**Cloud Computing for Data Analysis**

**Group Activity 03 – FP-Tree**

**Chapter 6 Exercise 8**

Build and mine FP-Tree using the data below (Min Support 3)

Table 6.24. Example of market basket transactions.

|  |  |
| --- | --- |
| Transaction ID | Items Bought |
| 1 | {a, b, d, e} |
| 2 | {b, c, d} |
| 3 | {a, b, d, e} |
| 4 | {a, c, d, e} |
| 5 | {b, c, d, e} |
| 6 | {b, d, e} |
| 7 | {c, d} |
| 8 | {a, b, c} |
| 9 | {a, d, e} |
| 10 | {b, d} |

**Answer:**

**Support of each item**

|  |  |
| --- | --- |
| Item | Support |
| a | 5 |
| b | 7 |
| c | 5 |
| d | 9 |
| e | 6 |

**Reordering the transactions as per the support order**

Hence we will be getting below table:-

|  |  |  |
| --- | --- | --- |
| Transaction ID | Items Bought | **Reordered set** |
| 1 | {a, b, d, e} | **d, b, e, a** |
| 2 | {b, c, d} | **d, b, c** |
| 3 | {a, b, d, e} | **d, b, e, a** |
| 4 | {a, c, d, e} | **d, e, a, c** |
| 5 | {b, c, d, e} | **d, b, e, c** |
| 6 | {b, d, e} | **d, b, e** |
| 7 | {c, d} | **d, c** |
| 8 | {a, b, c} | **b, a, c** |
| 9 | {a, d, e} | **d, e, a** |
| 10 | {b, d} | **d, b** |
|  |  |  |

**The Complete Frequent Pattern Growth Tree- Diagram:**

1

9

6

4

2

1

1

1

1

2

2

1

1

**Mining FP Tree for Frequent Item Sets**

Now mining the tree whose support is greater than 3.

1.

Suffix pattern – a  
Support = 5

Prefix Paths

(d, b, e, a, 2)

(d, e, a, 2)

(b, a, 1)

|  |  |
| --- | --- |
| D | 4 |
| E | 4 |
| B | 3 |

**Reordered Set**

{d, e, b}

{d, e}

{b}

**Tree:**

4

1

2

4

**Frequent Item Sets for a**

{d, e, a, 4}

{d, a, 4}

{b, a, 3}

**2.**

**Suffix Pattern – b, Support = 7**

Prefix path

{d, b, 6}

|  |  |
| --- | --- |
| d | 6 |

Reordered Set

{d}

6

**Frequent Item Set**

{d, b, 6}

**3.**

**Suffix Pattern – c, Support = 5**

Prefix Paths

{d, b, e, c, 1}

{d, b, c, 1}

{d, e, a, c, 1}

{d, c, 1}

{b, a, c, 1}

|  |  |
| --- | --- |
| d | 4 |
| b | 3 |
| a | 2 |
| e | 2 |

**Reordered Set**

{d, b, e}

{d, b}

{d, a, e}

{d}

{b, a}

**Tree**

4

1

1

2

1

1

1

**Frequent Item Sets**

{d, c, 4}

{b, c, 3}

**4.**

**Suffix Pattern – d, Support = 9**

Prefix Path – No prefix path for d because d is immediately next to root element

**No Frequent Item Set**

**5.**

**Suffix Pattern – e, Support = 6**

Prefix Paths

{d, b, e, 4}

{d, e, 2}

|  |  |
| --- | --- |
| d | 6 |
| b | 4 |

Reordered item set

{d, b}

{d}

**Tree**

6

4

**Frequent Item Sets**

{d, b, e, 4}

{d, e, 6}

{b, e, 4}