần tử và tần số xuất hiện của nó (support count) tương ứng như sau:

C1 = {

{20..25}:2,

{26..30}:5,

{31..40}:2,

{>40}: 1,

Giỏi: 4,

Khá: 4,

Trung bình: 2,

Có: 5,

Không: 5

}

* + - * Các tập phổ biến thỏa min\_sup 3 là: {26..30}, Giỏi, Khá, Có, Không
      * Các tập phổ biến 1 phần tử (tập các ứng viên thoải min\_support\_count):

L1 = {

{26..30},

{Giỏi},

{Khá},

{Có},

{Không}

}

Tập các ứng viên 2 phần tử và tần số xuất hiện của nó

C2 = {

{26..30, Giỏi}: 2,

{26..30, Khá}: 2,

{26..30, Có}:3,

{26..30, Không}: 2,

{Giỏi, Có}:3,

{Giỏi, Không}:1,

{Khá, Có}:1,

{Khá, Không}:3

}

* + - * Các tập phổ biến thỏa min\_sup 3 là: {26..30, Có}, {Giỏi, Có}, {Khá, Không}
      * Các tập phổ biến 2 phần tử (tập các ứng viên thoải min\_support\_count):

L2 = {{26..30, Có}, {Giỏi, Có}, {Khá, Không} }

* Tập các ứng viên 3 phần tử và tần số xuất hiện của nó

C3 = {{26..30, Giỏi, Có}: 2}

* + - * Các tập phổ biến thỏa min\_sup 3 là: không có
      * Các tập phổ biến 3 phần tử (tập các ứng viên thoải min\_support\_count):

L3 =

* 1. Tìm luật có về phải là thăng chức
* Xét tập phổ biến {26..30, Có} có min\_conf = 80%

{26..30}: 5, {Có}: 5, {26..30, Có}: 3

Ta có

|  |  |
| --- | --- |
| 26..30🡪Có | 3/5= 60% |

Với min\_conf = 80% dựa vào bảng trên ta không có luật kết hợp thỏa yêu cầu min\_conf. Nên tập {26..30, Có} không có luật kết hợp có về phải là thuộc tính Thăng chức.

* Xét tập phổ biến {Giỏi, Có} có min\_conf = 80%

{Giỏi}: 4, {Có}: 5, {Giỏi, Có}: 3

Ta có

|  |  |
| --- | --- |
| Giỏi🡪Có | 3/4= 75% |

Với min\_conf = 80% dựa vào bảng trên ta không có luật kết hợp thỏa yêu cầu min\_conf. Nên tập {Giỏi, Có} không có luật kết hợp có về phải là thuộc tính Thăng chức.

* Xét tập phổ biến {Khá, Không} Không min\_conf = 80%

{Khá}: 4, {Không}: 5, {Khá, Không}: 3

Ta Không

|  |  |
| --- | --- |
| Khá🡪Không | 3/4= 75% |

Với min\_conf = 80% dựa vào bảng trên ta không có luật kết hợp thỏa yêu cầu min\_conf. Nên tập {Khá, Không} không có luật kết hợp có về phải là thuộc tính Thăng chức.

* Không có luật kết hợp cần tìm với min\_conf = 80%

1. (Cơ bản) Cho bảng dữ liệu ở một công ty có các giao dịch như sau:

|  |  |
| --- | --- |
| TID | Items |
| T1 | A, B, C, D |
| T2 | A, B, C |
| T3 | A, B, C |
| T4 | A, B, D |
| T5 | A, B |
| T6 | A, C, D |
| T7 | A, D |
| T8 | B, C, D |

* 1. Tìm tập phổ biến bằng giải thuật Apriori
* Với min\_sup \_count > 1
* Tập các ứng viên 1 phần tử và tần số xuất hiện của nó (support count) tương ứng như sau:

C1 = {A:7, B:6, C:5, D:5}

* + - * Các tập phổ biến thỏa min\_sup 2 là: A, B, C, D
      * Các tập phổ biến 1 phần tử (tập các ứng viên thoải min\_support\_count):

L1 = {{A}, {B}, {C}, {D}}

* Tập các ứng viên 2 phần tử và tần số xuất hiện của nó

C2 = {{A, B}: 5, {A, C}: 4, {A, D}: 3, {B, C}: 4, {B, D}: 3, {C, D}: 3}

* + - * Các tập phổ biến thỏa min\_sup 2 là: {A, B}, {A, C}, {A, D}, {B, C}, {B, D}, {C, D}
      * Các tập phổ biến 2 phần tử (tập các ứng viên thoải min\_support\_count):

L2 = {{A, B}, {A, C}, {A, D}, {B, C}, {B, D}, {C, D}}

* Tập các ứng viên 3 phần tử và tần số xuất hiện của nó

C3 = {{A, B, C}: 3, {A, B, D}: 2, {A, C, D}: 2, {B, C, D}: 2}

* + - * Các tập phổ biến thỏa min\_sup 2 là: {A, B, C}, {A, B, D}, {A, C, D}, {B, C, D}
      * Các tập phổ biến 3 phần tử (tập các ứng viên thoải min\_support\_count):

L3 = {{A, B, C}, {A, B, D}, {A, C, D}, {B, C, D}}

* Tập các ứng viên 3 phần tử và tần số xuất hiện của nó

C4 = {{A, B, C, D}: 1

* + - * Không có tập phổ biến 4 phần tử thỏa min\_sup 2 là
      * Các tập phổ biến 4 phần tử (tập các ứng viên thoải min\_support\_count):

L4 =

* 1. Xây dựng cây FP
* Xây dựng cây FP
  + - Tập phổ biến 1 phần tử và tần suất xuất hiện của nó

|  |  |
| --- | --- |
| item | supp |
| A | 7 |
| B | 6 |
| C | 5 |
| D | 5 |

* + - Tập phổ biến 1 phần tử và tần xuất thỏa điều kiện ≥ min\_supp và đã được sắp xếp giảm dần

|  |  |
| --- | --- |
| item | supp |
| A | 7 |
| B | 6 |
| C | 5 |
| D | 5 |

* + - Điền cột Order Frequent Items vào

|  |  |  |
| --- | --- | --- |
| TID | items | Order Frequent Items |
| T1 | A, B, C, D | A, B, C, D |
| T2 | A, B, C | A, B, C |
| T3 | A, B, C | A, B, C |
| T4 | A, B, D | A, B, D |
| T5 | A, B | A, B |
| T6 | A, C, D | A, C, D |
| T7 | A, D | A, D |
| T8 | B, C, D | B, C, D |
| T1 | A, B, C, D | A, B, C, D |
| T2 | A, B, C | A, B, C |

* + - Vẽ cây FP
    - Header Table

|  |  |
| --- | --- |
| item | supp |
| A | 7 |
| B | 6 |
| C | 5 |
| D | 5 |

* Xây dựng cơ sở điều kiện cho mỗi mục phổ biến

|  |  |
| --- | --- |
| Item | Conditional Pattern Base |
| D | {A, B, C}: 1, {A}: 1, {A, B}: 1, {A, C}: 1, {B, C}: 1 |
| C | {A, B}: 3, {A}: 1, {B}: 1 |
| B | {A}: 5 |
| A | {} |

* Xây dựng cây FP điều kiện

|  |  |  |
| --- | --- | --- |
| item | Conditional Pattern Base | Conditional FP-Tree (supp ≥2) |
| D | {A, B, C}: 1, {A}: 1, {A, B}: 1, {A, C}: 1, {B, C}: 1 | {{A:3, B:2, C:1}, {A:1, C:1}, {B:1, C:1}} |D |
| C | {A, B}: 3, {A}: 1, {B}: 1 | {{A:4, B:3}, {B:1} |C |
| B | {A}: 5 | {{A}:5} |B |

* Kết quả các mẫu thường xuyên

|  |  |  |
| --- | --- | --- |
| item | Conditional FP-Tree | Frequent Pattern Buns |
| D | {{A:3, B:2, C:1}, {A:1, C:1}, {B:1, C:1}} |D | {D}: 5, {A, D}: 4, {B, D}: 3, {C, D}: 3,  {A, B, D}: 2, {A, C, D}: 2, {B, C, D}: 2 |
| C | {{A:4, B:3}, {B:1} |C | {C}: 5, {A, C}: 4, {B, C}: 4, {A, B, C}: 3 |
| B | {{A}:5} |B | {B}: 6, {A, B}: 5 |
| A | {} | {A: 7} |
| D | {{A:3, B:2, C:1}, {A:1, C:1}, {B:1, C:1}} |D | {D}: 5, {A, D}: 4, {B, D}: 3, {C, D}: 3,  {A, B, D}: 2, {A, C, D}: 2, {B, C, D}: 2 |

Lấy những phần tử có 2 phần tử trở lên và thỏa min\_supp ta được các tập hợp sau đây:

L2 = {{A, B}, {A, C}, {A, D}, {B, C}, {B, D}, {C, D}}

L3 = {{A, B, C}, {A, B, D}, {A, C, D}, {B, C, D}}