Experiment-1

Program

@relation exp1

@attribute customerid numeric

@attribute gender{male,female}

@attribute age numeric

@attribute annual_income numeric

@attribute spending_score numeric

@data

1,male,19,15,39

2,male,21,15,81

3,female,20,16,6

4,female,23,16,77

5,female,31,17,40

Output

```
=== Run information ===

Scheme:       weka.clusterers.HierarchicalClusterer -N 2 -L SINGLE -P -A "weka.core.Filtere
Relation:     exp1
Instances:    5
Attributes:   5
              customerid
              gender
              age
              annual_income
              spending_score
Test mode:    evaluate on training data


=== Clustering model (full training set) ===

Cluster 0
(((39.0:0.73431,81.0:0.73431):0.90381,77.0:1.63811):0.28319,6.0:1.9213)



Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      4 ( 80%)
1      1 ( 20%)
```

Experiment-2

Program

@relation exp1

@attribute emp_id numeric

@attribute gender{male,female}

@attribute age numeric

@attribute salary numeric

@attribute credit numeric

@data

111,male,28,1500000,39

222,male,25,1500000,27

333,female,26,1600000,42

444,female,25,1600000,40

555,female,30,1700000,64

666,male,29,2000000,72

Output

```
Initial starting points (random):

Cluster 0: 444,female,25,1600000,40
Cluster 1: 111,male,28,1500000,39
Cluster 2: 222,male,25,1500000,27

Missing values globally replaced with mean/mode

Final cluster centroids:
                                Cluster#
Attribute       Full Data              0            1            2
                    (6.0)          (3.0)        (1.0)        (2.0)
==================================================================
emp_id              388.5            444          666        166.5
gender               male         female         male         male
age               27.1667             27           29         26.5
salary            1650000   1633333.3333      2000000      1500000
credit            47.3333        48.6667           72           33




Time taken to build model (full training data) : 0 seconds

=== Model and evaluation on training set ===

Clustered Instances

0       3 ( 50%)
1       1 ( 17%)
2       2 ( 33%)
```

Experiment-3 naivebayes in breast_cancer

Output

```
Time taken to build model: 0 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances          69               71.134  %
Incorrectly Classified Instances        28               28.866  %
Kappa statistic                          0.3274
Mean absolute error                      0.3431
Root mean squared error                  0.4825
Relative absolute error                 79.9872 %
Root relative squared error            100.9522 %
Total Number of Instances               97

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.828    0.515    0.757      0.828    0.791      0.331    0.676     0.789     no-recurre
                 0.485    0.172    0.593      0.485    0.533      0.331    0.676     0.560     recurrence
Weighted Avg.    0.711    0.398    0.701      0.711    0.703      0.331    0.676     0.711

=== Confusion Matrix ===

  a  b   <-- classified as
 53 11 |  a = no-recurrence-events
 17 16 |  b = recurrence-events
```

Experiment-4

Program

@relation exp1

@attribute person{gopu,babu,baby,gopal,krishna,jai,dev,malini,hema,anu}

@attribute gender{yes,no}

@data

gopu,yes

babu,yes

baby,yes

gopal,no

krishna,yes

jai,no

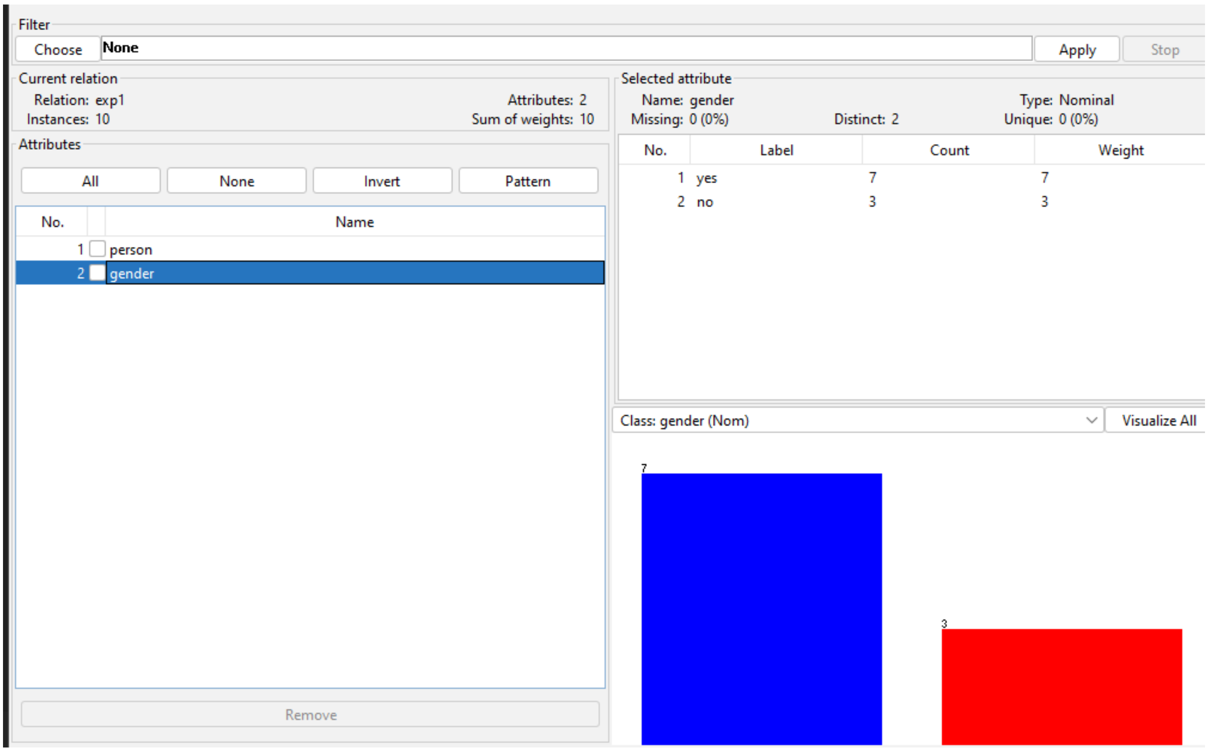
dev,no

malini,yes

hema,yes

anu,yes

output;



Experiment-6

Program

@relation exp1

@attribute bread{true,false}

@attribute cheese{true,false}

@attribute egg{true,false}

@attribute juice{true,false}

@attribute milk{true,false}

@attribute yogurt{true,false}

@data

true,true,true,true,false,false

true,true,false,true,false,false

true,false,false,false,true,true

true,false,false,true,true,false

false,true,false,true,true,false

output;

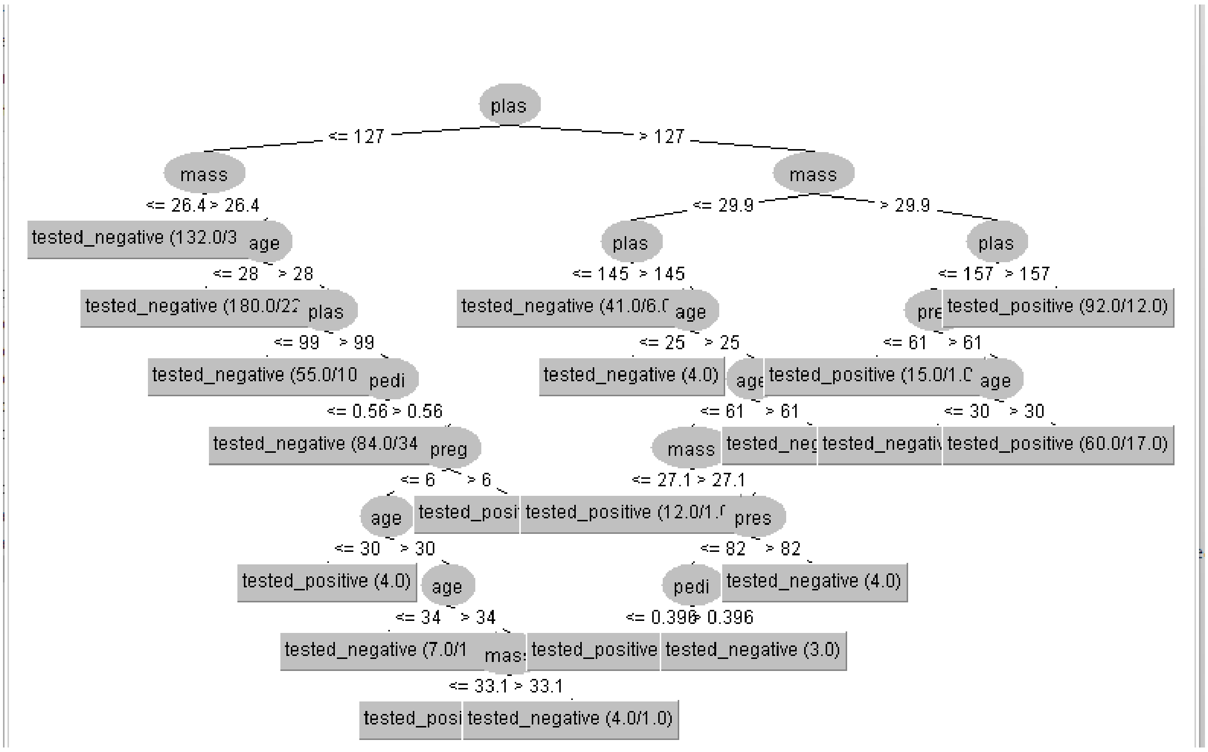
```
=== Run information ===

Scheme:       weka.associations.FPGrowth -P 2 -I -1 -N 10 -T 0 -C 0.75 -D 0.05 -U 2.0 -M 0.1
Relation:     exp1
Instances:    5
Attributes:   6
              bread
              cheese
              egg
              juice
              milk
              yogurt
=== Associator model (full training set) ===

FPGrowth found 14 rules (displaying top 10)

 1. [milk=false]: 2 ==> [yogurt=false]: 2   <conf:(1)> lift:(1.25) lev:(0.08) conv:(0.4)
 2. [bread=false]: 1 ==> [yogurt=false]: 1   <conf:(1)> lift:(1.25) lev:(0.04) conv:(0.2)
 3. [cheese=false]: 2 ==> [egg=false]: 2   <conf:(1)> lift:(1.25) lev:(0.08) conv:(0.4)
 4. [juice=false]: 1 ==> [egg=false]: 1   <conf:(1)> lift:(1.25) lev:(0.04) conv:(0.2)
 5. [bread=false]: 1 ==> [egg=false]: 1   <conf:(1)> lift:(1.25) lev:(0.04) conv:(0.2)
 6. [juice=false]: 1 ==> [cheese=false]: 1   <conf:(1)> lift:(2.5) lev:(0.12) conv:(0.6)
 7. [egg=false, milk=false]: 1 ==> [yogurt=false]: 1   <conf:(1)> lift:(1.25) lev:(0.04) conv:(0.2)
 8. [yogurt=false, cheese=false]: 1 ==> [egg=false]: 1   <conf:(1)> lift:(1.25) lev:(0.04) conv:(0.2)
 9. [bread=false]: 1 ==> [yogurt=false, egg=false]: 1   <conf:(1)> lift:(1.67) lev:(0.08) conv:(0.4)
10. [yogurt=false, bread=false]: 1 ==> [egg=false]: 1   <conf:(1)> lift:(1.25) lev:(0.04) conv:(0.2)
```

Experiment-7



Experiment-8

```r
library(dplyr)

library(tibble)

library(ggplot2)

data <- c(55, 60, 71, 63, 55, 65, 50, 55,58,59,61,63,65,67,71,72,75)

equi_depth <- quantile(data, probs = c(0, 1/3, 2/3, 1))

equi_depth_partitioned <- cut(data, breaks = equi_depth, labels = c("Low", "Medium", "High"),
include.lowest = TRUE)

min_value <- min(data)

max_value <- max(data)

width <- (max_value - min_value)/3

equal_width <- seq(min_value, max_value, by = width)

equal_width_partitioned <- cut(data, breaks = equal_width, labels = c("Low", "Medium", "High"),
include.lowest = TRUE)

kmeans_model <- kmeans(data, centers = 3)

cluster_assignments <- as.factor(kmeans_model$cluster)

levels(cluster_assignments) <- c("Low", "Medium", "High")


data_tibble <- tibble(data = data, equi_depth_partitioned = equi_depth_partitioned,
equal_width_partitioned = equal_width_partitioned, cluster_assignments = cluster_assignments)


ggplot(data_tibble, aes(x = data)) +

  geom_histogram(binwidth = 5) +

  facet_wrap(~equi_depth_partitioned, ncol = 1, scales = "free_x") +

  ggtitle("Histogram using Equi-Depth Partitioning")

ggplot(data_tibble, aes(x = data)) +

  geom_histogram(binwidth = 5) +

  facet_wrap(~equal_width_partitioned, ncol = 1, scales = "free_x") +

  ggtitle("Histogram using Equal-Width Partitioning")

ggplot(data_tibble, aes(x = data)) +

  geom_histogram(binwidth = 5) +

  facet_wrap(~cluster_assignments, ncol = 1, scales = "free_x") +
```

experiment-9

program

@relation exp1

@attribute SONY{true,false}

@attribute BPL{true,false}

@attribute LG{true,false}

@attribute SAMSUNG{true,false}

@attribute ONIDA{true,false}

@data

true,true,true,false,false

false,true,false,true,false

false,true,false,false,true

true,true,false,true,false

true,false,false,false,true

false,true,false,false,true

true,false,false,false,true

true,true,true,false,true

true,true,false,false,true

output;

```
=== Run information ===

Scheme:       weka.associations.FPGrowth -P 2 -I -1 -N 10 -T 0 -C 0.5 -D 0.05 -U 2.0 -M 0.1
Relation:     exp1
Instances:    9
Attributes:   5
              SONY
              BPL
              LG
              SAMSUNG
              ONIDA
=== Associator model (full training set) ===

FPGrowth found 10 rules (displaying top 10)

 1. [BPL=false]: 2 ==> [SAMSUNG=false]: 2   <conf:(1)> lift:(1.29) lev:(0.05) conv:(0.44)
 2. [BPL=false]: 2 ==> [LG=false]: 2   <conf:(1)> lift:(1.29) lev:(0.05) conv:(0.44)
 3. [SAMSUNG=false, SONY=false]: 2 ==> [LG=false]: 2   <conf:(1)> lift:(1.29) lev:(0.05) conv:(0.44)
 4. [BPL=false]: 2 ==> [SAMSUNG=false, LG=false]: 2   <conf:(1)> lift:(1.8) lev:(0.1) conv:(0.89)
 5. [SAMSUNG=false, BPL=false]: 2 ==> [LG=false]: 2   <conf:(1)> lift:(1.29) lev:(0.05) conv:(0.44)
 6. [LG=false, BPL=false]: 2 ==> [SAMSUNG=false]: 2   <conf:(1)> lift:(1.29) lev:(0.05) conv:(0.44)
 7. [SONY=false]: 3 ==> [SAMSUNG=false]: 2   <conf:(0.67)> lift:(0.86) lev:(-0.04) conv:(0.33)
 8. [ONIDA=false]: 3 ==> [LG=false]: 2   <conf:(0.67)> lift:(0.86) lev:(-0.04) conv:(0.33)
 9. [SONY=false]: 3 ==> [SAMSUNG=false, LG=false]: 2   <conf:(0.67)> lift:(1.2) lev:(0.04) conv:(0.67)
10. [LG=false, SONY=false]: 3 ==> [SAMSUNG=false]: 2   <conf:(0.67)> lift:(0.86) lev:(-0.04) conv:(0.33)
```

```
Apriori
=======

Minimum support: 0.6 (5 instances)
Minimum metric <confidence>: 0.5
Number of cycles performed: 8

Generated sets of large itemsets:

Size of set of large itemsets L(1): 5

Size of set of large itemsets L(2): 6

Size of set of large itemsets L(3): 1

Best rules found:

 1. ONIDA=true 6 ==> SAMSUNG=false 6    <conf:(1)> lift:(1.29) lev:(0.15) [1] conv:(1.33)
 2. LG=false ONIDA=true 5 ==> SAMSUNG=false 5    <conf:(1)> lift:(1.29) lev:(0.12) [1] conv:(1.11)
 3. LG=false SAMSUNG=false 5 ==> ONIDA=true 5    <conf:(1)> lift:(1.5) lev:(0.19) [1] conv:(1.67)
 4. SAMSUNG=false 7 ==> ONIDA=true 6    <conf:(0.86)> lift:(1.29) lev:(0.15) [1] conv:(1.17)
 5. SONY=true 6 ==> SAMSUNG=false 5    <conf:(0.83)> lift:(1.07) lev:(0.04) [0] conv:(0.67)
 6. ONIDA=true 6 ==> LG=false 5    <conf:(0.83)> lift:(1.07) lev:(0.04) [0] conv:(0.67)
 7. SAMSUNG=false ONIDA=true 6 ==> LG=false 5    <conf:(0.83)> lift:(1.07) lev:(0.04) [0] conv:(0.67)
 8. ONIDA=true 6 ==> LG=false SAMSUNG=false 5    <conf:(0.83)> lift:(1.5) lev:(0.19) [1] conv:(1.33)
 9. SAMSUNG=false 7 ==> SONY=true 5    <conf:(0.71)> lift:(1.07) lev:(0.04) [0] conv:(0.78)
10. LG=false 7 ==> BPL=true 5    <conf:(0.71)> lift:(0.92) lev:(-0.05) [0] conv:(0.52)
```

Experiment-10   Program

speed<-c(78,81,82,74,83,82,77,80,70)

time<-c(39,37,36,42,35,36,40,38,46)

a<-var(speed)

b<-var(time)

ss<-sqrt(var(speed))

st<-sqrt(var(time))

a

b

ss

st

output

```
> speed<-c(78,81,82,74,83,82,77,80,70)
> time<-c(39,37,36,42,35,36,40,38,46)
> a<-var(speed)
> b<-var(time)
> ss<-sqrt(var(speed))
> st<-sqrt(var(time))
> a
[1] 18.52778
> b
[1] 12.19444
> ss
[1] 4.304391
> st
[1] 3.492054
>
```

Experiment-11

a<-c(100,70,60,90,90)

min=0

max=1

v<-((a-min(a)/max(a)-min(a))*max+min)/100

summary(v,method=c("range"))

v

#zscore

w=mean(a)

s=sd(a)

z=(a-w)/s

summary(z)

z

output

```
R 4.3.1 · ~/
> a<-c(100,70,60,90,90)
> min=0
> max=1
> v<-((a-min(a)/max(a)-min(a))*max+min)/100
> summary(v,method=c("range"))
   Min. 1st Qu.  Median    Mean 
 -0.006   0.094   0.294   0.214 
3rd Qu.    Max. 
  0.294   0.394 
> v
[1]  0.394  0.094 -0.006  0.294  0.294
> #zscore
> w=mean(a)
> s=sd(a)
> z=(a-w)/s
> summary(z)
   Min. 1st Qu.  Median    Mean 
-1.3389 -0.7303  0.4869  0.0000 
3rd Qu.    Max. 
 0.4869  1.0954 
> z
[1]  1.0954451 -0.7302967 -1.3388774
[4]  0.4868645  0.4868645
>
```