

Data warehouse & Data mining lab

Assignment 3b

Submitted by :
Shiv Kumar
2016UIT2563
Semester 8
IT2

Experiment 10 : Write a procedure for Banking data using Farthest First Algorithm.

1. Data Loading

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... Open URL... Open DB... Generate... Undo Edit... Save...

Filter: Choose None Apply Stop

Current relation: Relation: bank Instances: 15 Attributes: 5 Sum of weights: 15

Attributes: All None Invert Pattern

No.	Name
1	<input checked="" type="checkbox"/> cust
2	<input type="checkbox"/> accno
3	<input type="checkbox"/> bankname
4	<input type="checkbox"/> location
5	<input type="checkbox"/> deposit

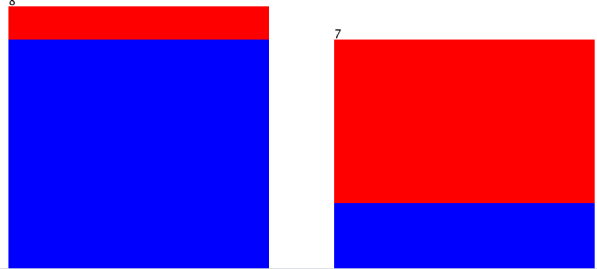
Remove

Selected attribute

Name: cust Missing: 0 (0%) Distinct: 2 Type: Nominal Unique: 0 (0%)

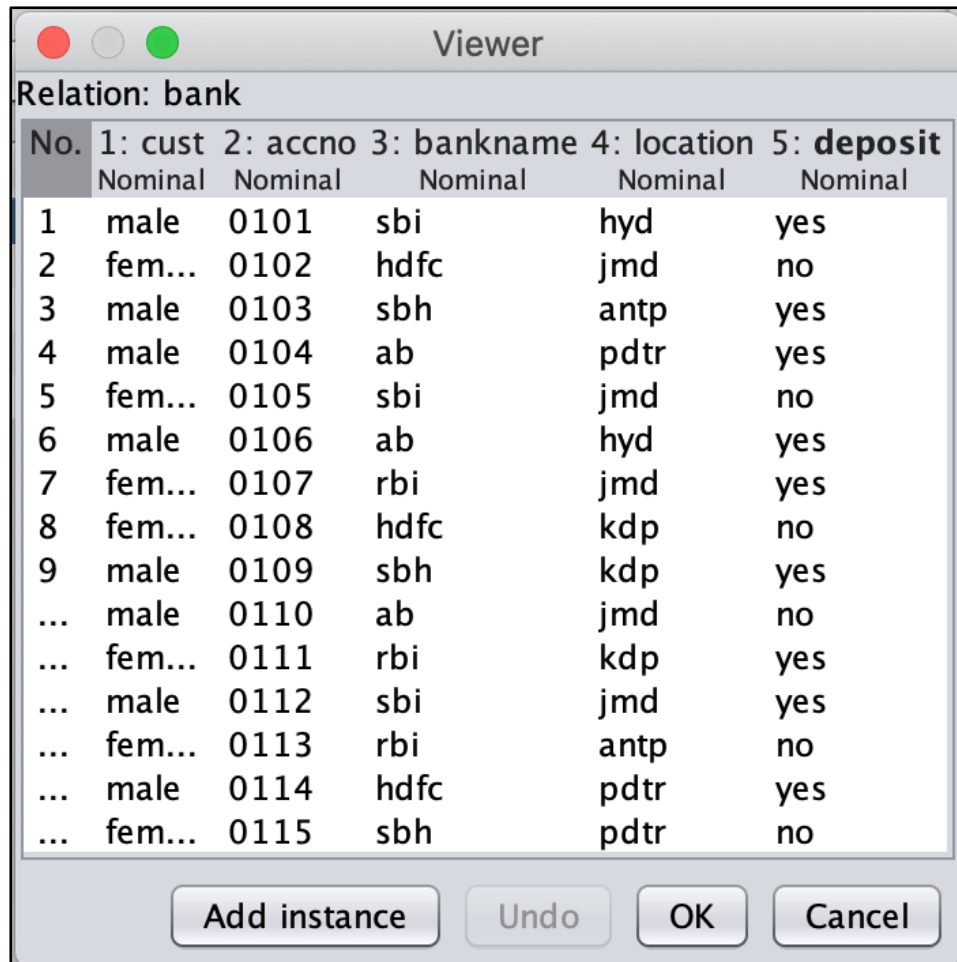
No.	Label	Count	Weight
1	male	8	8.0
2	female	7	7.0

Class: deposit (Nom) Visualize All



Status: OK Log x 0

2. Review of Loaded Data



Relation: bank

No.	1: cust Nominal	2: accno Nominal	3: bankname Nominal	4: location Nominal	5: deposit Nominal
1	male	0101	sbi	hyd	yes
2	fem...	0102	hdfc	jmd	no
3	male	0103	sbh	antp	yes
4	male	0104	ab	pdtr	yes
5	fem...	0105	sbi	jmd	no
6	male	0106	ab	hyd	yes
7	fem...	0107	rbi	jmd	yes
8	fem...	0108	hdfc	kdp	no
9	male	0109	sbh	kdp	yes
...	male	0110	ab	jmd	no
...	fem...	0111	rbi	kdp	yes
...	male	0112	sbi	jmd	yes
...	fem...	0113	rbi	antp	no
...	male	0114	hdfc	pdtr	yes
...	fem...	0115	sbh	pdtr	no

Add instance Undo OK Cancel

3. Result after Farthest First algorithm

The screenshot shows the Weka Explorer interface with the 'Cluster' tab selected. The 'Clusterer' dropdown is set to 'FarthestFirst -N 2 -S 1'. In the 'Cluster mode' section, 'Use training set' is selected, and 'Store clusters for visualization' is checked. The 'Clusterer output' pane displays the following information:

```
=== Run information ===  
  
Scheme:      weka.clusterers.FarthestFirst -N 2 -S 1  
Relation:    bank  
Instances:   15  
Attributes:  5  
             cust  
             accno  
             bankname  
             location  
             deposit  
Test mode:   evaluate on training data  
  
=== Clustering model (full training set) ===  
  
FarthestFirst  
=====
```

Cluster centroids:

```
Cluster 0  
      male 0101 sbi hyd yes  
Cluster 1  
      female 0102 hdfc jmd no
```

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

```
=== Model and evaluation on training set ===  
  
Clustered Instances  
  
0      8 ( 53%)  
1      7 ( 47%)
```

The 'Result list' on the left shows a single entry: '06:43:56 - FarthestFirst'. The 'Status' bar at the bottom indicates 'OK'.