

Data Warehousing and Data Mining Assignment - 5

**Submitted by: Shiv Kumar
2016UIT2563**

Decision Tree for Location Data (J48 classifier)

Dataset:

@relation location

@attribute Customer {m,f}

@attribute 'Bank Name' {Kotak,ICICI,LIC}

@attribute location {Mumbai,Delhi,Pune}

@attribute deposit {yes,no}

@attribute job {employed,unemployed}

@data

m,Kotak,Mumbai,yes,employed

m,ICICI,Delhi,yes,employed

f,ICICI,Pune,no,unemployed

m,LIC,Pune,yes,employed

f,Kotak,Mumbai,yes,employed

f,Kotak,Mumbai,no,employed

f,LIC,Delhi,yes,unemployed

m,ICICI,Pune,yes,unemployed

f,LIC,Delhi,yes,unemployed

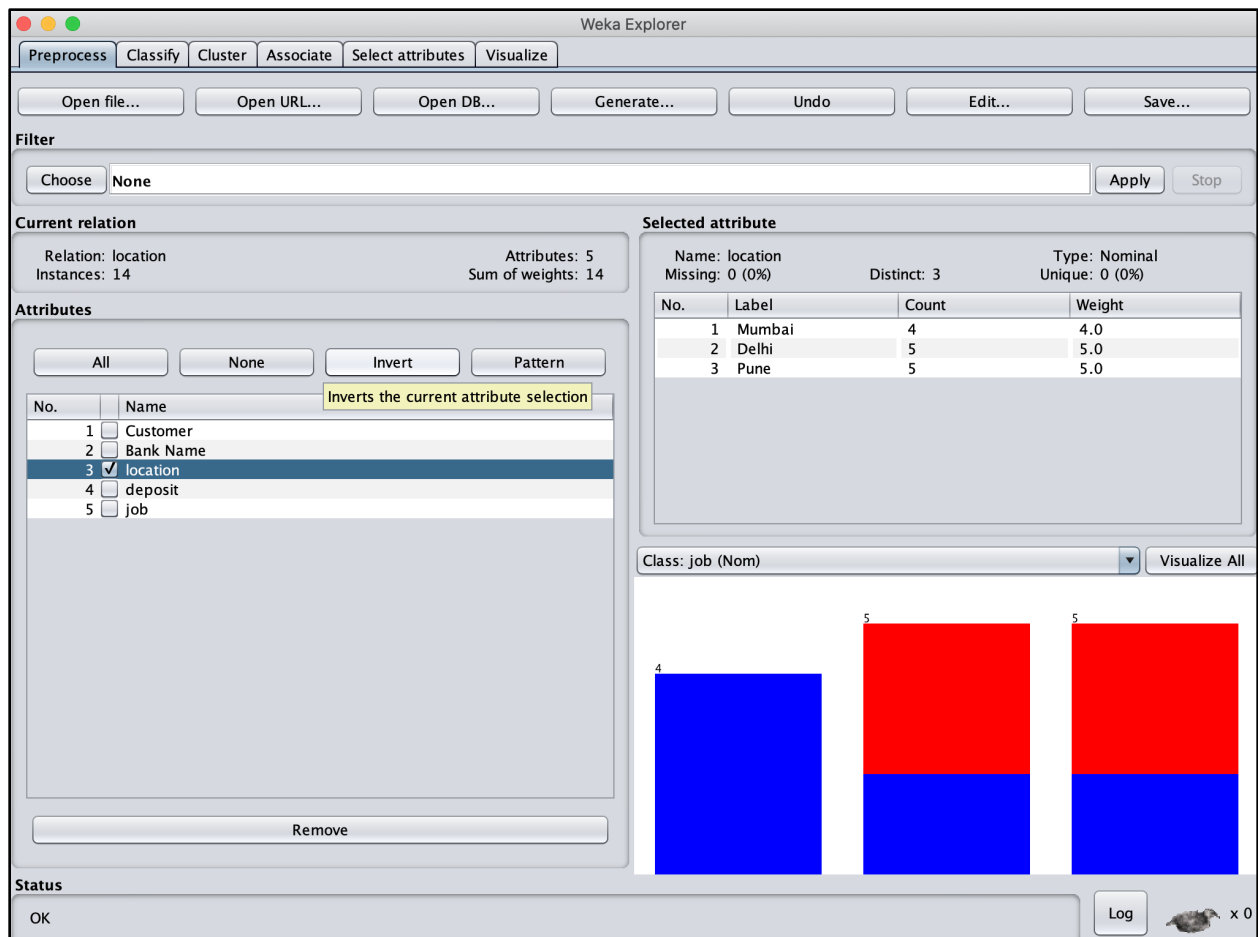
m,Kotak,Delhi,no,employed

m,LIC,Pune,yes,employed

f,Kotak,Delhi,yes,unemployed

m,ICICI,Mumbai,no,employed

f,ICICI,Pune,yes,unemployed



Viewer

Relation: location

No.	1: Customer	2: Bank Name	3: location	4: deposit	5: job
	Nominal	Nominal	Nominal	Nominal	Nominal
1	m	Kotak	Mumbai	yes	em...
2	m	ICICI	Delhi	yes	em...
3	f	ICICI	Pune	no	une...
4	m	LIC	Pune	yes	em...
5	f	Kotak	Mumbai	yes	em...
6	f	Kotak	Mumbai	no	em...
7	f	LIC	Delhi	yes	une...
8	m	ICICI	Pune	yes	une...
9	f	LIC	Delhi	yes	une...
...	m	Kotak	Delhi	no	em...
...	m	LIC	Pune	yes	em...
...	f	Kotak	Delhi	yes	une...
...	m	ICICI	Mumbai	no	em...
...	f	ICICI	Pune	yes	une...

Add instance Undo OK Cancel

After applying J48 classifier:

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier

Choose J48 -C 0.25 -M 2

Test options

☐ Use training set
☐ Supplied test set Set...
☒ Cross-validation Folds 10
☐ Percentage split % 66
 More options...

(Nom) location

Start Stop

Result list (right-click for options)

12:03:14 - trees.J48
12:22:12 - trees.J48

Classifier output

```

Bank Name = Kotak: Mumbai (4.0/1.0)
Bank Name = ICICI: Mumbai (2.0/1.0)
Bank Name = LIC: Pune (2.0)
job = unemployed
Bank Name = Kotak: Delhi (1.0)
Bank Name = ICICI: Pune (3.0)
Bank Name = LIC: Delhi (2.0)

Number of Leaves : 6
Size of the tree : 9

Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      5           35.7143 %
Incorrectly Classified Instances    9           64.2857 %
Kappa statistic                    0.0455
Mean absolute error                 0.407
Root mean squared error             0.5532
Relative absolute error             87.8933 %
Root relative squared error        112.2448 %
Total Number of Instances          14

=== Detailed Accuracy By Class ===

          TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
          0.500    0.400    0.333    0.500    0.400     0.091    0.525    0.310    Mumbai
          0.000    0.111    0.000    0.000    0.000    -0.207    0.522    0.443    Delhi
          0.600    0.444    0.429    0.600    0.500     0.149    0.667    0.542    Pune
Weighted Avg.   0.357    0.313    0.248    0.357    0.293     0.005    0.575    0.440

=== Confusion Matrix ===
 a b c  <-- classified as
 2 1 1 | a = Mumbai
 2 0 3 | b = Delhi
 2 0 3 | c = Pune
  
```

Status

OK Log x 0

Decision Tree

