Association Rule Mining

Bibek Sapkota

task 1: Install package

```
#install.packages("arules")
library("arules")
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'arules'
```

```
## The following objects are masked from 'package:base':
##
##     abbreviate, write
```

task 2:Importing and loading the data.

```
dataAR <- read.csv("weather.csv")
dataAR
```

```
##     outlook temperature humidity windy play
## 1     sunny         hot     high    no   no
## 2     sunny         hot     high   yes   no
## 3  overcast         hot     high    no  yes
## 4     rainy        mild     high    no  yes
## 5     rainy        cool   normal    no  yes
## 6     rainy        cool   normal   yes   no
## 7  overcast        cool   normal   yes  yes
## 8     sunny        mild     high    no   no
## 9     sunny        cool   normal    no  yes
## 10    rainy        mild   normal    no  yes
## 11    sunny        mild   normal   yes  yes
## 12 overcast        mild     high   yes  yes
## 13 overcast         hot   normal    no  yes
## 14    rainy        mild     high   yes   no
```

task 3: Applying apriori

```
rules <- apriori(dataAR)
```

```
## Warning: Column(s) 1, 2, 3, 4, 5 not logical or factor. Applying default
## discretization (see '? discretizeDF').
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.8    0.1    1 none FALSE            TRUE       5     0.1      1
##  maxlen target  ext
##      10  rules TRUE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 1
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[12 item(s), 14 transaction(s)] done [0.00s].
## sorting and recoding items ... [12 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 done [0.00s].
## writing ... [53 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

task 4: Inspect the dataset

```
inspect(rules)
```

```
##      lhs                    rhs                   support confidence  coverage     lift count
## [1]  {outlook=overcast}  => {play=yes}         0.2857143  1.0000000 0.2857143 1.555556     4
## [2]  {temperature=cool}  => {humidity=normal}  0.2857143  1.0000000 0.2857143 2.000000     4
## [3]  {play=no}           => {humidity=high}    0.2857143  0.8000000 0.3571429 1.600000     4
## [4]  {humidity=normal}   => {play=yes}         0.4285714  0.8571429 0.5000000 1.333333     6
## [5]  {outlook=overcast,
##       temperature=hot}   => {windy=no}         0.1428571  1.0000000 0.1428571 1.750000     2
## [6]  {outlook=overcast,
##       windy=no}          => {temperature=hot}  0.1428571  1.0000000 0.1428571 3.500000     2
## [7]  {outlook=overcast,
##       temperature=hot}   => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [8]  {temperature=hot,
##       play=yes}          => {outlook=overcast} 0.1428571  1.0000000 0.1428571 3.500000     2
## [9]  {outlook=sunny,
##       temperature=hot}   => {play=no}          0.1428571  1.0000000 0.1428571 2.800000     2
## [10] {temperature=hot,
##       play=no}           => {outlook=sunny}    0.1428571  1.0000000 0.1428571 2.800000     2
## [11] {outlook=sunny,
##       temperature=hot}   => {humidity=high}    0.1428571  1.0000000 0.1428571 2.000000     2
## [12] {temperature=hot,
##       play=no}           => {humidity=high}    0.1428571  1.0000000 0.1428571 2.000000     2
## [13] {temperature=hot,
##       play=yes}          => {windy=no}         0.1428571  1.0000000 0.1428571 1.750000     2
## [14] {outlook=overcast,
##       windy=yes}         => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [15] {outlook=overcast,
##       humidity=high}     => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
```

```
## [16] {outlook=overcast,
##       humidity=normal}   => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [17] {outlook=overcast,
##       windy=no}          => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [18] {outlook=rainy,
##       temperature=cool}  => {humidity=normal}  0.1428571  1.0000000 0.1428571 2.000000     2
## [19] {temperature=cool,
##       windy=yes}         => {humidity=normal}  0.1428571  1.0000000 0.1428571 2.000000     2
## [20] {temperature=cool,
##       windy=no}          => {humidity=normal}  0.1428571  1.0000000 0.1428571 2.000000     2
## [21] {temperature=cool,
##       play=yes}          => {humidity=normal}  0.2142857  1.0000000 0.2142857 2.000000     3
## [22] {temperature=cool,
##       windy=no}          => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [23] {outlook=sunny,
##       play=no}           => {humidity=high}    0.2142857  1.0000000 0.2142857 2.000000     3
## [24] {outlook=sunny,
##       humidity=high}     => {play=no}          0.2142857  1.0000000 0.2142857 2.800000     3
## [25] {windy=no,
##       play=no}           => {outlook=sunny}    0.1428571  1.0000000 0.1428571 2.800000     2
## [26] {outlook=sunny,
##       humidity=normal}   => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [27] {outlook=sunny,
##       play=yes}          => {humidity=normal}  0.1428571  1.0000000 0.1428571 2.000000     2
## [28] {outlook=rainy,
##       play=no}           => {windy=yes}        0.1428571  1.0000000 0.1428571 2.333333     2
## [29] {outlook=rainy,
##       windy=yes}         => {play=no}          0.1428571  1.0000000 0.1428571 2.800000     2
## [30] {temperature=mild,
##       play=no}           => {humidity=high}    0.1428571  1.0000000 0.1428571 2.000000     2
## [31] {windy=no,
##       play=no}           => {humidity=high}    0.1428571  1.0000000 0.1428571 2.000000     2
## [32] {outlook=rainy,
##       humidity=high}     => {temperature=mild} 0.1428571  1.0000000 0.1428571 2.333333     2
## [33] {outlook=rainy,
##       windy=no}          => {play=yes}         0.2142857  1.0000000 0.2142857 1.555556     3
## [34] {outlook=rainy,
##       play=yes}          => {windy=no}         0.2142857  1.0000000 0.2142857 1.750000     3
## [35] {temperature=mild,
##       humidity=normal}   => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [36] {humidity=normal,
##       windy=no}          => {play=yes}         0.2857143  1.0000000 0.2857143 1.555556     4
## [37] {outlook=overcast,
##       temperature=hot,
##       windy=no}          => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [38] {outlook=overcast,
##       temperature=hot,
##       play=yes}          => {windy=no}         0.1428571  1.0000000 0.1428571 1.750000     2
## [39] {temperature=hot,
##       windy=no,
##       play=yes}          => {outlook=overcast} 0.1428571  1.0000000 0.1428571 3.500000     2
## [40] {outlook=overcast,
##       windy=no,
##       play=yes}          => {temperature=hot}  0.1428571  1.0000000 0.1428571 3.500000     2
```

```
## [41] {outlook=sunny,
##       temperature=hot,
##       play=no}           => {humidity=high}    0.1428571  1.0000000 0.1428571 2.000000     2
## [42] {outlook=sunny,
##       temperature=hot,
##       humidity=high}     => {play=no}          0.1428571  1.0000000 0.1428571 2.800000     2
## [43] {temperature=hot,
##       humidity=high,
##       play=no}           => {outlook=sunny}    0.1428571  1.0000000 0.1428571 2.800000     2
## [44] {temperature=cool,
##       humidity=normal,
##       windy=no}          => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [45] {temperature=cool,
##       windy=no,
##       play=yes}          => {humidity=normal}  0.1428571  1.0000000 0.1428571 2.000000     2
## [46] {outlook=sunny,
##       windy=no,
##       play=no}           => {humidity=high}    0.1428571  1.0000000 0.1428571 2.000000     2
## [47] {outlook=sunny,
##       humidity=high,
##       windy=no}          => {play=no}          0.1428571  1.0000000 0.1428571 2.800000     2
## [48] {humidity=high,
##       windy=no,
##       play=no}           => {outlook=sunny}    0.1428571  1.0000000 0.1428571 2.800000     2
## [49] {outlook=rainy,
##       temperature=mild,
##       windy=no}          => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [50] {outlook=rainy,
##       temperature=mild,
##       play=yes}          => {windy=no}         0.1428571  1.0000000 0.1428571 1.750000     2
## [51] {temperature=mild,
##       windy=no,
##       play=yes}          => {outlook=rainy}    0.1428571  1.0000000 0.1428571 2.800000     2
## [52] {outlook=rainy,
##       humidity=normal,
##       windy=no}          => {play=yes}         0.1428571  1.0000000 0.1428571 1.555556     2
## [53] {outlook=rainy,
##       humidity=normal,
##       play=yes}          => {windy=no}         0.1428571  1.0000000 0.1428571 1.750000     2
```

task 5: Using Apriori algorithm mines association rules in dataAR with minimum 30% support and 30% confidence.

```
rules <- apriori(dataAR, parameter =list(supp=0.3, conf=0.3) )
```

```
## Warning: Column(s) 1, 2, 3, 4, 5 not logical or factor. Applying default
## discretization (see '? discretizeDF').
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.3    0.1    1 none FALSE            TRUE       5     0.3      1
```

```
##  maxlen target  ext
##      10  rules TRUE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 4
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[12 item(s), 14 transaction(s)] done [0.00s].
## sorting and recoding items ... [9 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 done [0.00s].
## writing ... [13 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

task 6:Display detailed information of the mined association rules.

```
inspect(rules)
```

```
##      lhs                   rhs                 support   confidence coverage
## [1]  {}                 => {outlook=sunny}     0.3571429 0.3571429  1.0000000
## [2]  {}                 => {play=no}           0.3571429 0.3571429  1.0000000
## [3]  {}                 => {outlook=rainy}     0.3571429 0.3571429  1.0000000
## [4]  {}                 => {windy=yes}         0.4285714 0.4285714  1.0000000
## [5]  {}                 => {temperature=mild}  0.4285714 0.4285714  1.0000000
## [6]  {}                 => {humidity=high}     0.5000000 0.5000000  1.0000000
## [7]  {}                 => {humidity=normal}   0.5000000 0.5000000  1.0000000
## [8]  {}                 => {windy=no}          0.5714286 0.5714286  1.0000000
## [9]  {}                 => {play=yes}          0.6428571 0.6428571  1.0000000
## [10] {humidity=normal} => {play=yes}          0.4285714 0.8571429  0.5000000
## [11] {play=yes}        => {humidity=normal}   0.4285714 0.6666667  0.6428571
## [12] {windy=no}        => {play=yes}          0.4285714 0.7500000  0.5714286
## [13] {play=yes}        => {windy=no}          0.4285714 0.6666667  0.6428571
##      lift     count
## [1]  1.000000 5
## [2]  1.000000 5
## [3]  1.000000 5
## [4]  1.000000 6
## [5]  1.000000 6
## [6]  1.000000 7
## [7]  1.000000 7
## [8]  1.000000 8
## [9]  1.000000 9
## [10] 1.333333 6
## [11] 1.333333 6
## [12] 1.166667 6
## [13] 1.166667 6
```

task 7:Finds association rules in dataAR with 30% support, 30% confidence, and minimum length of 2.

```
rules <- apriori(dataAR, parameter =list (supp=0.3, conf=0.3, minlen=2 ) )
```

```
## Warning: Column(s) 1, 2, 3, 4, 5 not logical or factor. Applying default
## discretization (see '? discretizeDF').
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.3    0.1    1 none FALSE            TRUE       5     0.3      2
##  maxlen target  ext
##      10  rules TRUE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 4
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[12 item(s), 14 transaction(s)] done [0.00s].
## sorting and recoding items ... [9 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 done [0.00s].
## writing ... [4 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

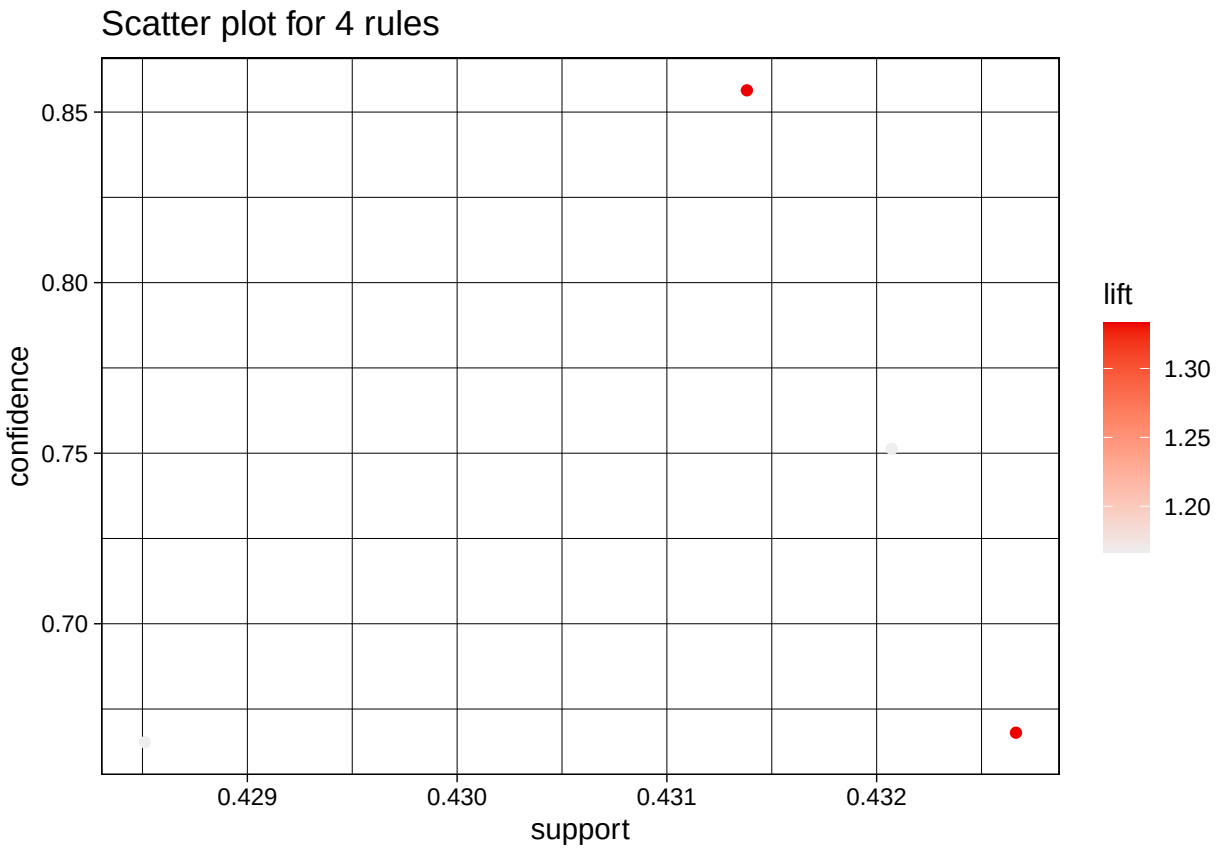
task 8: installing and loading the package

```
#install.packages("arulesViz")
library("arulesViz")
```

task 9:Ploting rules

```
plot(rules)
```

```
## To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
```

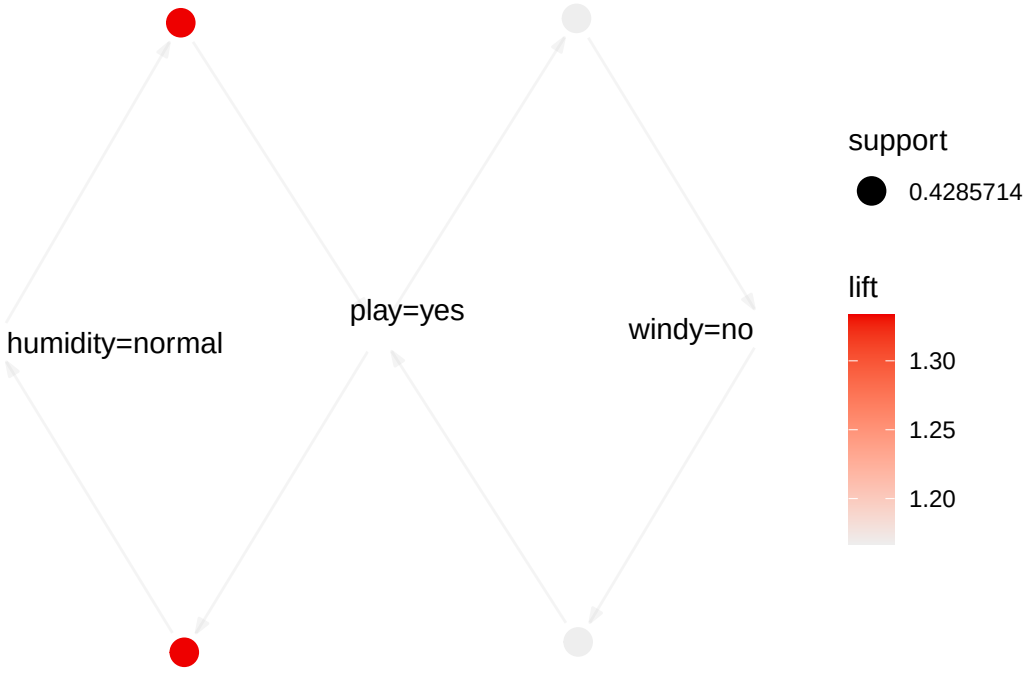


task 10:Using alternative graph

```
plot(rules, method="graph", control = list(type="items"))
```

```
## Warning: Unknown control parameters: type
```

```
## Available control parameters (with default values):
## layout    =  stress
## circular  =  FALSE
## ggraphdots    =  NULL
## edges     =  <environment>
## nodes     =  <environment>
## nodetext  =  <environment>
## colors    =  c("#EE0000FF", "#EEEEEEFF")
## engine    =  ggplot2
## max   =  100
## verbose   =  FALSE
```


humidity=normal
play=yes
windy=no
support
0.4285714
lift
1.30
1.25
1.20