**DAY 3 LAB EXPERIMENTS**

**NAME :K.Siva Naga Manoj Kumar**

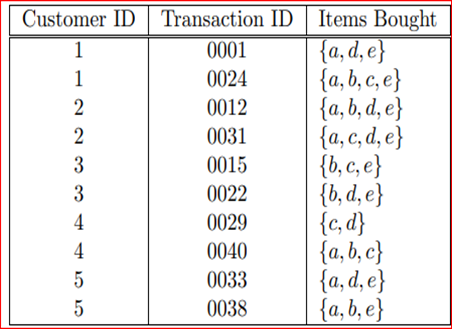
**REG NO.:192111630**

**SUB CODE :CSA1622**

**SUB NAME:DWDM**

**1.Consider the data set and perform the Apriori Algorithm and FP algorithm support:3 and**

**confidence=50%**



Input:

@relation dataset

@attribute a{true,false}

@attribute b{true,false}

@attribute c{true,false}

@attribute d{true,false}

@attribute e{true,false}

@data

true false false true true

true true true false true

true true false true true

true false true true true

false true true false true

false true false true true

false false true true false

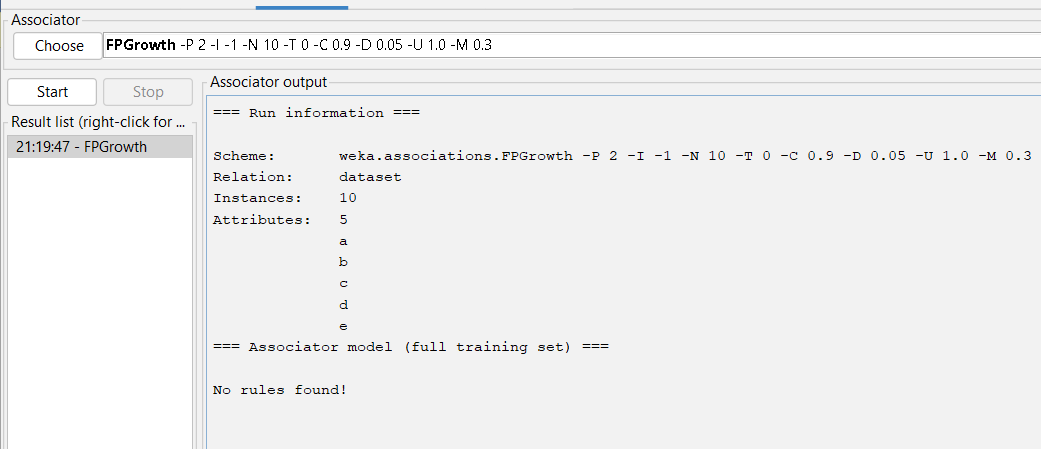
true true true false false

true false false true true

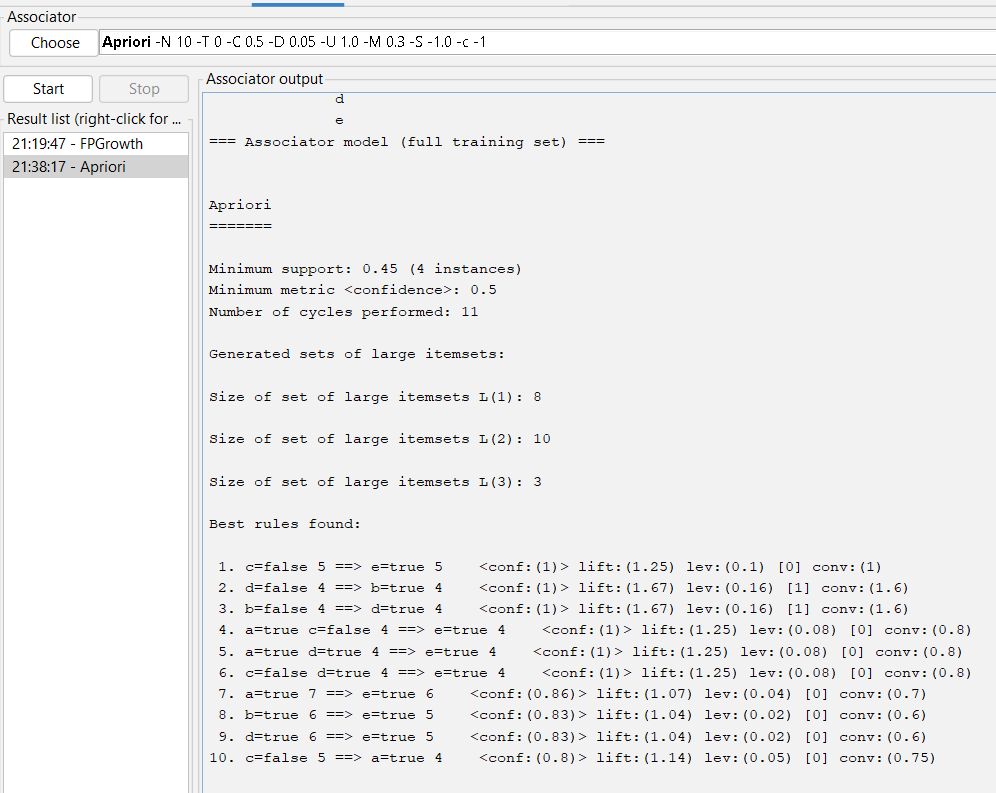
true true false false true

output:

FPGROWTH:



APRIORI ALGORITHM:



**2.Consider the data set and perform the Apriori Algorithm and FP algorithm support:3 and confidence=50%**

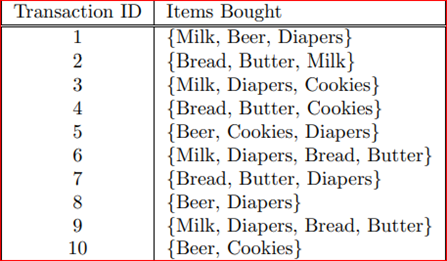
Consider the market basket transactions shown in the above table.

(a) What is the maximum number of association rules that can be extracted

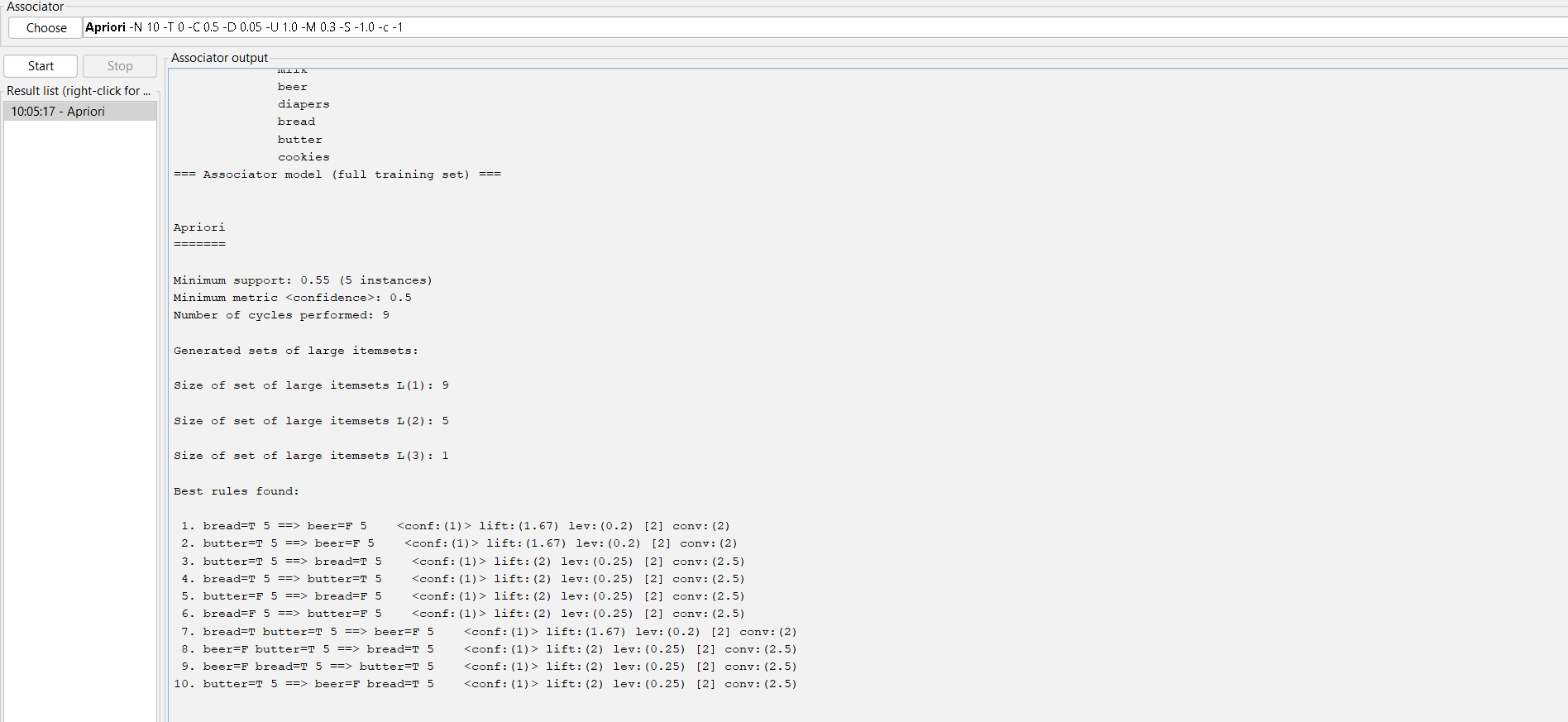
from this data (including rules that have zero support)?

(b) What is the maximum size of frequent itemsets that can be extracted

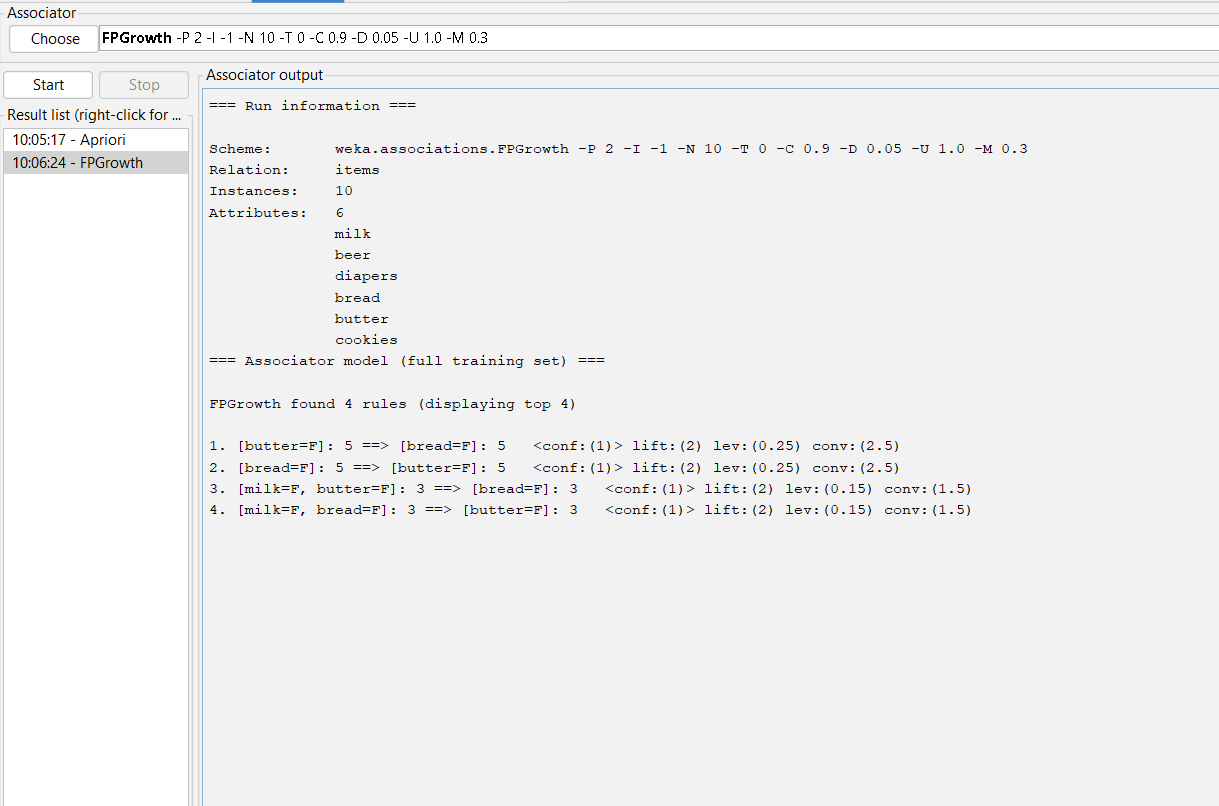
(assuming minsup > 0)?



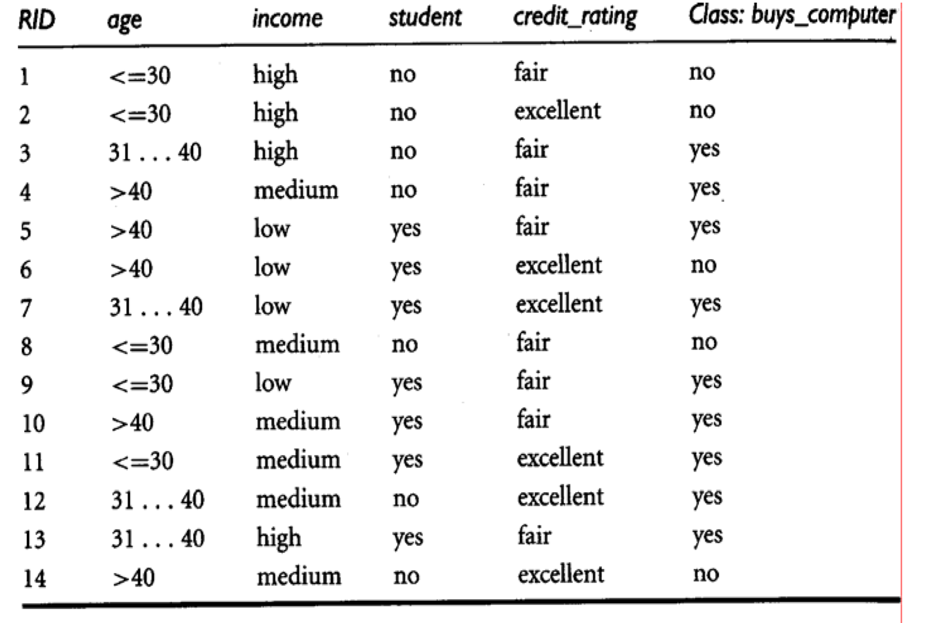
Apriori algorithm:



Fp growth algorithm:



3.Bayes classification and descion tree (using training and test data)



Input:

@relation decision\_tree

@attribute age{young,middle,old}

@attribute income{low,medium,high}

@attribute student{yes,no}

@attribute Creit\_rating{fair,excellent}

@attribute class{yes,no}

@data

young high no fair no

young high no excellent no

middle high no fair yes

old medium no fair yes

old low yes fair yes

old low yes excellent no

middle low yes excellent yes

young medium no fair no

young low yes fair yes

old medium yes fair yes

young medium yes excellent yes

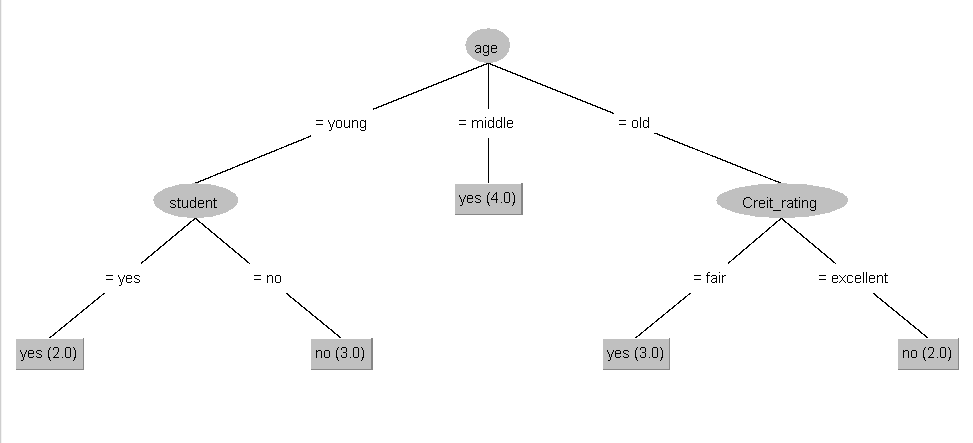
middle medium no excellent yes

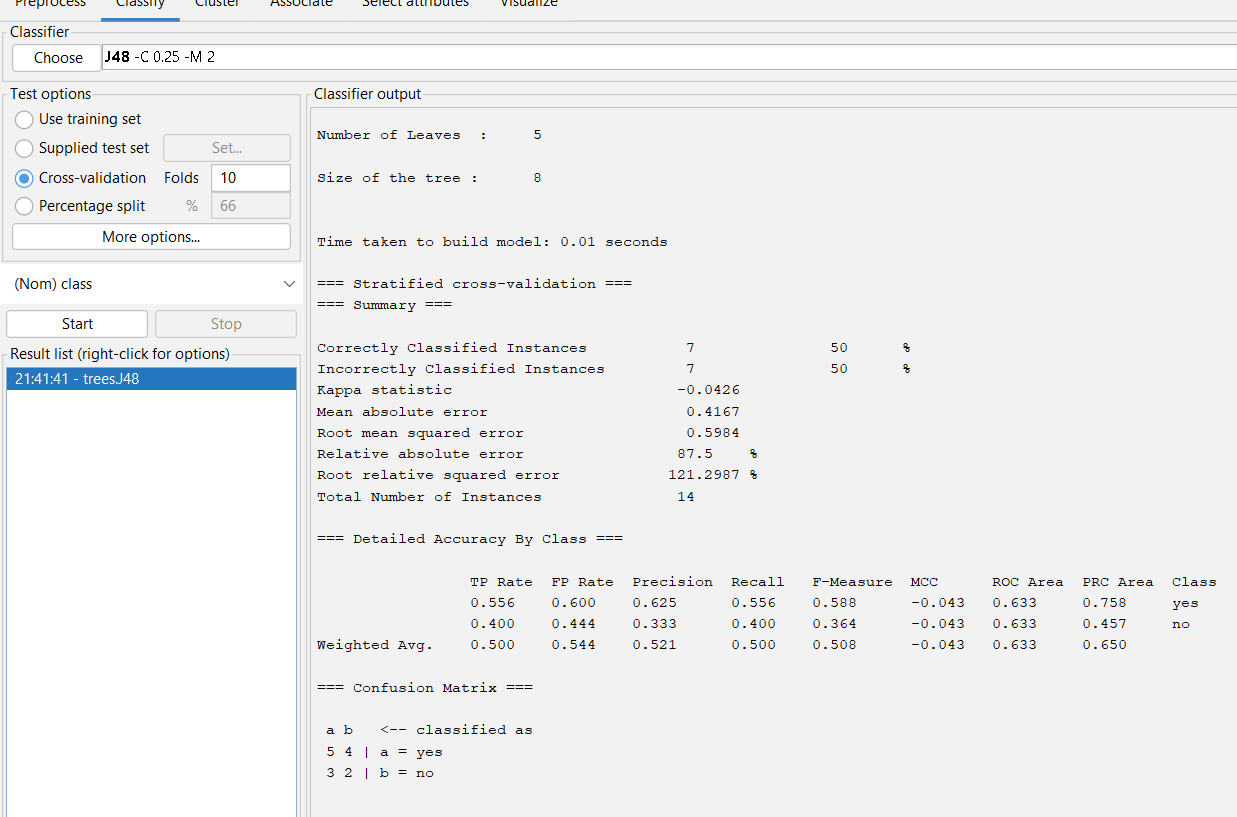
middle high yes fair yes

old medium no  excellent no

output:

tree:





**4.Analysis the dataset “diabetes. csv” how the diabetes trend is for different age people, using linear regression and multiple regression.**

Input:

data<-read.csv("C:/Users/Hari Naidu/Desktop/POM/download papers/diabetes.csv")

data

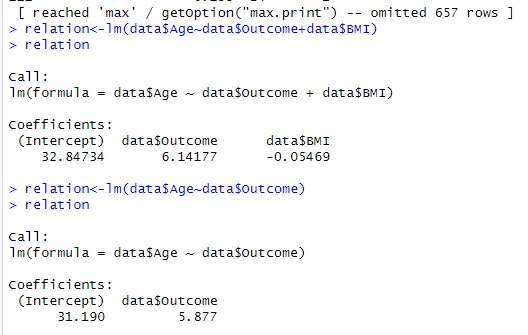
relation<-lm(data$Age~data$Outcome)

relation

relation<-lm(data$Age~data$Outcome+data$BMI)

relation

output:



**5.Implement using WEKA for the given Suppose a database has five transactions. Let min sup= 50%(2) and min con f = 80%.**

**Transactions Items**

T1 (M, O, N, K, E, Y)

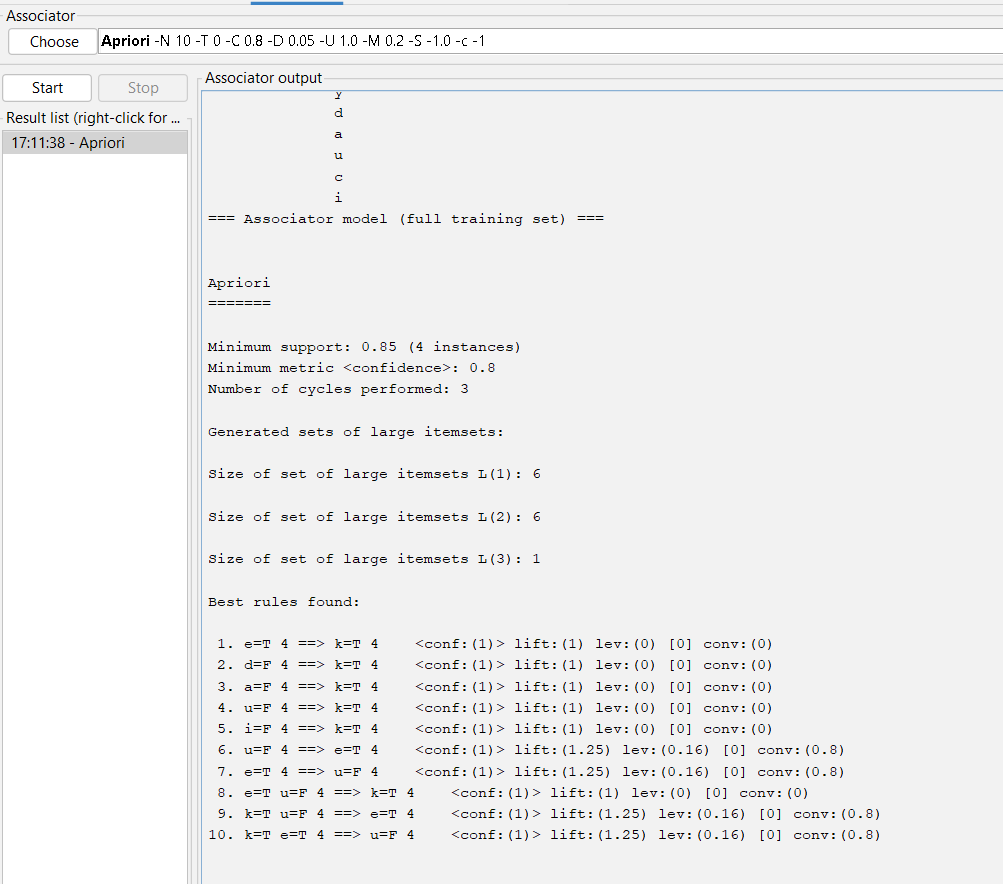
T2 (D, O, N, K, E, Y)

T3 (M, A, K, E)

T4 (M, U, C, K, Y)

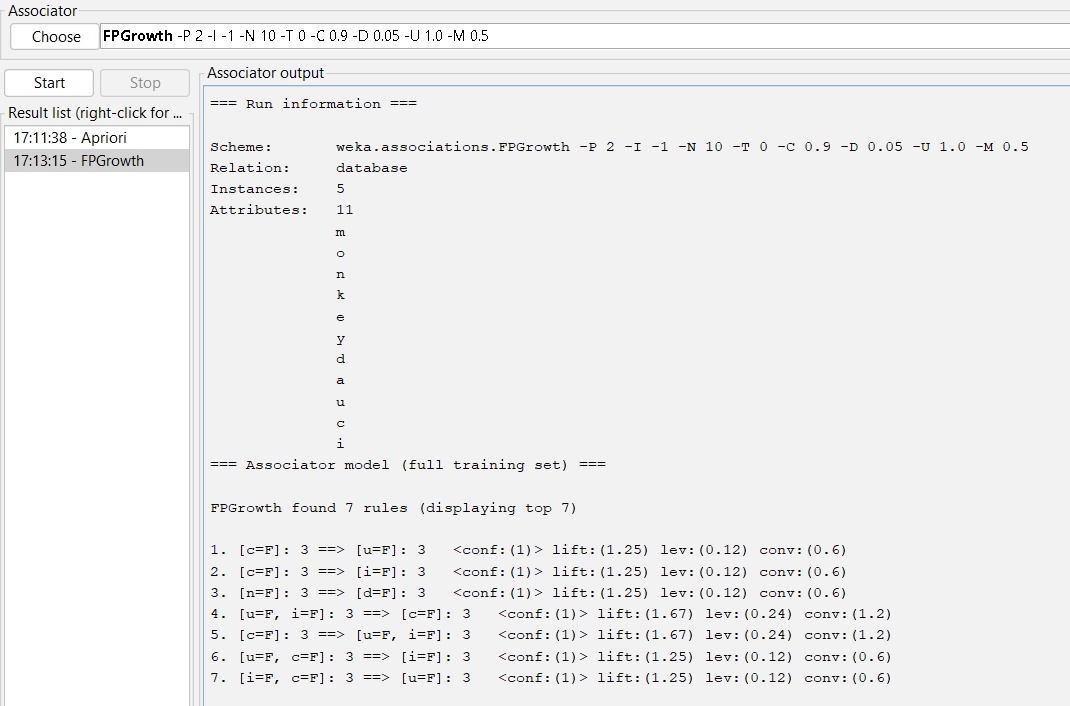
T5 (C,O, O, K, I ,E)

* Find all frequent item sets using Apriori algorithm
* Also draw FP-Growth Tree

Input:

Apriori algorithm:

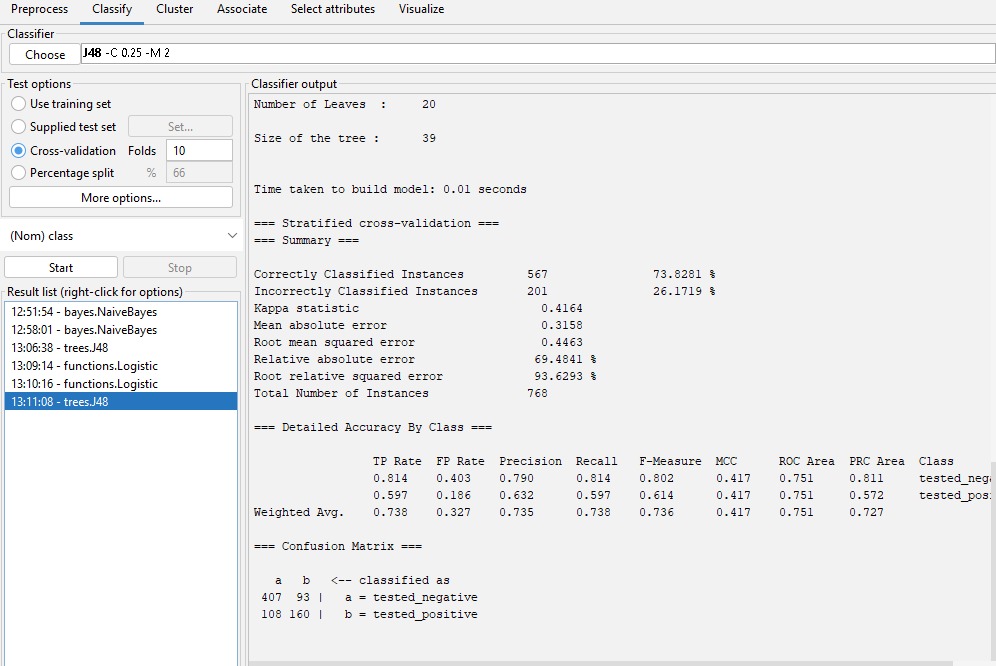
Fpgrowth algorithm:



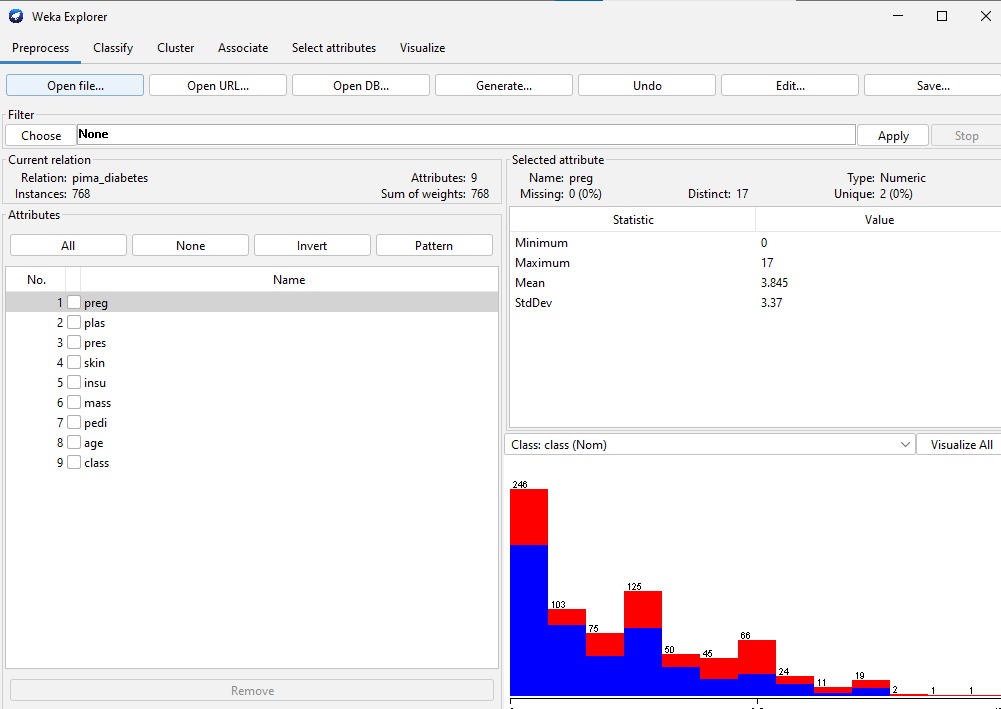
**6. Prediction of Categorical Data using Decision Tree Algorithm through WEKA using any datasets. a) Tree b) Preprocess c) Logistic**

Output:

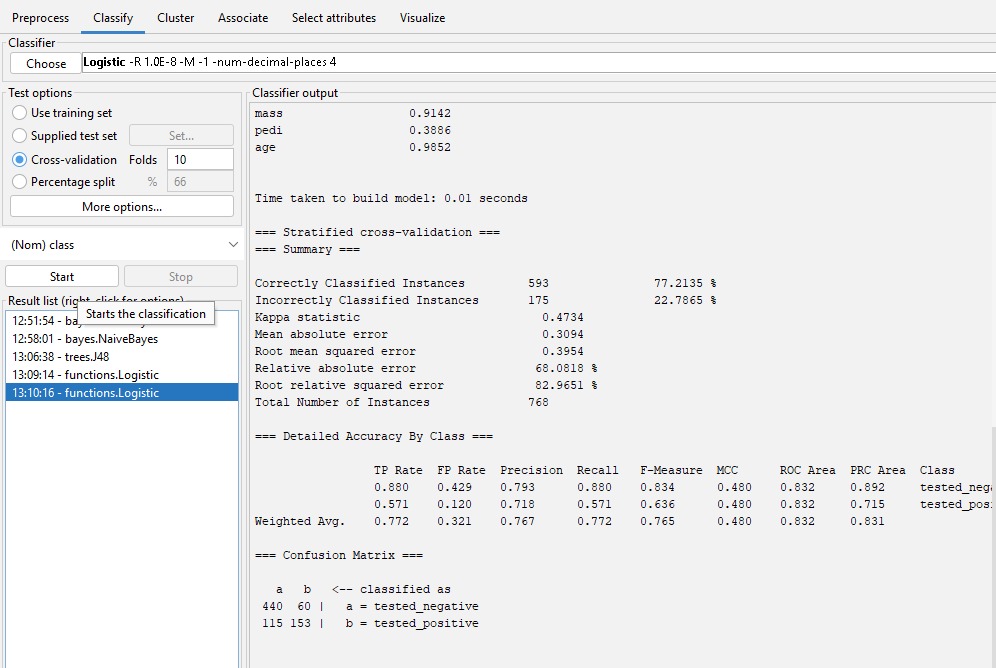
Tree:

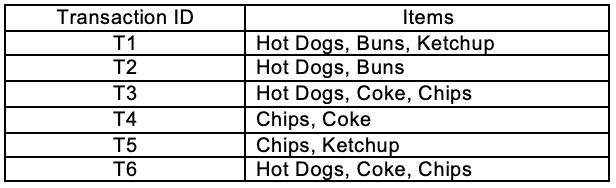


Preprocessor:



Logistic:



**7.Create the dataset using ARFF file** format:

a.Find the **frequent itemsets** and generate **association rules** on this. Assume that minimum support threshold (s = 33.33%) and minimum confident threshold (c = 60%).

b.List the various rule generated by apriori and FP tree algorthim ,mention wheather accepted or rejcted.

Input:

@relation hotdogs

@attribute hotdogs{t,f}

@attribute buns{t,f}

@attribute ketchup{t,f}

@attribute coke{t,f}

@attribute chips{t,f}

@data

t t t f f

t t f f f

t f f t t

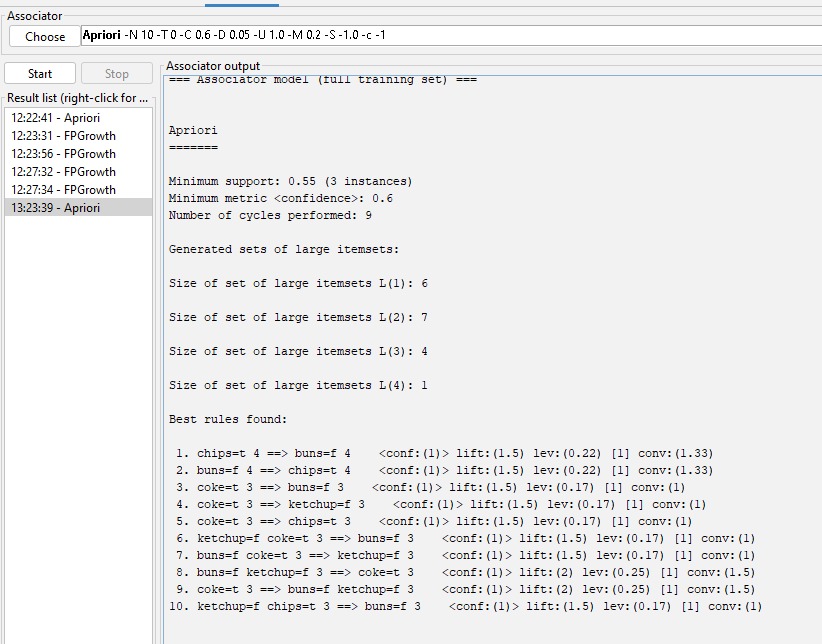
f f f t t

f f t f t

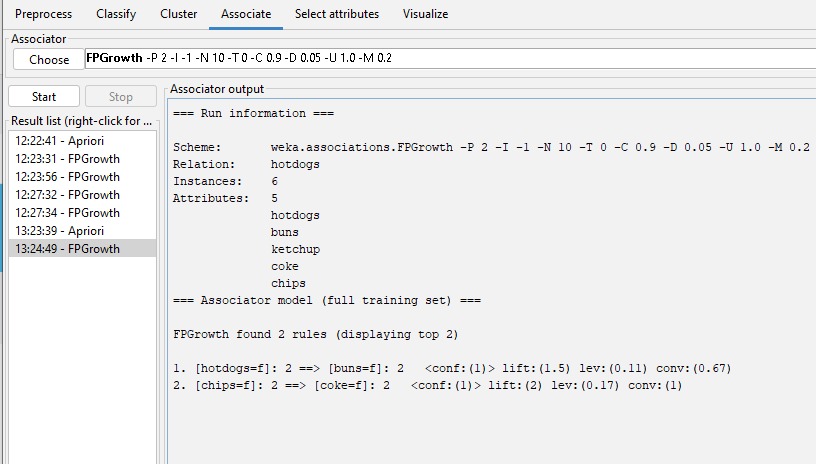
t f f t t

output:

apriori algorithm:



Fp growth:



**8.Prediction of Categorical Data using Rule base classification and decision tree classification through WEKA using any datasets. Compare the accuracy using two algorithm and plot the graph**

Input:

@relation decision\_tree

@attribute age{young,middle,old}

@attribute income{low,medium,high}

@attribute student{yes,no}

@attribute Creit\_rating{fair,excellent}

@attribute class{yes,no}

@data

young high no fair no

young high no excellent no

middle high no fair yes

old medium no fair yes

old low yes fair yes

old low yes excellent no

middle low yes excellent yes

young medium no fair no

young low yes fair yes

old medium yes fair yes

young medium yes excellent yes

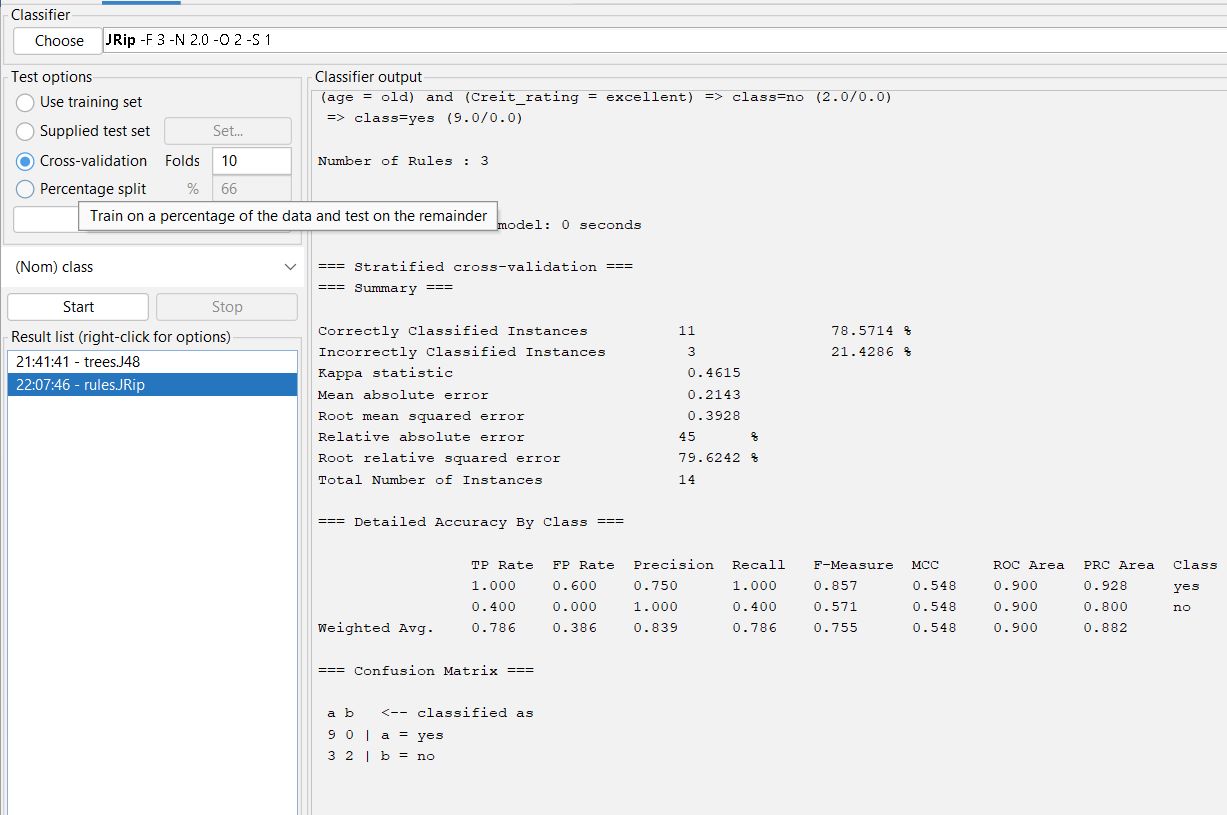
middle medium no excellent yes

middle high yes fair yes

old medium no  excellent no

Output:

Rule based classification:



Decision tree:

