Application of Neural Network based on Clustering Analysis

Haogui CHEN
Hunan City University, Yiyang, Hunan, 413000

Abstract: As the 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 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].

Neural network is designed for random simulated input mode, and it has a very wide range of applications. By warning value's adjustments, 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. 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 modeling. Therefore, there is a need of improvements for traditional 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 cannot be applied and so on. Wang Li-Juan proposed two fuzzy clustering methods which based on weighted features respectively. 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. 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. 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 cannot 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.

2. The Traditional ART2 Neural Network

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).

Wherein, the function u (x) of the equation (3) is a non-linear processing function for the transmitted signal's nonlinear transformation. The combination of non-linear processing and normalization processing determines the noise criteria, and can tell the noise from the network. The degree of the nonlinear for non-lineal processing function u (x) determines the degree of contrast enhancement and noise suppression, the function u (x) is defined as

$$u(x)=\begin{cases}\dfrac{2\lambda y^2}{y^2+o^2}, 0\le t\le\varphi \\ y, y\ge\varphi\end{cases} \quad (1)$$

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,

$$y_j=\max\{l_i\}, i-1,2,...,n, l_i=\sum_{i=1}^{m}t_i l_{il} \quad (2)$$

3. Improved ART2 Neural Network Method

The main idea of this method is: when the data sample entering into F1-layer, after F1-layer's self-stabilizing, amplitude information prototype Ii identifies the winning neuron by conducting competitive learning; neurons with the shortest distance is to win. The winning neuron feeds back a signal strength from the top to down, using the feedback information P= (p1, p2,...,p n) and the phase information u= (u1, u2,..., u n) of stable F1 layer

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 .

$$u(x)=\begin{cases}0, 0\le|t|\le\vartheta \\ t, |t|\ge\vartheta\end{cases} \quad (3)$$

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.

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 network respectively to cluster the data sample.

The clustering results are shown in Figures 4 and 5, 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	11	12	0.14	0.95	0	0.1	×
Improved ART2	13	12	0.15	0.96	0	0.98	0.36

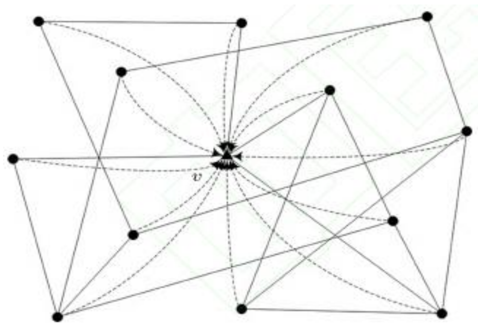

Figure 1. Clustering results of traditional ART2 network

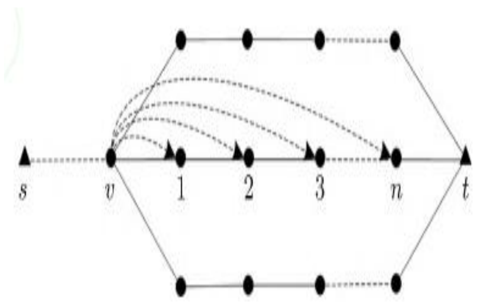

Figure 2. Clustering results of improved ART2 network

Conventional 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 4, the data samples with the same or similar phase are divided in the same class; the phase cannot be identical or similar to distinguish two classes. Improved 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 5, the improved 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 network. In the F1-layer of the traditional network, non-positive real number of sample data is suppressed to 0, so the traditional network cannot effectively classify data samples locating two, three, four quadrants.

This paper made a comparative analysis of data samples' clustering as to the traditional network and improved 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 6, the effect of the traditional ART2 network is poor in four quadrants clustering of data samples. While Figure 7 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.

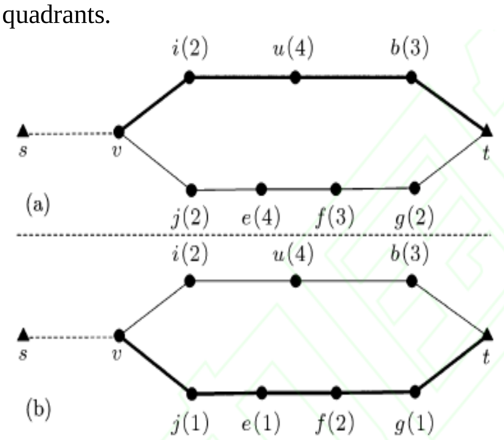

Figure 3. Traditional ART2 data clustering in the four quadrants

5 . Conclusions

Through the theoretical and experimental results, it shows that when the improved neural network is in the same phase of the two clusters, the performance is better than the traditional. 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 network is to be significantly better than the traditional 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