Chapter 5. Adaptive Resonance Theory (ART)

- **ART1**: for binary patterns; **ART2**: for continuous patterns
- Motivations: Previous methods have the following problem:
  1. Training is non-incremental:
     - with a fixed set of samples,
     - adding new samples often requires re-train the network with the enlarged training set until a new stable state is reached.
  2. Number of class nodes is pre-determined and fixed.
     - Under- and over- classification may result from training
     - No way to add a new class node (unless these is a free class node happens to be close to the new input).
     - Any new input $\boldsymbol{x}$ has to be classified into one of an existing classes (causing one to win), no matter how far away $\boldsymbol{x}$ is from the winner. no control of the degree of similarity.

- Ideas of ART model:
  - suppose the input samples have been appropriately classified into k clusters (say by competitive learning).
  - each weight vector $\boldsymbol{W}_{\bullet j}$ is a representative (average) of all samples in that cluster.
  - when a new input vector $\boldsymbol{x}$ arrives
    1. Find the winner j among all k cluster nodes
    2. Compare $\boldsymbol{W}_{\bullet j}$ with $\boldsymbol{x}$

       if they are sufficiently similar ($\boldsymbol{x}$ resonates with class j),

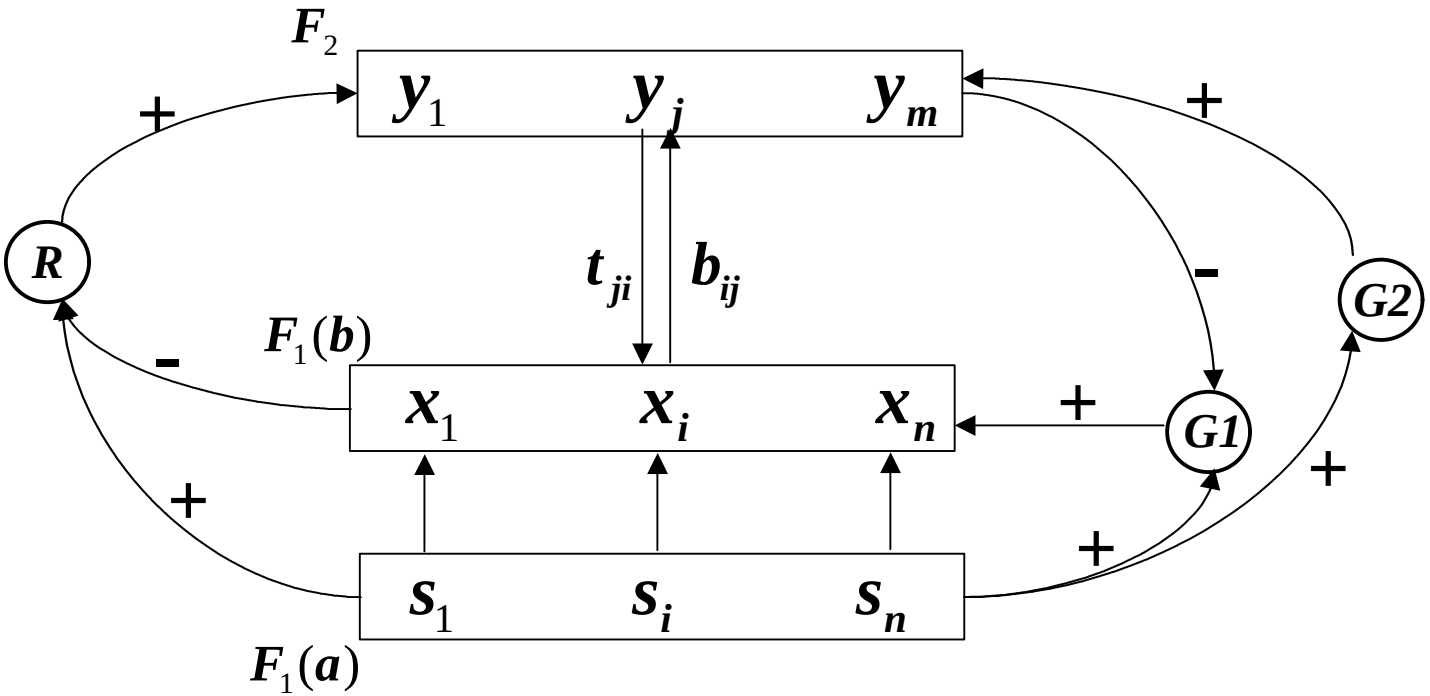
       then update $\boldsymbol{W}_{\bullet j}$ based on $\boldsymbol{x} - \boldsymbol{W}_{\bullet j}$

       else, find/create a free class node and make $\boldsymbol{x}$ as its

       first member.

- To achieve these, we need:
  - a mechanism for testing and determining similarity.
  - a control for finding/creating new class nodes.
  - need to have all operations implemented by units of local computation.

ART1 Architecture



$F_1(a)$: input units

$F(b)_1$: interface units

F_2 : cluster units

$F_1(a)$ to $F_1(b)$: pair - wise connection

between F_2 and $F_1(b)$: full connection

b_{ij} : bottom up weights
from x_i to y_j (real value)

t_{ji} : top down weights
from y_j to x_i (representing class j
binary/bipolar)

R, $G1$, $G2$: control units

- F_2 cluster units: competitive, receive input vector x through weights b: to determine winner j.
- $F_1(a)$ input units: placeholder or external inputs
- $F_1(b)$ interface units:
  - pass s to x as input vector for classification by F_2
  - compare x and $t_{j\bullet}$ (projection from winner y_j)
  - controlled by gain control unit $G1$
- Nodes in both $F_1(b)$ and F_2 obey 2/3 rule (output 1 if two of the three inputs are 1)

  G_1

  Input to $F_1(b): s_i, G1, t_{ji}$ Input to $F_2: t_{ji}, G2, R$
- Needs to sequence the three phases (by control units $G1$, $G2$, and R)

$$G_1 = \begin{cases} 1 & \text{if } s \neq 0 \text{ and } y = 0 \\ 0 & \text{otherwise} \end{cases}$$

$G_1 = 1$: $F_1(b)$ open to receive $s \neq 0$

$G_1 = 0$: $F_1(b)$ open for t_J.

$$G_2 = \begin{cases} 1 & \text{if } s \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$G_2 = 1$ signals the start of a new classification for a new input

$$R = \begin{cases} 0 & \text{if } \dfrac{\|x\|}{\|s\|} \geq r \\ 1 & \text{otherwise} \end{cases}$$

$o < r < 1$ vigilance parameter

$R = 0$: resonance occurs, update $b_{\cdot J}$ and t_J.

$R = 1$: fails similarity test, inhibits J from further computation

Working of ART1

- Initial state: nodes on $F_1(b)$ and F_2 set to zeros
- ***Recognition phase***: determine the winner cluster for input s

  $s \neq 0$ is applied to $F_1(a)$ and stay there (clamped)

  $G_1 = 1 \quad (\because y = 0, s \neq 0)$

  $G_2 = 1 \quad (\because s \neq 0)$

  $F_1(b)$ is open to receive s

  $R = 0 \quad (|x|/|s| = 1 > r\ )$

  F_2 is open to receive $x \bullet b_{\bullet j}$

  determine winner J

  x is tentatively classified to cluster J

  $Y_J = 1 \quad Y_{k \neq J} = 0$

- ***Comparison phase***:

  $\boldsymbol{t_{J\bullet}}$ is sent down from $\boldsymbol{F}_2$

  $\boldsymbol{G}_1 = 0 \quad (\boldsymbol{y} \neq 0)$

  new $\boldsymbol{x}$ appears on $\boldsymbol{F}_1(\boldsymbol{b})$: $\boldsymbol{x_i} = \boldsymbol{s_i} \wedge \boldsymbol{t_{Ji}}$ (2/3 rule)

  if $\|\boldsymbol{x}\| / \|\boldsymbol{s}\| \geq r \quad \boldsymbol{R} = 0$

  a resonance occurs, the classification is accepted

  if $\|\boldsymbol{x}\| / \|\boldsymbol{s}\| < r \quad \boldsymbol{R} = 1$

  classification is rejected

  reset signal is sent (from $\boldsymbol{R}$ to $\boldsymbol{F}_2$)

  $\boldsymbol{y_J}$ is permanently disabled

  all other $\boldsymbol{y_k}$ is set to zero

  goes back to recognition phase

  search for other possible match

- ***Weight update/adaptive phase***
  - Initial weight: (no bias)

    bottom up: $0 < b_{ij}(0) < L/(L-1+n)$ (L usually 2)

    top down: $t_{ji}(0) = 1$
  - When a resonance occurs with y_J, update $b_{\bullet J}$ and $t_{J\bullet}$

    s : current input (on $F_1(a)$)

    x : comparison result between s and $t_{J\bullet}$ (on $F_1(b)$)

    $$x_i = s_i \cdot t_{Ji}$$

    new $t_{Ji} = x_i$ new $b_{ij} = \dfrac{Lx_i}{L-1+\|x\|}$
  - If k sample patterns are clustered to node y_J. then

    $t_{\bullet J}$ = pattern whose 1's are common to all these k samples

    $t_{\bullet J} = s(1) \wedge s(2).....s(k)$

    $b_{\bullet J}$ (new) $\neq 0$ iff $x_i \neq 0$ only if $s_i \neq 0$

    $t_{J\bullet} = x, b_{\bullet J}$ is a normalized $t_{\bullet J}$

– Winner may shift:

$$L = 2,\ b_{ij}(0) = 0.2$$
$$s(1) = (1\ 1\ 1\ 0) \qquad t_{1\bullet} = (1\ 1\ 1\ 0)$$
$$s(2) = (1\ 1\ 0\ 0) \qquad t_{1\bullet} = (1\ 1\ 0\ 0)$$
$$s(3) = (1\ 0\ 0\ 0) \qquad t_{1\bullet} = (1\ 0\ 0\ 0)$$
$$s(1) = (1\ 1\ 1\ 0) \qquad \frac{\|x\|}{\|s\|} = \frac{1}{3} < r$$

– What to do when failed to classify into any existing cluster?
  – report failure/treat as outlier
  – add a new cluster node y_{m+1} with

$$b_{i,m+1} = \frac{L}{L-1+n},\ t_{m+1,i} = 1$$

Notes

1. Classification as a search process
2. No two classes have the same ***b*** and ***t***
3. Different ordering of sample input presentations may result in different classification.
4. Increase of ρ increases # of classes learned, and decreases the average class size.
5. Classification may shift during search, will reach stability eventually.
6. ART2 is the same in spirit but different in details.