

Anonymous

Nalezni všechny časté itemsety (frequent itemsets) délky 1, jejichž podpora je minimálně 2.
První řádek považuj za názvy sloupců.

354C31,DBCEEF,6CE255,199D26,D59553,331981,869F50
B2895C,B0FAAD,9ED0E8,D11884,269963,ADC8A2,999999
1520D1,49E95E,A6463C,D11884,269963,ADC8A2,999999
0BEA8D,B0FAAD,A6463C,D11884,269963,ADC8A2,999999
7EA60D,B0FAAD,C5B9CC,D11884,269963,420666,999999

Gemini

Časté itemsety délky 1 s minimální podporou 2:

Názvy sloupců:

1. 354C31
2. DBCEEF
3. 6CE255
4. 199D26
5. D59553
6. 331981
7. 869F50
8. B2895C
9. B0FAAD
10. 9ED0E8
11. D11884
12. 269963
13. ADC8A2
14. 999999

Časté itemsety:

- **B0FAAD** (podpora: 4)
- **D11884** (podpora: 4)
- **269963** (podpora: 4)
- **999999** (podpora: 4)

Výpočet:

Pro výpočet častých itemsetů délky 1 s minimální podporou 2 můžeme použít algoritmus Apriori. Algoritmus Apriori prochází databázi transakcí a počítá podporu pro každý kandidátní itemset. Kandidátní itemset je sada položek, která se vyskytuje v minimálním počtu transakcí (v tomto případě 2). Algoritmus Apriori se zastaví, když nenajde žádné další kandidátní itemsety s požadovanou minimální podporou.

V tomto případě můžeme algoritmus Apriori spustit ručně. Prvním krokem je identifikace všech kandidátních itemsetů délky 1. V tomto případě existuje 14 kandidátních itemsetů, jeden pro každý sloupec v databázi.

Dále spočítáme podporu pro každý kandidátní itemset. Podpora itemsetu je počet transakcí, ve kterých se itemset vyskytuje. V tomto případě můžeme podporu spočítat ručně nebo pomocí nástroje pro analýzu dat.

Kandidátní itemset s podporou menší než 2 je vyloučen. V tomto případě jsou vyloučeny všechny kandidátní itemsety kromě B0FAAD, D11884, 269963 a 999999.

Zbylé itemsety (B0FAAD, D11884, 269963 a 999999) jsou časté itemsety délky 1 s minimální podporou 2.

Další kroky:

Můžeme dále analyzovat nalezené časté itemsety. Například můžeme:

- Vypočítat podporu a důvěru asociací mezi itemsety.
- Vizualizovat nalezené asociace pomocí grafu.
- Využít nalezené asociace k predikci budoucího chování.

Nástroje pro analýzu dat:

Existuje mnoho nástrojů pro analýzu dat, které lze použít k výpočtu častých itemsetů. Mezi populární nástroje patří:

- Python knihovna `apriori`
- R balíček `arules`
- Weka
- RapidMiner

Reference:

- Apriori algorithm: https://en.wikipedia.org/wiki/Apriori_algorithm
- Frequent itemset mining: [neplatná adresa URL byla odstraněna]