A PSO Based Incremental Hyper-Sphere Partitioning Approach to Classification

Binge Zheng, Department of Computer Science and Software Engineering, Liverpool University, Jiangsu, China

Sheng-Uei Guan, Department of Computer Science and Software Engineering, Xian Jiaotong-Liverpool University, Jiangsu, China

Jinghao Song, Department of Computer Science and Software Engineering, Liverpool University, Jiangsu, China

ABSTRACT

This paper proposes an incremental hyper-sphere partitioning approach for classification problems. Hyper-spheres that are close to the classification boundaries of a given problem are searched using an incremental approach based upon Particle Swarm Optimization (PSO). This new algorithm is proposed to tackle the difficulty of classification problems caused by the complex pattern relationship with a simplified expert rule structure. We solve classification problems through a combination of hyper-sphere partitioning and a Euclidean-distance based partitioning approach. Moreover, an incremental approach combined with output partitioning and pattern reduction is applied to cope with the curse of dimensionality. The algorithm is tested with seven datasets. The experimental results show that this proposed algorithm outperforms ILEGA (our former research work) and normal GA significantly in the final classification accuracy. In terms of the time complexity, it also gains significant improvement in comparison with ILEGA.

Keywords: Algorithm, Classification, Hyper-Sphere, Incremental Approach, Particle Swarm Optimization

INTRODUCTION

This section gives a brief literature review of previous work on classification problems.

In the machine learning field, Evolutionary Algorithms, Neural Networks, Fuzzy Logic and some other methods were widely used to derive solutions for classification problems. Data classification is probably the most applied supervised data mining technique. This technique starts from a set of learning instances labeled with known classes and then form a matching rule from instances to corresponding classes, which can be used to classify new instances.

Particle Swarm Optimization (PSO) is regarded as a typical algorithm among evolutionary algorithms. It is a heuristic technique which is suited for the search of solutions based
on the concept of swarm. PSO is known from the literature to be effective in solving optimization problems. However, so far there exist only a few papers that applied PSO to classification problem solving. Classification problems can be easily converted to multi-attribute optimization problems, which could be solved effectively by PSO. In the multi-dimensional space of a particular problem, each class has a centroid that can be represented in the form of coordinates. Therefore, the original classification problem is converted to optimize the centroid coordinates for each class.

In the research work of Sousa, Silva and Neves (2004), a comparison of PSO, Genetic Algorithm and a Tree Induction algorithm for classification in database was reported. PSO was applied to classify images by Omran et al. (2002). A binary version of PSO was proposed for classification in Cervantes, Galván and Isasi’s (2005) work. PSO was also used in the construction of classification rules by Gandhi et al. (2010). Moreover, according to the research of Falco et al. (2007), PSO was used to classify patterns through comparisons of their Euclidean-distance from cluster centroids.

Furthermore, many ideas adopted in previous work of GA might be also useful in PSO, one of which is the incremental approach. To reduce the complexity of deriving a reasonable solution for classification, an incremental approach is adopted for GA-based classifiers in a dynamic environment where training samples or new attributes may become available over time (Zhu & Guan, 2005). Similar incremental approach has also been exploited with artificial neural network to enhance the intelligence of the learnt models (Guan & Li, 2001). Incremental learning methods were proposed by Yamauchi et al. (1999) earlier for a different purpose, i.e. incremental pattern learning. In their approach, a small part of past learnt patterns will be learnt with new patterns. In contrast, the algorithm introduced in this paper will only incrementally add new experts and learnt patterns will not be learnt so that the training complexity will be reduced.

In this paper, a PSO-based incremental hyper-sphere partitioning approach (IHSP-PSO) for classification is proposed. This method uses hyper-sphere to partition the problem space incrementally until all patterns are learnt. The classification results are compared with ILEGA, which is a GA-based algorithm proposed in an earlier work (Yang et al., 2013). This algorithm is effective in training complex solutions and performs better in higher-dimension problems according to the final experimental results we got in comparison with ILEGA and normal GA.

The structure of this paper is as follows. Section II presents the working scheme of the original PSO. Section III illustrates the design and encoding of the proposed algorithm. The experimental results upon 4 artificial datasets and 3 benchmark datasets from UCI are reported in Section IV. Finally, in Section V the conclusions and suggestions for further studies are presented.

PARTICLE SWARM OPTIMIZATION

This section gives a brief introduction to the original PSO.

Particle swarm optimization is inspired from the observation of social behaviors in various animal groups such as bird flocking and fish schooling. Reynold (1987), Heppner and Grenander (1990) presented the simulations of bird flocking. Sociobiologist Wilson (1975) gave the hypothesis that during the process of searching food, individuals of a fish school could take advantage of the experience gained by other group members. This statement suggests that sharing of information among social members might offer evolutive advantages. These have been fundamental to PSO’s development.

In the PSO model, a swarm consists of n individuals. Each member of the swarm is regarded as a particle and the location of food is treated as the minimum or maximum value of a function. Each particle represents a possible solution to a problem with N dimensions. The original equations to update the velocity and
Related Content

A Rough-Sets Approach to Kansei Evaluation Modeling and Design Support
Hongli Ju and Yoshiteru Nakamori (2012). *Systems Approaches to Knowledge Management, Transfer, and Resource Development* (pp. 296-307).
[www.igi-global.com/chapter/rough-sets-approach-kansei-evaluation/68226?camid=4v1a](www.igi-global.com/chapter/rough-sets-approach-kansei-evaluation/68226?camid=4v1a)

Self-Evolvable Protocol Design Using Genetic Algorithms
[www.igi-global.com/article/self-evolvable-protocol-design-using/40903?camid=4v1a](www.igi-global.com/article/self-evolvable-protocol-design-using/40903?camid=4v1a)

Opinion Dynamics: A Multidisciplinary Review and Perspective on Future Research
[www.igi-global.com/article/opinion-dynamics-multidisciplinary-review-perspective/61135?camid=4v1a](www.igi-global.com/article/opinion-dynamics-multidisciplinary-review-perspective/61135?camid=4v1a)
An Interval Linguistic Distribution Model for Multiple Attribute Decision Making Problems with Incomplete Linguistic Information
