Ministry of Education and Science of the Russian Federation
Peter the Great St. Petersburg State Polytechnic University
Institute of Computer Sciences and Technologies
Graduate School of Cyber-Physical Systems and Control

Practice Task – Ch 3

Adaptive Resonance Theory
Discipline: Intellectual Computing
15 February 2017

Student Group: 13541/8

Christopher W. Blake

Professor

Kuchmin A.Y

St. Petersburg
2017

Contents

Christopher W. Blake
February 15, 2017

Introduction

Chapter 3 of "AI Application Programming" by M. Tim Jones is about Adaptive Resonance Theory. This method is essentially used for identification of clusters of groups in data. A dataset is organized as a system of feature vectors such that all items follow a fixed scheme. This scheme is then traversed and most-similar clusters are identified. This is performed via a proximity equation, which determines membership and a vigilance function which determines modifications of the cluster.

A sample C# program has been created to show this methodology. A small dataset of 9 customers and their purchases is provided. This set of 9 customers is then broken into clusters. Finally, the vigilance factor is adjusted to show its effect on the cluster size.

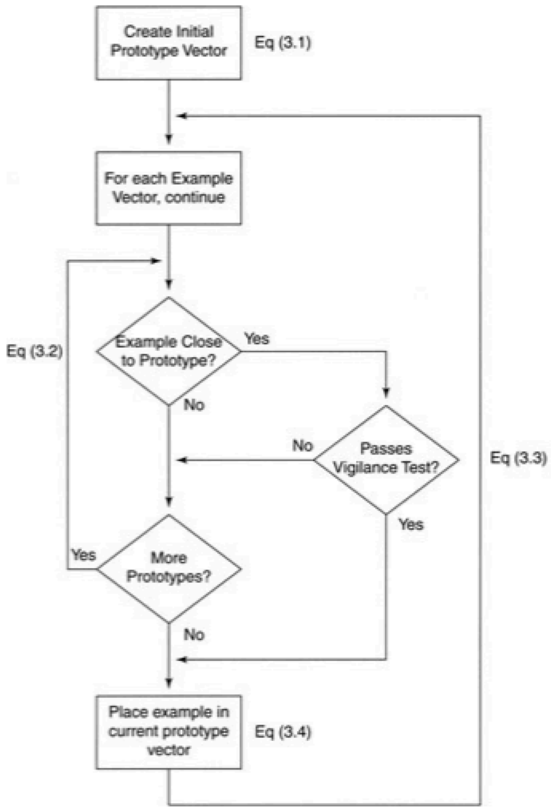
Background

Algorithm

The main algorithm flow can be seen in the figure below (borrowed from the book). Additionally, the steps are described in the below process steps.

Processing Steps

1) Compare all existing customer data to existing cluster vectors.
   a) If passing the proximity test, add it as a member of the cluster.
   b) If passing the vigilance test, update the cluster vector.
   c) If not passing, create a new cluster using the customer feature vector.
2) Repeat process to verify cluster membership.
3) Repeat process until all feature vectors are associated with a cluster and no further changes are required.



Equations

There are two main equations in the ART process. The proximity equation (3.2), which determines membership of a feature vector to a cluster, and the vigilance equation (3.3) which determines necessity of updates to a cluster's feature vector prototype.

Equations Components

d = Size of a feature vector
P = Cluster prototype feature vector
E = Customer feature vector
B = Beta factor for proximity equation
P = Vigilance factor

3.2 Proximity Equation

$$\frac{\|P_i \cap E\|}{\beta + \|P_i\|} > \frac{\|E\|}{\beta + d}$$

3.3 Vigilance Equation

$$\frac{\|P_i \cap E\|}{\|E\|} < \rho$$

Vigilence Parameter Testing

As mentioned in the processing description, the vigilence factor is used for determining updates to the the cluster prototype feature vector. As such, this parameter is varied from 0.1 to 0.9 to test the influence. This is perfomed on the below customer purchases. A one represents a purchased item and a zero is a not-purchased item. The number to the right is the row number for referencing.

```
Column Names
"Hammer", "Paper", "Snickers", "ScrewDriver",  "Pen",  "Kit-Kat",  "Wrench",
"Pencil",  "Heath Bar",  "Tape Measure",  "Binder"

Data
Hmr  Ppr  Snk  Scr  Pen  Kkt  Wrn  Pcl  Hth  Tpm  Bdr
0,   0,   0,   0,   0,   1,   0,   0,   1,   0,   0    //0
0,   1,   0,   0,   0,   0,   0,   1,   0,   0,   1    //1
0,   0,   0,   1,   0,   0,   1,   0,   0,   1,   0    //2
0,   0,   0,   0,   1,   0,   0,   1,   0,   0,   1    //3
1,   0,   0,   1,   0,   0,   0,   0,   0,   1,   0    //4
0,   0,   0,   0,   1,   0,   0,   0,   0,   0,   1    //5
1,   0,   0,   1,   0,   0,   0,   0,   0,   0,   0    //6
0,   0,   1,   0,   0,   0,   0,   0,   1,   0,   0    //7
0,   0,   0,   0,   1,   0,   0,   1,   0,   0,   0    //8
0,   0,   1,   0,   0,   1,   0,   0,   1,   0,   0    //9
```

Results

Initial Results

The initial results are performed with a high vigilance factor (0.9), similar to the book. This is used to ensure realistic results. After the ART clustering operation, the data is shown to be separated into 4 clusters.

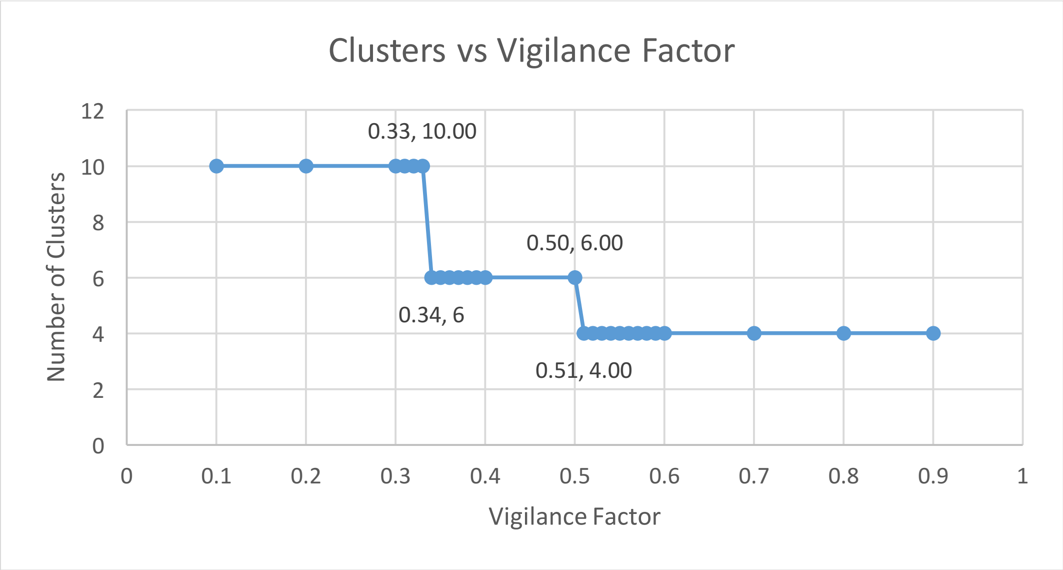
```
CLUSTERS
             Hmr Ppr Snk Scr Pen Kkt Wrn Pcl Hth Tpm Bdr
 Prototype:  0   0   0   0   0   0   0   0   1   0   0
Customer 0:  0   0   0   0   0   1   0   0   1   0   0
Customer 7:  0   0   1   0   0   0   0   0   1   0   0
Customer 9:  0   0   1   0   0   1   0   0   1   0   0

 Prototype:  0   0   0   0   1   0   0   0   0   0   0
Customer 5:  0   0   0   0   1   0   0   0   0   0   1
Customer 8:  0   0   0   0   1   0   0   1   0   0   0

 Prototype:  0   0   0   1   0   0   0   0   0   0   0
Customer 2:  0   0   0   1   0   0   1   0   0   1   0
Customer 4:  1   0   0   1   0   0   0   0   0   1   0
Customer 6:  1   0   0   1   0   0   0   0   0   0   0

 Prototype:  0   0   0   0   0   0   0   1   0   0   1
Customer 1:  0   1   0   0   0   0   0   1   0   0   1
Customer 3:  0   0   0   0   1   0   0   1   0   0   1
```

Summarized Results

The vigilance factor has been modified in 0.1 steps and the number of clusters recorded. Additionally, during the transition, the resolution was increased to 0.01 steps to determine a more exact point when the cluster size changes.

It can be seen in the below chart there are two critical vigilance factor values 0.33 and 0.5. This results in three results and is not continuous as previously expected.

- Less than 0.33: All items are placed in separate groups. i.e. 10 clusters
- Between 0.33 and 0.5: 6 clusters
- Above 0.5: 4 clusters



Conclusion

It can be seen that the vigilance factor makes a significant difference in producing different numbers of clusters from the data. As the vigilance factor is decreased, the number of clusters increases. As this vigilance factor approaches zero, all feature vectors become separated as separate groups and the idea of clustering is lost. Hence it is best to use a higher value (between 0.5 and 0.9) to reduce the number of clusters.

NOTE: It should be noted that this experiment was performed with a limited number of sample data (10 items). Hence better results are likely obtainable if much more data is utilized.