REALIZATION OF GENERALIZED RBF NETWORK

NungKion, Lee
Faculty of Cognitive Science and Human Development, UNIMAS
94300 Kota Samarahan, Sarawak, Malaysia
+60 82679276
nklee@fcs.unimas.my

DianHui, Wang
Dept. of Computer Science and Eng.
La Trobe University
3080 VIC
+61 39479 3034
dhwang@cs.latrobe.edu.au

ABSTRACT
Neural classifiers have been widely used in many application areas. This paper describes generalized neural classifier based on the radial basis function network. The contributions of this work are: i) improvement on the standard radial basis function network architecture, ii) proposed a new cost function for classification, iii) hidden units feature subset selection algorithm, and iv) optimizing the neural classifier using the genetic algorithm with a new cost function. Comparative studies on the proposed neural classifier on protein classification problem are given.

Keywords
Radial basis function, Neural Network, Classification, Proteins Classification

1. INTRODUCTION
Classification is one of the areas in data mining. After training or learning, a classification system assigns an unknown test data to a known class or group in the problem domain. Different types of classifier or classification system can be used for classification. The most commonly used are the decision tree [1], neuro-fuzzy system [2], K-Nearest neighborhood [3], and neural networks [4]. Each of the classifier has its own strengths and weaknesses in some problem domain. For example, the decision tree is not suitable for noisy data whereas the neural networks are more tolerant to noise. This paper describes a neural network classifier applied to the protein sequences domain.

Neural networks have been used in many application areas, such as industrial process modeling and control, pattern recognition, financial forecasting, and bioinformatics. The multilayer-perceptron (MLP) [4] is the most popular neural architecture because it is easy to set up, achieve reasonable results and its universal approximation property [5]. Despite being widely used, the MLP network has a few limitations. The number of hidden units and the number of layers for MLP network are very subjective decisions that required trial and error. It is also rather hard to include prior knowledge into the network architecture although attempts have been made. The standard MLP network has many parameters (i.e. connection weights) to update during training, which resulted in long training time and local minima.

An alternative to MLP network is the Radial Basis Function (RBF) network. The RBF network is becoming more popular by looking at the number of researches, publications and applications that increased dramatically in recent years. The increasing popularity of this network architecture is due to its several interesting properties. Compared to the MLP network, the RBF network has only one layer of connection weights. This results in shorter training time and better generalization. Like the MLP network, the RBF network also posses the universal approximation property [6]. The standard RBF network consists of three layers; although more than three layer architecture has been considered. The main difficulty in using the RBF network is in determining the number of hidden units and their parameters. There have been many method proposed for this purpose, such as using subset of training data [7], clustering [8], mixture models and orthogonal least square [6]. In comparison to the MLP network, there are mixed results on which architecture can achieve better performance. For example in [9, 10], the RBF network can achieved good results, whereas in [11, 12], the MLP classifier is better.

The neural networks training required a cost function to update the connection weight and other parameters. The Mean Squared Error (MSE) [6] cost function is the most commonly used. This cost function calculates the error between the desired outputs and actual outputs and use them for updating the network parameters. By carefully studying the MSE cost function, it is noticed that this cost function has little relationship to the classification. Instead, it is more suitable for function approximation because its desired outputs are known. In classification, the desired outputs have to be artificially created, for example by using the 1-of-p coding. The assumption that this output vectors is optimal (e.g. standard one) might not be true in many problems.

This paper describes a generalized RBF (GRBF) neural network architecture for classification. This work is an extension on [13] earlier work with several improvements and modifications. The motivations of the GRBF neural network architecture are to improve the interpretability of the architecture and overcome some of the limitations of the standard RBF network. The proposed GRBF network is optimized using the genetic algorithm using a new objective function. This objective function is intended to achieve higher classification rate and reduce mis-classification rate. A modified feature subset selection algorithm is also proposed to reduce the number of input dimensions in the hidden units.

This paper is organized as follows. Section 2 review some related works. Section 3 describes the GRBF model, Section 4 gives the GRBF objective function, Section 5 describes the implementations, Section 6 gives some experimental results and the last section will be the conclusions.

2. RELATED WORK
Data driven initialization are commonly used to determine the number of hidden units, its cluster center and shapes of RBF network. The covariance matrix of a cluster determines the cluster shape (could be sphere, or ellipsoidal[+]) and orientation. Clusters center could be found by using any of the clustering techniques such as k-means [8], fuzzy-clustering [14], Expectation-Maximization [15], or agglomerate [8]. Beside RBF network, the

[+] Other cluster shapes are not considered in this paper.