International Congress Series 1291 (2006) 189–192

www.ics-elsevier.com

A hybrid real-coded genetic algorithm with forgetting and its applications

H. Zhang *, M. Ishikawa

Department of Brain Science and Engineering, Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu, Kitakyushu 808-0196, Japan

Abstract. This paper proposes a hybrid real-coded genetic algorithm with forgetting for improving the generalization ability of classification models. A crucial idea here is the introduction of structural learning with forgetting into a hybrid real-coded genetic algorithm. The proposed method has two advantages: (1) finding near optimal classification models efficiently by a hybrid technique and (2) improving the generalization ability of the resulting classification models by the forgetting technique. Applications of the proposed method to an iris classification problem well demonstrate its effectiveness. Our results indicate that it has not only high learning performance for training data, but also high generalization ability for the test data compared with conventional algorithms such as backpropagation and structural learning with forgetting.

Keywords: Real-coded genetic algorithm; Multilayer preceptron; Local search; Structural learning with forgetting

1. Introduction

Finding good classification models is an important issue in many application disciplines such as pattern recognition. The goal of designing classification models is to achieve the best classification performance in the task concerned. Numerous studies on data mining have been done in machine learning, Neural Networks [1,2,4,5].

For solving real-valued optimization problems efficiently, we have proposed a hybrid real-coded genetic algorithm with local search (HRGA/LS) [6,7]. Since a hybrid technique of global search and local search is adopted, the resulting classification models have high learning performance. However, their generalization ability is not so high even with the group search.

* Corresponding author. Tel./fax: +81 93 695 6112.
E-mail address: zhang@brain.kyutech.ac.jp (H. Zhang).

0531-5131/
doi:10.1016/j.ics.2006.01.040

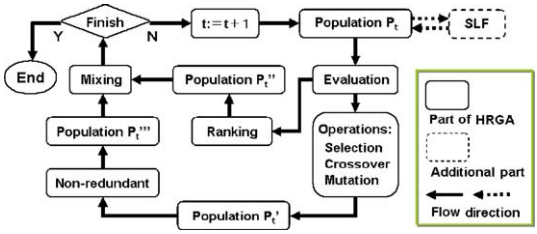

Fig. 1. Flow chart of hybrid real-coded genetic algorithm with forgetting.

The backpropagation (BP) is the most frequently used learning method in Neural Networks for acquiring classification models. Due to its characteristics of distributed representation, the generalization ability of the resulting classification models, however, is not high in general. In contrast to BP, structural learning with forgetting (SLF) [1,2] aims at getting skeletal structured networks reflecting regularities in data by eliminating unnecessary connections. It has demonstrated the effectiveness in classification of mushrooms, MONKs and so forth.

Therefore, we propose to combine HRGA/LS and SLF, named a hybrid real-coded genetic algorithm with forgetting (HRGA/F), for improving the generalization ability of classification models by HRGA/LS. The proposed method has two advantages: (1) finding an optimal classification model efficiently by a hybrid technique and (2) improving the generalization ability of obtained classification models by the forgetting technique.

To demonstrate the effectiveness of HRGA/F, simulation experiments of iris classification are carried out. We investigate how to determine the amount of forgetting for better generalization ability and show experimental results.

2. Hybrid real-coded genetic algorithm with forgetting

In cases where a fitness function for evaluating individuals can be represented by multilayer perceptrons (MLP), a local search technique is applicable to RGA, which is called a hybrid real-coded genetic algorithm with local search, HRGA/LS [6,7]. Accordingly, SLF is also applicable to HRGA/LS by introducing regularization.

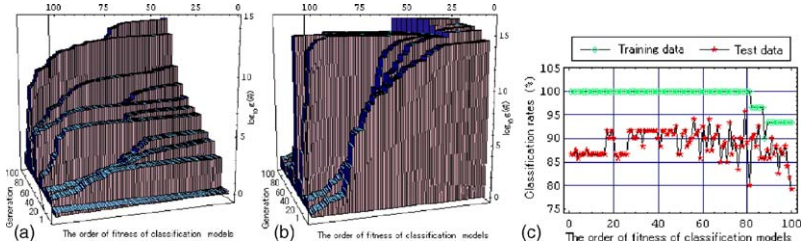

Fig. 2. Fitness of each classification model with HRGA/F at a search process and the correct classification rate for both training data and test data. (a) A result with the criterion function, $E'_1(w)$; (b) a result with the criterion function, $E'_2(w)$; (c) the correct classification rate according to each classification model for both training data and test data at 100th generation.