Application of Optimized Neural Network in Data Processing

Jianjun WU
Hunan City University, Yiyang, 413000, CHINA

Abstract: As the ART2 neural network clustering occurs normalization in the data inputting mode by vector and nonlinear transformation pretreatment process is easy to be filtered as a substrate for an important, but a minor component of the noise, while there are still phenomenon of the drifting mode in the learning process due to the correction of the value of weight, this paper proposes an improved method of ART2 neural network. The improved method stores the amplitude information in the learning process, and it is considering the shortest distance of being inputted into the center of the cluster, increasing a threshold limit value for determining outliers at the same time and eliminating the influence of outliers of the clustering results. Finally, the clustering of data samples experimental results show that: the improved ART2 network can handle negative data, the four quadrants of data can be effectively clustered, the performance is superior to the traditional ART2 network.

Keywords: Adaptive resonance; Learning algorithm; Neurons; Resonance

1. Introduction

With the rapid growth of network information, the amount of daily published text has a exponential growth [1]. It indicates that providing an effective way of text organization becomes increasingly important, and clustering of text can do more scientific and reasonable clustering analysis and processing for text data, and it is effective in helping people to get a variety of information what you want [2-4].

Adaptive Resonance Theory is a self-organizing and unsupervised neural network, having a self-organized manner to quickly input pattern recognition and clustering, and it can pick up the background noise signal that put into the various approximation and even carry on to strengthen other advantages.

But when the inputting pattern of the traditional ART2 neural network is undergoing the clustering of recognition, the inputting mode only uses the phase information, while ignoring its effect of amplitude information, that is, when handling the same phase with the different amplitude of the inputting modes clusters, it is so difficult to tell them apart that the clustering result is not satisfactory; there are certain restrictions of choosing the values in the inputting model of each neuron [5-7]. In the layer of F1, a positive real number will be 0 in the inputting mode, resulting in some loss of information in inputting mode, and affecting the clustering results [8-10].

ART2 neural network is designed for random simulated input mode, and it has a very wide range of applications. By warning value's adjustments, ART2 neural network can classify simulated input sample through arbitrary precision. However, because the network has the characteristics of preconceptions; that is, it often has a gradual process of input samples with lower sensitivity, and the initial input network mode will play a decisive role in subsequent input mode if the degree of similarity with the initial mode is high enough, then it is classified into the same category. The small differences with the initial mode will only cause a little change in the memory mode. ART2 network has classified them into the same categories. As the memory mode is continuous tuning, it may lead to making the subsequent input model and the initial mode very different, and it still be classified into the same category, so there is mistake of classification. This is also clearly demonstrated in the use of ART2 modeling. Therefore, there is a need of improvements for traditional ART2 network.

The vector space model (vector space model, VSM), proposed by G. Salton, represented by the example of TD-IDF algorithm to extract the weight term, thereby forming a matrix representation of the text setting. Since many words appeared in the text, the matrix of text features tends to exhibit the huge dimension, resulting in the problems, such as the curse of dimensionality, computational complexity of text clustering, some classical statistical algorithms can not be applied and so on. Wang Li-Juan proposed two fuzzy clustering methods which based on weighted features respectively [11]. They have something in common, that is, before the first, by using a clustering supervised or unsupervised learning process, it has been able to reflect the characteristics of the internal structure of the dataset weight vector [12-13]. On this basis, it formed the weighted feature of distance function. Then, based on the framework of FCM algorithm, the data setting of fuzzy partition is obtained. Weighted feature, in the learning process of the data setting, revealed the structural characteristics of each class, which is conducive to the subsequent clustering process [14-16]. However, in the process of clustering, the feature weights no longer change. Therefore, the clustering process is affected by the constraints of the previous learning outcomes. In order to be valid for the clustering process of text data, people have used a number of effective clustering methods, such as the classic k-means clustering algorithm which was clustering algorithm, based on SOM neural network text. However, these methods often require a lot of previous knowledge to determine the numbers of clusters. It can not dynamically start to learn and learning the new vector will affect the learned vectors and other issues. According to ART2 neural network can dynamically learn efficiently, and realize the balance of memory and learning, but also determine the number of clusters adaptively. But ART2 network remains worthy of improvement, such as the sensitive entry for data of ART2 network will greatly affect the clustering results.

In order to carry out the clustering process to the text data, people have used a number of effective clustering methods, such as the classic k-means clustering algorithm, which wad based on text clustering algorithm of SOM neural network. However, these methods often require a lot of previous knowledge to determine the number of clusters; it is not a way of dynamic learning, and learning the new vector will affect the learnt vectors and other issues. According to the advantages of ART2 neural network, it can efficiently go on the dynamic learning, and it not only realize the balance of memory and learning, but also determine the number of clusters adaptively. But ART2 network remains worthy improvements, such as the entering sensitivity to data will greatly affect the clustering results of ART2 network.

2. The Traditional ART2 Neural Network

2.1. ART2 neural network's structure

The basic idea of ART2 neural network is competitive learning mechanism and stabilized self-learning mechanism. ART2 neural network's structure has two-layer, that is, the layer of F1 and F2, wherein the layer of F1 has n inputting nodes, F2 layer m outputting nodes. There are both feed forward connection weights w i_j (i =1, 2,..., n, j==1,2,..., m) between two layers of the network layer from F1 to F2-layer, and the feedback connection weights t i_j (i =1, 2,..., n, j==1,2,..., m) from the layer of F1 to F2. In addition, the network also includes a reset signal R to achieve F2 layer's resetting. The typical structure of a single neuron in ART2 neural network is shown in Figure 1:

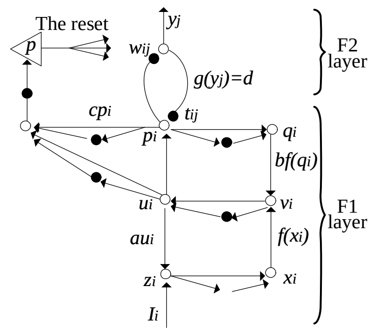

Figure 1. Structure of individual neurons in ART2 neural network

ART2 neural network can also be divided into attention subsystems and orientation subsystem. Attention subsystem is to complete the competition from the bottom up vector selection and comparison of similarity between vectors, the orientation subsystem to check the similarity of satisfactory standard and to do the appropriate action.

2.2. ART2 neural network's learning algorithm

The layer of F1 contains the upper, middle and lower sub-layers. Wherein, the middle sub-layer and the lower sub-layer form a closed loop, the upper sub-layer and the middle sub-layer an another closed loop. When the n-dimensional inputting pattern I= (I1, I 2 , ...,I n) enters into the F1 layer, the nodes have to wait until the output signal is stable, then inputs F1 layer's Short Time Memory (STM) P= (p1,p2,...,p n) into the F2. In calculating the steady STM variable P, it goes through repeated iterations to get the result of u= (u1, u2,...,u n). The absolute value of the results between the former and the latter's difference must be small enough, each iteration is carried on according to the sequence of equation (1) - (6):

$$B_i = I_i + cu_j \quad (1)$$

$$C_i = \frac{B_i}{c+\|x\|} \approx \frac{B_i}{\|x\|} \quad (2)$$

$$x_i = t(x_i) + ng(t_i) \quad (3)$$

In ART2 neural networks, F2 layer is mainly used for category's division; each neuron represents a class. In Figure 1, $w_m \times n$ represents weight matrix of long-term memory from the layer F1 to F2, w_{ij} the connection weights of F1 layer's i neuron and F2 layer's j neuron, the initial value of its elements is set to $1/\left[\left(1-y\sqrt{x}\right)\right]$

After learning, the j column indicates the phase information its class j. $T_m \times n$ represents the weight matrix of long-term memory from the layer of F1 F2; t_{ij} represents the connection weights between the layer of F2's j neurons and F1's i neurons; the initial value is set to 0. When the variable P of STM in F1 enters into F2, it compete with each other to select the neuron j of greatest similarity, namely, $y_j = \max\{l_i\}, i-1,2,...,n, l_i = \sum_{i=1}^{m} t_i l_{il}$

Winning neuron is activated, then the other neurons are in the state of inhibition, namely,

$$t_i = \begin{cases} 1, l_j = \max\{l_j\} \\ 0, \mathit{others} \end{cases} \quad (4)$$

When the strength of the feedback signal from the layer of F1 and F2 is

$$t(u_i) = \begin{cases} t, y_i = \max\{y_i\} \\ 0, \ \mathit{others} \end{cases} \quad (5)$$

3. Improved ART2 Neural Network Method

The traditional ART2 neural network's clustering is based on the phase information, irrelevant to amplitude information. The effect of the traditional ART2 neural network is far from ideal when dealing with the same phase information and amplitude information of two different clusters. Some also are mentioned in the article by comparing the weights of the model and the inputting samples to recover the amplitude information, but this does not reflect the weight of the prototype model's amplitude information, so it is still unable to use the amplitude information.

For the data samples that the original data are both positive and negative, due to the limitations of inputting fields of the traditional ART2 network. In the F1-layer of the traditional ART2 network, non-positive real number of sample data is suppressed to 0, so the traditional ART2 network can not effectively classify data samples locating two, three, four quadrants.

Meanwhile, the traditional ART2 neural network is not sensitive to the presence of outliers. In order to try to eliminate outliers' influence on clustering results, this improved algorithm, by using the outlier as an additional class, reduce the effects of outlier on the clustering results.

As for such existing shortcomings in the traditional ART2 neural network, in the inputting activation process, the study calculates the shortest distance to each cluster from the center, considering its amplitude information. Only when both the values of phase and amplitude exceed the corresponding threshold limit alerting value, the resonance occurs and the adjustment of the weights begins. So does non-linear transfer function. So much so that it can properly handle negatives' input and retain its negative form after the stability of F1-layer, avoiding loss of information in the inputting mode; in order to eliminate the impact of outliers on clustering results, the paper also carried out on the sure of outliers in inputting mode and increased a threshold limit value $R-dis$ to detect outliers. Improved ART2 neural network matches the phase and amplitude as shown in Figure 2.

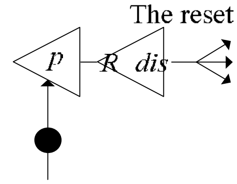

Figure 2. The matching of phase and amplitude of improved ART2 neural network

Improved ART2 network also includes attention subsystem and orientation subsystem. Attention subsystem includes STM, two short storage units of F1 and F2, and long storage unit LTM connecting F1 and F2 layers, that is, connected weight vectors $w_n \times m$ and $T_n \times m$. In this time, the connection weight vector w_{ij} from the bottom to up is a record of the amplitude information of the cluster's center; the j denotes $j's$ center point. The role of the orientation subsystem is to calculate the phase degree of matching between the inputting mode and the memory mode, that is, the matching extent between intermediate mode U that the F1-layer is stable and the winning neuron's feedback mode P from the top to down is compared, as well as detecting inputting mode's outliers. It is for determining the next action of network: resonate or reset.

First, initialize the network's settings. In the improved ART2 network, the initialization of F1-layer from the top to down and the initialization of weight vector $T_n \times m$ is the same with the traditional ART2 network. The number of clusters m is set to 1, the connection weights vector $w_n \times m$ from the bottom to up initializes the first inputting pattern as the first cluster center, i, e.

$w_n \times m = (t_1^1, t_2^1, ..., t_n^1)$

It also need set two threshold limit values ρ and $R-dis$. ρ is the alerting value of the phase matching, $R-dis$ as outliers' alerting values for determination.

When the n-dimensional inputting mode t= (t_1, t_2 ,..., t_n) entering into F1-layer, the formula (1) to (6) F1 layer can calculate the steady states of F1-layer. As non-positive real number in the traditional network ART2 is unified t_y processed as 0, nonlinear transferring function takes it as a noise to deal with, making the network lost that part of the information and affecting the whole results of the clustering. Therefore, there is a need to adjust non-linear processing function to correctly handle non-positive real numbers, to prevent the misuse of the useful information. The non-linear processing function is adjusted to

$$u(x)=\begin{cases}\frac{2\vartheta t^2}{t^2+\vartheta^2}, 0\le t\le\vartheta \\ t, |t|\ge\vartheta \\ -\frac{2\vartheta t^2}{t^2+\vartheta^2}, -\vartheta\le t\le 0\end{cases} \quad (6)$$

Or

$$u(x)=\begin{cases}0, 0\le|t|\le\vartheta \\ t, |t|\ge\vartheta\end{cases} \quad (7)$$

When F1-layer reaches a steady state, the inputting mode I connects with F2-layer through a bottom-up weight vector $w_n \times m$ and conducts competitive learning for finding neurons with the shortest distance as the winning neuron, i,e.

$$w_j=\min\{w_i\}, i=1,2,...,n, w_j=\sum_{i=1}^{m}\left\|i_1-w_{ij}\right\|$$

Winning neuron is activated, while the other neurons are in the state of inhibition. F2-layer selects the winner neuron j and returns a feedback signal, calculating the degree of phase matching ‖R‖ between the processed STM signal U of F1-layer and the feedback value LTM P of active neuron signal. Since ‖R‖ reflects the overall matching degree between U and P, regardless of the difference between P and U, which are the various components. Equation (11) is used for phase-matching calculation in the paper. If ‖R‖ is greater than the setting threshold limit value ρ, then the inputting mode is determined by the outliers; if k_j is larger than a presetting threshold limit value $R-dis$, then the inputting mode is processed as the outlier point, and the inputting mode is considered as a separated class; the inputting mode is classified into the class j. When the network enters into the learning phase, bottom-up weight vector w_{ij} is updated to a new class $j's$ center point. The effect is to be that it is the average value of data samples' weights, weight vector t_{ij} from the top to down is to be updated according to formula (13).

As for the traditional ART2 network, when the inputting mode enters into F1-layer for resonance to make the F1-layer is in a stable state, there is no feedback information from F2-layer at that time. In fact, only 1 to 2 times is needed to make F1 layer go into resonated steady state, and then enter the F2-layer, computing the similarity of m neurons in F2-layer, the neuron with maximum similarity wins; the winning neuron feedbacks a signal and starts matching calculations.

If the degree of matching is less than a predetermined threshold limit value, then reset F2-layer and find out the neuron with the second similarity. In the worst case, the number of resetting F2-layer is m times, otherwise the network goes into the learning phase. Therefore, the time complexity of the algorithm, O (mn), wherein, n is the number of the inputting mode; improved ART2 network still requires the resonance of only 1 to 2 times to reach a steady state in F1-layer. In that case, the inputting mode in stable F1-layer is transmitted to layer-F2 for calculating a distance among m neurons; neuron with shortest distance will win; A winning neuron feedbacks signal and starts matching calculation. If the matching degree is smaller than the preset threshold limit value, then reset F2-layer; find out neurons with the second largest similarity. The worst case is that the number of resetting F2-layer is m times, otherwise the network o_{nc} carries on the outliers determination. The determination times of the improved algorithm F2-layer is more than traditional ART2's, that is, the times of resetting F2 layer in the worst case is m +1. Therefore, the time complexity of the algorithm is still O (mn).Although the complexity of algorithm is unchangeable in the order of magnitude, it is not the same phase in the process of the case. And the accuracy has improved significantly, compared with the traditional ART2 network.

4. Experiment and Analysis

In this paper, the above algorithm for the horizontal and vertical coordinates in [0, 1] is generated randomly within five characteristics distinct groups, each containing 30 data samples cluster, using traditional ART2 and improved ART2 network respectively to cluster the data sample. The clustering results are shown in Figures 3 and 4, wherein each parameter is set as shown in Table 1, where the parameters a, b, c, d, e values of both parameters, these parameters may also be determined by the experience of experiments.

Table 1. Setting Table of Network Parameters

	a	b	c	d	e	t	$R-dis$
Traditional ART2	10	10	0.12	0.93	0	0.98	×
Improved ART2	10	10	0.12	0.93	0	0.94	0.21

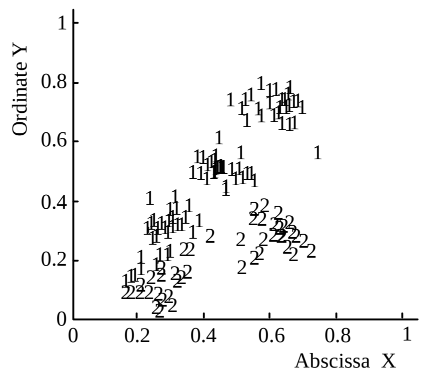

Figure 3. Clustering results of traditional ART2 network

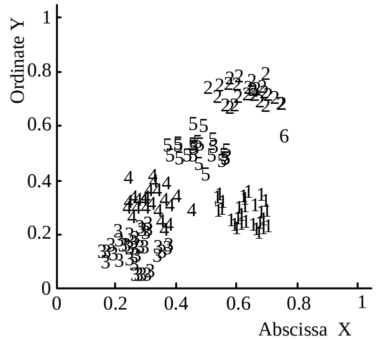

Figure 4. Clustering results of improved ART2 network

Conventional ART2 network normalized sample data in the process of data, keeping only the phase information of the data. The phase information clustered through the competitive learning, the clustering results of data samples just considered phase information and ignored amplitude information of the data, putting the same or similar phase information of sample data into the same class. It can be seen that from Figure 3, the data samples with the same or similar phase are divided in the same class; the phase can not be identical or similar to distinguish two classes. Improved ART2 network, the data processing process not only make the data normalized, while still retaining the amplitude information of the data prototypes. In the process of competitive learning, by a combination of both amplitude and phase information, it can cluster the same phase and amplitude of the two different classes correctly and effectively. It can also be seen that from Figure 4, the improved ART2 network id able to effectively identify outliers; category 6 indicates the shortest distance of the data point to the center with the other five categories is larger than the setting threshold limit value $R-dis$, the data points are treated as outliers.

For the data samples that the original data are both positive and negative, due to the limitations of inputting fields of the traditional ART2 network. In the F1-layer of the traditional ART2 network, non-positive real number of sample data is suppressed to 0, so the traditional ART2 network can not effectively classify data samples locating two, three, four quadrants.

This paper made a comparative analysis of data samples' clustering as to the traditional ART2 network and improved ART2 network which are located in the four quadrants. The data samples, the horizontal and vertical coordinates in [0, 1], are generated randomly within five characteristics distinct groups, each containing 30 data samples cluster. The settings of each network parameters are shown in Table 1. As it can be seen from Figure 5, the effect of the traditional ART2 network is poor in four quadrants clustering of data samples. While Figure 6 is a network using the improved ART2 to cluster, it is clear that the improved ART2 network can classify data effectively in the four quadrants.

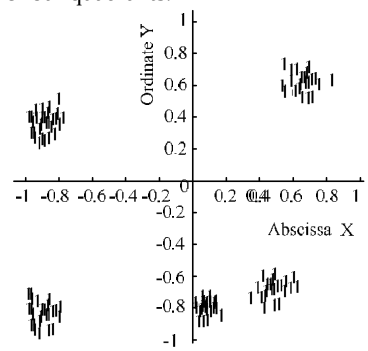

Figure 5. Traditional ART2 data clustering in the four quadrants

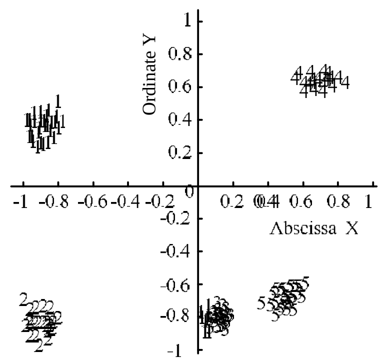

Figure 6. Improved clustering of data samples in four quadrants

5. Conclusion

Through the theoretical and experimental results, it shows that when the improved ART2 neural network is in the same phase of the two clusters, the performance is better than the traditional ART2. Meanwhile, the network takes the phase information of data and amplitude information of a prototype data into account and eliminates outliers of the clustering results. By changing the nonlinear transforming function, an improved ART2 network can handle negative data, and the four quadrants of the data can be efficiently clustered. The experiments show that the improved ART2 network is to be significantly better than the traditional ART2 network in dealing with outliers amplitude information and data samples' performance.

References

[1] Xin Huang, Xiao Ma, Bangdao Chen, Andrew Markham, Qinghua Wang, Andrew William Roscoe.Human Interactive Secure ID Management in Body Sensor Networks. Journal of Networks, Vol 7, No 9 (2012), 1400-1406

[2] Jian Wu, Jie Xia, Jian-ming Chen, Zhi-ming Cui, "Moving Object Classification Method Based on SOM and K-means. Journal of Computers", vol.6, no.8 , pp.1654-1661, 2011.

[3] Jimin Li, Wei Zhang, KunLun Li, "A Novel Semi-supervised SVM based on Tri-training for Intrusion Detection", Journal of Computers, vol. 5, no. 4, pp. 638-645, 2010.

[4] Tan x, Triggs B. Enhanced local texture feature sets for face recongnition under difficult illuminationing conditions [J]. IEEE Trans on Image Processing, 2010, 19(6): 1635-1650.

[5] Kuo-Feng Huang, Shih-Jung Wu, Real-time-service-based Distributed Scheduling Scheme for IEEE 802.16j Networks. Journal of Networks, Vol 8, No 3 (2013), 513-517

[6] Provenzi E Gatta C Fierro M et al. A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(10):1757-1770.

[7] Zhou S, Aggarwal G, Chellapa R. Appearance characterization of linear lamebrain object, Generalized photometric stereo and illumination- Invariant face recognition [J]. IEEE Trans on PAMI, 2007, 29(2): 230-245.

[8] Muhammad J. Mirza, Nadeem Anjum.Association of Moving Objects Across Visual Sensor Networks.Journal of Multimedia, Vol 7, No 1 (2012), 2-8

[9] Haiping Huang, Hao Chen, Ruchuan Wang, Qian Mao, Renyuan Cheng.(t, n) Secret Sharing Scheme Based on Cylinder Model in Wireless Sensor Networks.Journal of Networks, Vol 7, No 7 (20 12), 1009-1016

[10] Muhammad J. Mirza, Nadeem Anjum,Association of Moving Objects across Visual Sensor Networks. Journal of Multimedia, Vol 7, No 1 (2012), 2-8

[11] Zhao Liangduan, Zhiyong Yuan, Xiangyun Liao, Weixin Si, Jianhui Zhao.3D Tracking and Positioning of Surgical Instruments in Virtual Surgery Simulation. Journal of Multimedia, Vol 6, No 6 (2011), 502-509

[12] Young k P, Seokl P, Joong K K. Retinex method based on adaptive smoothing for illumination base on adaptive smoothing for illumination invariant face recognition[J]. Signal Processing, 2008,88: 1929-1945.

[13] Guang Yan, Zhu Yue-Fei, Gu Chun-Xiang, Fei Jin-long, He Xin-Zheng, A Framework for Automated Security Proof and its Application to OAEP. Journal of Networks, Vol 8, No 3 (2013), 552-558

[14] R. Berangi, S. Saleem, M. Faulkner, et al. TDD cognitive radio femtocell network (CRFN) operation in FDD downlink spectrum. IEEE, 22nd International Symposium on Personal, Indoor and Mobile Radio Communications, 2011: 482-486

[15] Land E. The Retinex[J]. American Scientist, 1964, 52(1): 247-264.

[16] Kasman Suhairi, Ford Lumban Gaol, The Measurement of Optimization Performance of Managed Service Division with ITIL Framework using Statistical Process Control. Journal of Networks, Vol 8, No 3 (2013), 518-529

[17] Aricit,D Ikbas S, Altunbasak Y. A histogram modification framework and its application for image contrast enhancement [J]. IEEE Trans on Image Processing, 2009, 18(9):1921-1935

[18] Q. Zhang, B. Li, "Discriminative K-SVD for dictionary learning in face recognition", Proceeding(s) of Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2691-2698, 2010.

[19] Mao X. Nguyen, Quang M. Le, Vu Pham, Trung Tran, Bac H. Le , "Multi-scale Sparse Representation for Robust Face Recognition", Proceeding(s) of Conference on Knowledge and Systems Engineering (KSE), pp. 195-199, 2011.

[20] Yikui Zhai, Junying Gan, Jingwen Li, "Study of occluded robust face recognition approach based on homotopy algorithm and color information fusion", Signal Processing, vol. 21, no. 11, pp.1762-1768, 2011.

[21] Yang M,Zhang Lei, "Gabor Feature based Sparse Representation for Face Recognition with Gabor Occlusion Dictionary", Proceedings of the European Conference on Computer Vision, ECCV, pp. 448-461, 2010.