Knowledge Discovery and Data Mining

Unit # 14

Adaptive Resonance Theory

- The ART1 (adaptive resonance theory) algorithm is a simple, unsupervised learning algorithm with biological motivation.
- It works with objects called feature vectors.
- A feature vector is nothing more than a collection of binary values that represent some type of information.
- An example of a feature vector is a customer's purchase data

Hammer	Paper	Pen	Kit-Kat	Pencil	Binder	Snickers
1	0	0	1	0	0	1

Working of ART1

- We being with a set of feature vectors and a set of initialized prototype vectors (P1….PN).
- The prototype vector is the center of the cluster.
- The number of prototype vector, N, is the maximum number of clusters that can be supported.
- We initialize a vigilance parameter (ρ) to a small value between 0 and 1 and a beta parameter to a small positive integer.
- The parameter (d) represents the dimension of the vectors.

ART1 Conditions

- Initially, no prototype vectors exist, so at the start fo the algorithm an initial prototype vector is created with the first example vector.
- We then check all subsequent example feature vectors against each existing prototype vector for its proximity.
- Proximity Test
  - $||P_i \cap E||/(\beta+||Pi||) > ||E||/(\beta+d)$
- Vigilance Test
  - $||P_i \cap E||/(||E||) > \rho$

Example

- P0: {1, 0, 0, 1, 1, 0, 1}
- P1: {1, 1, 0, 0, 0, 1, 0}
- E: {1, 1, 1, 0, 0, 1, 0}
- $\beta = 1.0$, $\rho = 0.6$, d = 7
- Proximity Test with P0 ?
- Proximity Test with P1 ?
- Vigilance Test with P1 ?
- P1 AND E {1,1,0,0,0,1,0} AND {1,1,1,0,0,1,0} = {1,1,0,0,0,1,0}

Example (Cont'd)

- Proximity Test
  - $||P_i \cap E||/(\beta+||Pi||) > ||E||/(\beta+d)$
- Vigilance Test
  - $||P_i \cap E||/(||E||) > \rho$
- P & E = 3, P = 3, E = 4, beta + d = 8
- 3 / 4 > 4 / 8
- 3 / 4 > 0.6

Euclidean ART

- Let
  - Cluster Set = CS= Empty Set
  - Dataset = D = Total customer accounts
  - Data row =R = A single data record
  - Vigilance= v = parameter to control distance between data points in cluster
  - n= Total Number of iterations specified by the user

Euclidean ART (Cont'd)

- While (Iterations < n)
  - Begin
    - Create a single cluster and add it to CS
    - Initialize its centroid to the first record
    - Foreach record R in dataset D
      - Compute Euclidean distance between R and centroid of all the clusters in CS
      - Find the minimum distance and denote it as mindist
      - If the (mindist < v)
        - » Add R into the cluster c and recompute the centroid of c
        - » Increment cluster size of c
      - Else
        - » Create a new cluster and add this new cluster to CS
        - » Set R as the initial centroid of the new cluster
    - End // foreach
    - Reshuffle the items in data set D so as to remove any bias and repeat the above steps
  - End // total iterations

Termination Condition

- As example feature vectors are tested against the prototype vectors, new clusters are created or existing clusters are modified at the inclusion of an example.
- This process, known as “resonance”, indicates the process of learning within the algorithm.
- When the algorithm reaches equilibrium (that is, no further changes occur with the prototype vectors), learning is complete and the data set is classified.

ART Summary

- ART1 is both conceptually simple and easy to implement.
- Earlier algorithms, such a k-means clustering algorithm, though much simpler, have some significant drawbacks.
- For example, k-means does not allow the creation of new clusters (the clusters are statically defined at the start).
- Also, no parameter exists within k-means to adjust the class size of the result clusters.
- A drawback to both algorithms (ART1 and k-means) is that the final set of clusters (and prototype vectors) can be influenced based on the order in which training is performed.