Available online at www.sciencedirect.com

SCIENCE DIRECT®

Applied Soft Computing 6 (2005) 72–84

Applied Soft Computing

www.elsevier.com/locate/asoc

Improving recognition and generalization capability of back-propagation NN using a self-organized network inspired by immune algorithm (SONIA)

Muhammad R. Widyanto [a,*], Hajime Nobuhara [a], Kazuhiko Kawamoto [a], Kaoru Hirota [a], Benyamin Kusumoputro [b]

[a] Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, G3-49, 4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan

[b] Faculty of Computer Science, University of Indonesia, Depok Campus, West Java, Indonesia

Received 5 April 2003; received in revised form 14 September 2004; accepted 22 October 2004

Abstract

To improve recognition and generalization capability of back-propagation neural networks (BP-NN), a hidden layer self-organization inspired by immune algorithm called SONIA, is proposed. B cell construction mechanism of immune algorithm inspires a creation of hidden units having local data recognition ability that improves recognition capability. B cell mutation mechanism inspires a creation of hidden units having diverse data representation characteristics that improves generalization capability. Experiments on a sinusoidal benchmark problem show that the approximation error of the proposed network is 1/17 times lower than that of BP-NN. Experiments on real time–temperature-based food quality prediction data shows that the recognition capability is 18% improved comparing to that of BP-NN. The development of the world first time–temperature-based food quality prediction demonstrates the real applicability of the proposed method in the field of food industry.

Keywords: Back-propagation; Immune algorithm; Self-organization; Food quality prediction

1. Introduction

Single hidden layer back-propagation neural networks (BP-NN) [1] are the most widely used method among various learning methods for neural networks. The networks are applied both to the function approximation and pattern classification. The networks estimate relation between input and output of sample patterns by updating iteratively the strength of connections between neural units so that an error between the actual and desired outputs of the networks will be minimized.

Since response to an input is given globally as multiplication and summation of all the connection

* Corresponding author. Fax: +81 45 924 5676.
E-mail address: widyanto@hrt.dis.titech.ac.jp (M.R. Widyanto).

1568-4946/$ – see front matter
doi:10.1016/j.asoc.2004.10.008

strengths of hidden units, BP-NN are hard to efficiently find relation between input vector and pattern category. This makes BP-NN cannot achieve an optimal recognition capability. Moreover, the strengths of the connections from input layer to hidden layer are updated based on the error between the actual and desired outputs only. As a result, there is a possibility that training data will be memorized by the network, this is called overfitting problem. In addition, if training data are not uniformly distributed in input space, there will be a region in the input space, in which the training data are not available enough for the network to correctly approximate the function in this region. This also leads to a poor generalization capability.

A method to improve recognition as well as generalization capability of BP-NN as a hidden layer self-organization inspired by immune algorithm [2] is proposed. This method serves as a hidden unit creation in BP-NN. B cell construction mechanism of immune algorithm inspires a creation of hidden units having local data recognition ability that improves recognition capability. Moreover, these hidden units enable the networks to memorize the characteristics of training data, therefore avoiding overfitting problem. B cell mutation mechanism inspires a creation of hidden units having diverse data representation characteristics that improves generalization capability. Experiments on the sinusoidal benchmark problem and a time–temperature-based food quality prediction show that the proposed method improves recognition and generalization capability of BP-NN.

In Section 2, the limitations of BP-NN are mentioned. The self-organized network inspired by immune algorithm (SONIA) is proposed in Section 3. Experimental results on the sinusoidal problem are presented in Section 4.1. In Section 4.2, experimental results on the time–temperature-based food quality prediction are shown.

2. Discussions on BP-NN

In back-propagation neural networks, there are issues that make the networks cannot reach an optimal generalization and recognition capability. The issues are summarized in the following two subsections.

2.1. Generalization capability

A capability of BP-NN to respond properly to previously untrained data is called generalization. Generalization problem associated with function approximation problem where generalization means the ability to interpolate in a meaning way through the training data in input space. Let $(\boldsymbol{s}_n, \boldsymbol{u}_n)$ $(n = 1, \ldots, N, N \in \mathbb{N})$ be a given set of pairs of input and output to be tested. Generalization capability is measured by approximation error $\mathrm{AE} \in \mathbb{R}_+$ defined by:

$$\mathrm{AE} = \frac{1}{N}\sum_{n=1}^{N}(\boldsymbol{u}_n - f(\boldsymbol{s}_n))^2. \tag{1}$$

As an example to see how the generalization capability of BP-NN, consider a sinusoidal benchmark problem [3,4]

$$h(x) = 0.4\sin(2\pi x) + 0.5, \tag{2}$$

where $x \in [0, 1]$. As in [3,4], $M = 36$ independent input–output vectors $(\boldsymbol{s}_m, \boldsymbol{u}_m)$ $(m = 1, \ldots, M)$ are generated to train BP-NN, with $\boldsymbol{s}_m$ taken randomly in [0,1] and $\boldsymbol{u}_m = h(\boldsymbol{s}_m) + r_m$, where r_m is noise to the mth output with noise value randomly taken in [−0.05, 0.05]. The network is tested using uniformly distributed data without noise $(\boldsymbol{s}_n, \boldsymbol{u}_n)$ $(n = 1, \ldots, N)$. Variable $\boldsymbol{s}_n$ is taken sequentially with difference 0.05 in [0, 1], resulting 21 input vectors $N = 21$ and $\boldsymbol{u}_n = h(\boldsymbol{s}_n)$.

The problem is simulated using BP-NN in Matlab neural networks toolbox [5]. The number of hidden units and other parameters are adjusted therefore the networks produce the best result. The number of hidden units is 13, this configuration was also used in [3,4]. The networks are trained until the mean squared error (M.S.E.) of training was constant where the training arrived at 4680 iterations with M.S.E. equals to 0.0018. The approximation error resulted is 0.01994. Fig. 1 shows the training data and the approximation result, it can be seen that BP-NN cannot achieve an optimal generalization performance.

In BP-NN, the strengths of the connection from input layer to hidden layer are updated based on the error between output of training data and output of the network. There is a possibility that training data will be memorized by the network, this is called overfitting problem. In addition, if training data are not uniformly distributed in input space (as shown by Fig. 1), there