17-4

Fuzzy ART Relieving Vigilance

Haruka Isawa, Haruna Matsushita and Yoshifumi Nishio
(Tokushima University)

1. Introduction

Fuzzy ART (Adaptive Resonance Theory) is one of the unsupervised neural networks introduced by G.A.Carpenter and S.Grossberg. Fuzzy ART is applied to association, clustering and memory of input pattern. This model put focus on stability- plasticity dilemma by using a vigilance parameter. In this study, we propose a new type of Fuzzy ART algorithm, which is called Fuzzy ART Relieving Vigilance (R-FART). The important feature of R-FART is that the vigilance parameter varies. We investigate the behavior of R-FART and compare R-FART with the conventional Fuzzy ART.

2. Proposed Fuzzy ART (R-FART)

R-ART is composed of F_1 (input layer) and F_2 (category layer). F_1 and F_2 are connected by the bottom-up-weight vector $\boldsymbol{w}_{ij}$ and the top-down-weight vector $\boldsymbol{w}_{ji}$. m neurons of the input layer F_1 correspond to the an input vector $\boldsymbol{I}$.
Input vector: Each input $\boldsymbol{I}$ is an m-dimensional vector $\boldsymbol{I} = (I_1, \cdots, I_m)$, where $I_i \in [0,1]^m$.
Weight vector: Each category j corresponds to a vector $\boldsymbol{w}_j = (w_{j1}, \cdots, w_{jm})$, $(j = 1, \cdots, n)$ of adaptive weight.
Parameters: R-FART dynamics are determined by a *choice parameter* $\alpha > 0$; a *learning parameter* $\beta \in [0,1]$; and the *vigilance parameter* $\rho(t) \in [0,1]$. The vigilance parameter is fixed value in the conventional Fuzzy ART. However, in R-FART, the vigilance parameter is varied with learning.
(RFART1) An input vector $\boldsymbol{I}$ is inputted to the category layer F_2 from the input layer F_1.
(RFART2) A winning category J is chosen. For the input vector $\boldsymbol{I}$ and category j, *choice function* T_j is defined by

$$T_j(\boldsymbol{I}) = \frac{\mid \boldsymbol{I} \wedge \boldsymbol{w}_j \mid}{(\alpha + \mid \boldsymbol{w}_j \mid)}, \tag{1}$$

where the fuzzy AND operator $\wedge$ and the norm $\mid \cdot \mid$ are defined by

$$(\boldsymbol{p} \wedge \boldsymbol{q})_i \equiv \min(p_i, q_i), \quad \mid \boldsymbol{P} \mid \equiv \sum_{i=1}^{m} \mid p_i \mid. \tag{2}$$

The winning category J, whose T_j is maximum, is found;

$$T_J = \max\{T_j : j = 1 \cdots n\}. \tag{3}$$

If more than one T_j is maximal, the category j with the smallest index is chosen as the winner J.
(RFART3) The similarity of $\boldsymbol{I}$ and the current winning category $\boldsymbol{w}_J$ is measured by the vigilance criterion. We check whether

$$\frac{\mid \boldsymbol{I} \wedge \boldsymbol{w}_J \mid}{\mid \boldsymbol{I} \mid} \geq \rho(t). \tag{4}$$

If Eq. (4) is not satisfied, namely

$$\frac{\mid \boldsymbol{I} \wedge \boldsymbol{w}_J \mid}{\mid \boldsymbol{I} \mid} < \rho(t), \tag{5}$$

a new index J is chosen by Eq. (3). The search process continues until the chosen J satisfies Eq. (4). In R-FART algorithm, the vigilance parameter $\rho(t)$ is varied according to the learning step t as following gaussian function;

$$\rho(t) = \frac{1}{\sqrt{2\pi\sigma}} \exp\{\frac{-(t-\mu)^2}{2\sigma^2}\}. \tag{6}$$

(RFART4) $\boldsymbol{w}_J$ is updated by

$$\boldsymbol{w}_J^{\text{new}} = \beta(\boldsymbol{I} \wedge \boldsymbol{w}_J^{\text{old}}) + (1-\beta)\boldsymbol{w}_J^{\text{old}}, \tag{7}$$

if there is J which satisfies Eq. (4). On the contrary, if all available F_2 nodes do not satisfy Eq. (4), a new category is established in F_2;

$$\boldsymbol{w}_{n+1} = \boldsymbol{I}. \tag{8}$$

(RFART5) The steps from (FART1) to (FART4) are repeated for all the input data.

(Complement coding) To prevent weight vector from monotone decreasing, let input $\boldsymbol{I}$ preprocess into the complement coding form.

$$\boldsymbol{I} = (\boldsymbol{a}\ ,\ \boldsymbol{a}^c) = (a_1, \cdots, a_M,\ a_1^c, \cdots, a_M^c), \tag{9}$$

4. Simulation Results

We apply R-FART to the real world clustering problem. We use the iris data as real data. This data set is widely used to be found in pattern recognition literatures, and contains three classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are not linearly separable from each other. The parameters for the learning are chosen as follows; (For Fuzzy ART) $\alpha = 0.2,\ \beta = 1,\ \rho = 0.8$, (For R-FART) $\alpha = 0.2,\ \beta = 1,\ N(0,1)$. We perform the learning simulation 10 times. Table 1 shows percent recognition for Fuzzy ART and R-FART. The correct answer rate is obtained by;

$$\frac{N_{c_r}}{N_c} \times 100\ [\%], \tag{10}$$

where N_r is the obtained number of the desired input data within the class c and N_c is the total number of the input data belonging to the class c. From the simulation result of Fuzzy ART, Fuzzy ART can not recognize class 3 well because iris data has linearly inseparable classes. In contrast, we can see that the recognition percent is improved in every case from the simulation result of R-FART.

Iris data		Class 1	Class 2	Class 3
Fuzzy ART	Best [%]	88	68	10
	Worst [%]	74	22	2
	Average[%]	80.0	38.6	8.8
R-FART	Best [%]	100	100	72
	Worst [%]	98	54	30
	Average[%]	99.5	92.0	56.2

4. Conclusions

In this study, we have proposed Fuzzy ART Relieving Vigilance (R-FART). We have investigated its behaviors with application to iris data, and confirmed the efficiency of R-FART.