

Solutions Frequent Pattern Mining 2026

Exercise 1: Frequent Item Set Mining

(a) Level 1:

candidate	support	frequent?
<i>A</i>	3	✓
<i>B</i>	5	✓
<i>C</i>	6	✓
<i>D</i>	5	✓
<i>E</i>	2	✓
<i>F</i>	1	✗

Level 2:

candidate	support	frequent?
<i>AB</i>	3	✓
<i>AC</i>	2	✓
<i>AD</i>	2	✓
<i>AE</i>	0	✗
<i>BC</i>	4	✓
<i>BD</i>	3	✓
<i>BE</i>	0	✗
<i>CD</i>	4	✓
<i>CE</i>	1	✗
<i>DE</i>	1	✗

Level 3:

candidate	support	frequent?
<i>ABC</i>	2	✓
<i>ABD</i>	2	✓
<i>ACD</i>	1	✗
<i>BCD</i>	2	✓

ABE is not a level-3 candidate, because its level-2 subsets AE and BE are not frequent. Notice that since AE is not frequent, AB and AE won't even be combined into a "pre-candidate" by the algorithm. For similar reasons, BCE and CDE are not level-3 candidates.

$ABCD$ is not a level-4 candidate because its level-3 subset ACD is not frequent.

(b) The closed frequent itemsets are

itemset	support
AB	3
B	5
C	6
D	5
E	2
ABC	2
ABD	2
BC	4
BD	3
CD	4
BCD	2

(c) The maximal frequent item sets are: E , ABC , ABD , and BCD .

(d)

$$\text{confidence}(A \rightarrow C) = \frac{s(AC)}{s(A)} = \frac{2}{3}$$

$$\text{lift}(A \rightarrow C) = \frac{s(AC) \times |db|}{s(A) \times s(C)} = \frac{2 \times 8}{3 \times 6} = \frac{16}{18} = \frac{8}{9},$$

where $|db|$ denotes the number of transactions in the data base. Since the lift is smaller than 1, the rule would not be considered interesting according to this measure.

The formula for lift used above can be derived from its definition as follows:

$$\text{lift}(X \rightarrow Y) = \frac{P(Y | X)}{P(Y)} = \frac{P(X, Y)}{P(X)P(Y)} = \frac{s(XY)/|db|}{s(X)/|db| \times s(Y)/|db|} = \frac{s(XY) \times |db|}{s(X) \times s(Y)}$$

Note that we write $s(XY)$ as a shorthand of $s(X \cup Y)$.

Exercise 2: Frequent Sequence Mining

(a) Level 1:

candidate	support	frequent?
A	5	✓
U	4	✓
H	3	✓
B	3	✓

Level 2:

candidate	support	frequent?
AA	2	✗
AU	4	✓
AH	3	✓
AB	2	✗
UA	2	✗
UU	3	✓
UH	3	✓
UB	2	✗
HA	2	✗
HU	3	✓
HH	0	✗
HB	1	✗
BA	1	✗
BU	0	✗
BH	0	✗
BB	0	✗

Level 3:

candidate	support	frequent?
AUU	3	✓
AUH	3	✓
AHU	3	✓
UUU	0	✗
UUH	0	✗
UHU	3	✓
HUU	0	✗

Level 4:

candidate	support	frequent?
AUHU	3	✓

There is no level 5 candidate. We can combine AUHU with itself to generate pre-candidate AUHUU but this contains the infrequent subsequence HUU.

(b) AUHU and B are maximal.

(c) A, B, AU, AUHU.

Exercise 3: Anti-monotonicity

No, for example A is a subsequence of AB, but AB occurs twice as a subsequence in ABB, and A just once.

Alternative definition of “distinct occurrence”: $\phi_1(i) \neq \phi_2(i)$ for all positions i in the pattern sequence.

Exercise 4: Transitivity of the subsequence relation

If the subsequence relation is transitive, then from $\mathbf{r}_1 \subseteq \mathbf{r}_2$, and $\mathbf{r}_2 \subseteq \mathbf{s}$ (where $\mathbf{s} \in \mathbf{D}$) it follows that $\mathbf{r}_1 \subseteq \mathbf{s}$ and therefore that the support of $\mathbf{r}_1$ is at least as big as the support of $\mathbf{r}_2$.

It is given that $\mathbf{q} \subseteq \mathbf{r}$, and $\mathbf{r} \subseteq \mathbf{s}$. So there exist one-to-one mappings ϕ_{qr} and ϕ_{rs} that preserve the labels and order. Now define the mapping $\phi_{qs}(i) = \phi_{rs}(\phi_{qr}(i))$. This mapping has the following properties:

1. $\mathbf{q}[i] = \mathbf{s}[\phi_{qs}(i)]$ because $\mathbf{q}[i] = \mathbf{r}[\phi_{qr}(i)]$ and $\mathbf{r}[\phi_{qr}(i)] = \mathbf{s}[\phi_{rs}(\phi_{qr}(i))]$.
2. $i < j \Rightarrow \phi_{qs}(i) < \phi_{qs}(j)$, because $i < j \Rightarrow \phi_{qr}(i) < \phi_{qr}(j)$ and $\phi_{qr}(i) < \phi_{qr}(j) \Rightarrow \phi_{rs}(\phi_{qr}(i)) < \phi_{rs}(\phi_{qr}(j))$.

Also, since ϕ_{qr} is one-to-one, and ϕ_{rs} is one-to-one, it follows that ϕ_{qs} is one-to-one as well. Hence, the mapping ϕ_{qs} satisfies all the required properties.

Note: Actually, it follows from $i < j \Rightarrow \phi(i) < \phi(j)$ that ϕ is one-to-one, so the latter restriction is superfluous. The “proof” above that ϕ_{qs} is one-to-one is therefore also superfluous: we have already shown that ϕ_{qs} preserves the order. Thanks to Max Hessey and Casper Hagenars for noting this.

Exercise 5: Variations on a theme

We define the subsequence relation as follows: $S^1 = (X_1X_2 \dots X_k)$ is a subsequence of $S^2 = (Y_1Y_2 \dots Y_m)$ (denoted $S^1 \preceq S^2$) if there exists a mapping

$$\phi : [1, k] \rightarrow [1, m],$$

such that

1. $X_i \subseteq Y_{\phi(i)}$, and
2. $i < j \Rightarrow \phi(i) < \phi(j)$.

We define support as:

$$\text{sup}(R) = |\{S^i \in D : R \preceq S^i\}|$$

To show that the anti-monotonicity property holds, it suffices to show that the subsequence relation is transitive, that is,

$$S^1 \preceq S^2 \text{ and } S^2 \preceq S^3 \Rightarrow S^1 \preceq S^3$$

So assume that $S^1 \preceq S^2$, and $S^2 \preceq S^3$. This means there exist mappings ϕ_{12} and ϕ_{23} that satisfy the subset and order constraints. Now define the mapping $\phi_{13}(i) = \phi_{23}(\phi_{12}(i))$. This mapping has the following properties:

1. $S^1_i \subseteq S^3_{\phi_{13}(i)}$ because $S^1_i \subseteq S^2_{\phi_{12}(i)}$ and $S^2_{\phi_{12}(i)} \subseteq S^3_{\phi_{23}(\phi_{12}(i))}$.
2. $i < j \Rightarrow \phi_{13}(i) < \phi_{13}(j)$, because $i < j \Rightarrow \phi_{12}(i) < \phi_{12}(j)$ and $\phi_{12}(i) < \phi_{12}(j) \Rightarrow \phi_{23}(\phi_{12}(i)) < \phi_{23}(\phi_{12}(j))$.

Therefore, S^1 is a subsequence of S^3 . From this the anti-monotonicity property

$$S^1 \preceq S^2 \Rightarrow \text{sup}(S^1) \geq \text{sup}(S^2)$$

follows.

Notice that in this proof we left out the requirement that ϕ is a one-to-one mapping. We explained in the note at exercise 5 why it is superfluous.