Megan Cusey

Unit 3 Homework

Summer 2019

Pages 243-271

6.6 (a only – Apriori Solution)

A database has five transactions. Let *min\_sup* = 60% and *min\_conf* = 80%

|  |  |
| --- | --- |
| **TID** | **Items\_bought** |
| T100 | {M, O, N, K, E, Y} |
| T200 | {D, O, N, K, E, Y} |
| T300 | {M, A, K, E} |
| T400 | {M, U, C, K, Y} |
| T500 | {C, O, O, K, I, E} |

1. Find all frequent itemset using Apriori and FP-growth, respectively. Compare the efficiency of the two mining processes.
2. lhs rhs support confidence lift
3. [1] {} => {K} 1.0 1 1.00
4. [2] {M} => {K} 0.6 1 1.00
5. [3] {O} => {E} 0.6 1 1.25
6. [4] {O} => {K} 0.6 1 1.00
7. [5] {Y} => {K} 0.6 1 1.00

R Script

##load arules

library(arules)

T100 <- c('M','O','N','K','E','Y')

T200 <- c('D','O','N','K','E','Y')

T300 <- c('M','A','K','E')

T400 <- c('M','U','C','K','Y')

T500 <- c('C','O','O','K','I','E')

data <- list(T100,T200,T300,T400,T500)

data2 <-as(data,"transactions")

inspect(data2)

##Convert transactions to transaction ID Lists

tl <- as(data2,"tidLists")

inspect(tl)

summary(tl)

## Pass in the min support and min confidence

rules <- apriori(tl, parameter =

list(supp=.6, conf =.8))

rules <- sort(rules, by="confidence", decreasing=TRUE)

option(digits=2)

inspect(rules[1:5])

## Remove Data List

rm(list=ls())