Data Clustering Using Hybrid Neural Network

Donghai Guan, Andrey Gavrilov, Weiwei Yuan, Sungyoung Lee, Young-Koo Lee
Dept. of Computer Engineering, Kyung Hee University
e-mail : donghai@oslab.khu.ac.kr

Abstract

Clustering plays an indispensable role for data analysis. Many clustering algorithms have been developed. However, most of them suffer poor performance of learning. To archive good clustering performance, we develop a hybrid neural network model. It is the combination of Multi-Layer Perceptron (MLP) and Adaptive Resonance Theory 2 (ART2). It inherits two distinct advantages of stability and plasticity from ART2. Meanwhile, by combining the merits of MLP, it improves the performance for clustering. Experiment results show that our model can be used for clustering with promising performance.

1. Introduction

Clustering plays an indispensable role for data analysis. It aims to organize a collection of data items into clusters, such that items with a cluster are more similar to each other than t hey are to items in the other clusters. In clustering, no labeled data are available

Stability and plasticity are important for clustering. For most existing methods, they are either stable but not capa ble of forming new clusters, or plastic but unstable. The conflict between stability and plasticity is called the stab ility-plasticity dilemma.

Adaptive Resonance Theory is a clustering method which is s pecially designed to overcome the stability-plasticity dilemma [1].

Considering the distinct merits of ART, in this work, we pr opose a hybrid neural network (HNN) based on it. The main motivation is to improve the clustering performance of ART, meanwhile, inherit stability and plasticity from it.

HNN is the sequential combination of Multi-Layer Perceptron (MLP) and ART. ART is the key component in HNN and it provides clustering function. MLP is the secondary component used for data preprocessing. Data preprocessing here refers to data transformation. When data pass through the hidden layers of MLP, its value and dimension might be changed. Then the changed data will be used by ART for clustering instead of original data.

2. Adaptive Resonance Theory

Adaptive Resonance Theory (ART) is a family of different ne ural architectures. The first and most basic architecture is AR T1. ART1 can learn and recognize binary patterns. ART2 is a n analog version of ART, it can cluster real-valued input vectors. Our proposed HNN model is the combi nation of MLP and an ART system. In this work, we will pres ent HNN based ART2. The ART2 neural network was chosen for its ability to process both binary-valued and analog-valued input patterns. Original ART2 is rather complex and a bout ten parameters need to be initialized for clustering. This hinders ART2 to be widely used in concrete applications. To make our hybrid model more applicable and clearer, we com bine MLP with the simplified ART2 instead of original one. I n the remaining of this paper, ART2 refers to its simplified ve rsion.

As shown in Fig. 1, ART2 consists of two layers, input layer and output layer. There are n neurons in input layer and m neurons in output layer. Each neuron in output layer represents one class. When an input data item X $(X \in R^n)$ is inputted to ART2, ART2 will try to cluster X to the most relevant class among m classes.

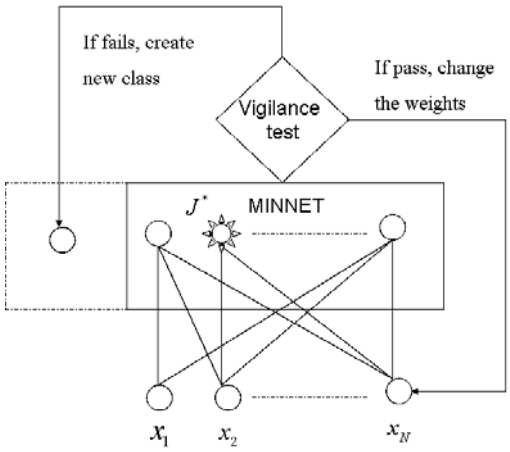

Fig. 1. Architecture of Adaptive Resonance Theory 2

3. Our Model: Hybrid Neural Network

As shown in Fig. 2, our proposed hybrid neural network is a combination of MLP and ART2 [2] with MLP in front and ART2 back. When it is used for data clustering, the unlabeled data will be sent to the input layer of MLP first. Then the output of MLP will be the input of ART2.

In HNN, MLP could be treated as a data preprocessing layer, because it can provide data conversion through its hidden layers. Appropriate data conversion depends on the weights of MLP. In our model, error back propagation algorithm (EBP) will be used for training MLP. Usually we are familiar with MLP training in supervised learning. We know that the labeled data will be used to train MLP. The purpose of training is to provide an accurate characterization of unobserved samples, so long time training is always needed. It should be noted that training in our model is totally different with it in supervised learning, because our model is devised for unsupervised learning. In our model, the teacher of MLP is ART. And the goal of training is to provide some additional help to ART through data

transformation. The different goal of training determined long time training is avoided in our model.

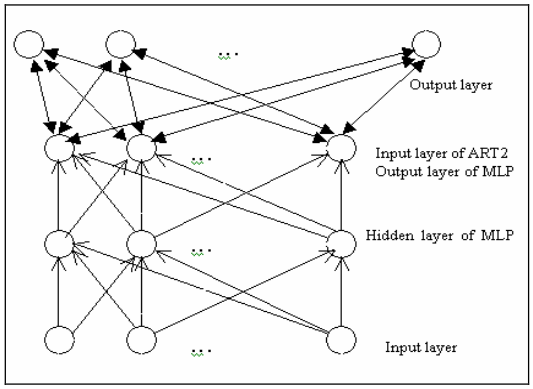

Fig. 2 Structure of our hybrid neural network

Clustering algorithm of HNN proceeds as follows:
Stage 1: HNN initialization.
1) MLP initialization.
2) ART2 initialization. (Num. of output neuron is zero)
Stage 2: Clustering
This part is almost same with clustering in ART2.
One additional part is MLP training by error back propa gation (EBP).
In detail, for each sample, the output of ART2 is treate d as the expected output of MLP, while the output of MLP is the real output. Then, the difference between th em will be used for MLP training.

4. Experiment Result

The dataset in this part is iris, which is one of the most popul ar data sets to examine the performance of novel methods in pattern recognition and machine learning. There are three cat egories in the data set, each having 50 patterns with four feat ures.

Using our proposed method, 4 ones among the 150 p atterns are mis-clustered. Hence the Percentage of error s is 2.7%. This result is superior to most of clustering r esults reported in the literature [3].

The parameters used in our experiment is shown in T able 1.

Table 1. Parameters in the experiment

MLP	1 hidden layer; 4 neurons in hidden layer
	4 neurons in output layer
	Sigmoid activation function, with a=1
	Learning rate=0.1
	Iterations=1
ART	Vigilance value R=0.08

5. Conclusions

In this paper, we propose a new data clustering method. It is a hybrid neural network, which combines Multi-Layer Perceptron and Adaptive Resonance Theory 2. The exp eriment on Iris shows that compared with most existing meth ods, our method can provide better clustering accuracy.

Reference

[1] Gail A. Carpenter. “Neural Network Models for Pattern Recognition and Associative Memory”, In IEEE Transaction on Neural Networks, (1989) 243-257
[2] G.A. Carpenter and S.Crossberg. “ART2: Self-organization of stable category recognition codes for analog input patters”, In Appl. Opt., Vol. 26, (1987) 4919-4930
[3] R. Xu and D. Wunsch. “Survey of Clustering Algori thms”, In IEEE Transaction on Neural Networks, (2005) 645-678