

Esercizi algoritmo Apriori

Eseguire l'algoritmo a priori.

Dataset

AE
DE
BDE
ACDE
ABEF
AB
BCF
BEF
ADF
ACDF
minsup = 20%
minconf = 30%

K=1

A 6
B 5
C 3
D 5
E 6
F 5

Tutti
maggiori
di minsup

↓
Non elimino
nessun
elemento

K=2

AB 2
AC 2
AD 3
AE 3
AF 3
BC 1
BD 1
BE 3
BF 3
CD 2
CE 1
CF 2
DE 3
DF 2
EF 2

Non considero i doppiomi tipo AA perché in un set ogni elemento compare max 1 volta; inoltre, l'ordine non conta: AB=BA
Genero tutte le coppie perché non ho eliminato nessun elemento nell'iterazione precedente

K=3

ABC
ABD
ABE 1
ABF 1
ACD 2
ACE
ACF 1
ADE 1
ADF 2
AEF 1
BEF 2
CDF 1
CDE
DEF 0

Uniamo sottosequenze di lunghezza K-1 (=2)

con almeno K-2 (=1) elemento in comune

Effettuiamo il pruning:

ABC generato da AB e AC →

debiamo controllare AB → non frequente

→ lo eliminiamo

K=4

ACDF

ACD + ADF

MA $\sigma(ACF) = 1$

⇒ pruning

Abbiamo finito, l'algoritmo si ferma; abbiamo ottenuto i tre insiemi evidenziati in verde

Abbiamo ottenuto le seguenti regole per ACD:

$AC \rightarrow D$ $c = \frac{\sigma(ACD)}{\sigma(AC)} = \frac{2}{2} = 1$; $L = \frac{P(ACD)}{P(AC)P(D)} = \frac{2/10}{2/10 \cdot 5/10} = 2$

$AD \rightarrow C$ $c = \frac{2}{3} = 0,6$

$CD \rightarrow A$ $c = \frac{2}{2} = 1$

$A \rightarrow CD$ $c = \frac{2}{4} = 0,5$

$C \rightarrow AD$ $c = \frac{2}{3} = 0,6$

$D \rightarrow CA$ $c = \frac{2}{3} = 0,6$

Confidence di $X \rightarrow Y = \frac{\sigma(X,Y)}{\sigma(X)}$
 (= probabilità condizionata)

lift di $X \rightarrow Y = \frac{P(X \wedge Y)}{P(X)P(Y)}$

Domande: con n diversi pattern, qual è il numero di tutti i possibili pattern? $2^n - 1$ (perché non consideriamo l'insieme vuoto)

Qual è il numero di tutti i possibili pattern di lunghezza k ? $\binom{n}{k}$ COEFFIC. BINOMIALI

In un pattern con k elementi, qual è il numero di tutte le possibili regole? $2^k - 2$ (perché non consideriamo $X \rightarrow \phi$ o $\phi \rightarrow Y$)

Altro dataset

BE
ER
BER
ACR
ABEF
AB
BFR
BEF
AFR
ACFR
minsup = 20
minconf = 70

K=1
A 5
B 6
C 2
E 5
F 5
R 6

K=2
AB 2
AC 2
AE 1
AF 3
AR 3
BC 0
BE 4
BF 3
BR 2
CE 0
CF 1
CR 2
EF 2
ER 2
FR 3

K=3
ABC
ABF 1
ABR \emptyset
ABE
ACR 2
AFR 2
BEF 2
BF R 1
BRC
EFR \emptyset

Regole BEF:

$BE \rightarrow F$ $c = 2/4$

$BF \rightarrow E$ $c = 2/3$

$EF \rightarrow B$ $c = 2/2$

$B \rightarrow EF$ $c = 2/6$

$E \rightarrow BF$ $c = 2/5$

$F \rightarrow BE$ $c = 2/5$

lift = $\frac{P(X \wedge Y)}{P(X)P(Y)} = \frac{2/10}{2/10 \cdot 4/10} = \frac{5}{3}$