# R Code (Association Rules)

##Association rules with color dataset

install.packages("arules")

library(arules)

#Read dataset

color <- read.csv("C:/MA 299/R/color.csv")

#factor all predictors that use numbers to represent different levels

color$Red <- factor(color$Red)

color$White <- factor(color$White)

color$Blue <- factor(color$Blue)

color$Orange <- factor(color$Orange)

color$Green <- factor(color$Green)

color$Yellow <- factor(color$Yellow)

#Association rules (default minimum support = 0.10

#and minimum confidence = 0.80)

apriori.parameter=list(support = 0.10, confidence = 0.50)

rules = apriori(color, parameter = apriori.parameter,

appearance = list(both=c("Red=1", "White=1", "Blue=1", "Orange=1", "Green=1", "Yellow=1"),

default = "none"))

summary(rules)

#Sort rules by confidence

rules.sorted <- sort(rules, by="confidence")

inspect(rules.sorted[1:10])

#Sort rules by Lift ratio

rules.sorted <- sort(rules, by="lift")

inspect(rules.sorted[1:10])

##Association rules with CourseTopics dataset

library(arules)

#Read dataset

course <- read.csv("C:/MA 299/R/CourseTopics.csv")

#factor all predictors that use numbers to represent different levels

course$Intro <- factor(course$Intro)

course$DataMining <- factor(course$DataMining)

course$Survey <- factor(course$Survey)

course$Cat.Data <- factor(course$Cat.Data)

course$Regression <- factor(course$Regression)

course$Forecast <- factor(course$Forecast)

course$DOE <- factor(course$DOE)

course$SW <- factor(course$SW)

#Association rules (default minimum support = 0.10

#and minimum confidence = 0.80)

#Note: If we set minimum support = 0.05, the number of generated rules = 0

apriori.parameter=list(support = 0.01, confidence = 0.50)

rules = apriori(course, parameter = apriori.parameter,

appearance = list(both=c("Intro=1", "DataMining=1", "Survey=1", "Cat.Data=1", "Regression=1", "Forecast=1", "DOE=1", "SW=1"),

default = "none"))

summary(rules)

#Sort rules by Lift ratio

rules.sorted <- sort(rules, by="lift")

inspect(rules.sorted[1:10])

#Association rules for classification

library(arules)  
BankA <- read.csv("C:/MA 299/R/UniversalBankAssociationRules.csv")  
#factor all predictors that use numbers to represent different levels  
BankA$Education <- factor(BankA$Education)  
BankA$Securities <- factor(BankA$Securities)  
BankA$CD <- factor(BankA$CD)  
BankA$Online <- factor(BankA$Online)  
BankA$CreditCard <- factor(BankA$CreditCard)  
BankA$PersonalLoan <- factor(BankA$PersonalLoan)  
summary(BankA)

#Set up apriori.appearance and apriori.parameter

apriori.appearance = list(rhs = c("PersonalLoan=0",   
 "PersonalLoan=1"))  
apriori.parameter = list(support = 0.05, confidence = 0.90)  
rules = apriori(BankA, parameter = apriori.parameter,  
 appearance = apriori.appearance)  
summary(rules)

#Sort rules by confidence

rules.sorted <- sort(rules, by="confidence")  
inspect(rules.sorted[1:5])

#Generate scatter plot

library(arulesViz)  
plot(rules)