18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Amsterdam 1996
7.4.1: Clinical Decision Making

IMPROVING THE CLASSIFICATION ACCURACY OF CARDIAC PATIENTS

J. Goddard , I. Lopez*, Hugo.L. Rufiner**
UAMI, Dept Electrical Eng., Av. Michoacán y La Purisima S/N, Col. Vicentina, CP (09340), Mexico D.F., México, jgc@xanum.uam.mx, *Centro de Cómputo Universitario, Universidad Autonoma Chapingo, Chapingo, Mexico, ilopez@mail.internet.com.mx, **Universidad Nacional de Entre Rios, Argentina, routa 10, Km 11, Oro Verde, Entre Rios, Argentina,rufi@xanum.uam.mx

Abstract

In medicine, the cost of erroneously classifying an ill subject as healthy could have disastrous consequences. Many classification techniques however, do not try to improve the classification on one of the classes involved. In this paper different strategies are proposed to solve this problem using Artificial Neural Networks (NN), and they are applied to a Heart disease dataset. An important reason for using NN is that the methods we propose can be directly incorporated into the learning process. To achieve this, two forms of NN training are used. One involves a slight modification of the usual backpropagation algorithm (BP), and in the other, the training is achieved through an Evolution Program (EP). The latter is applied to take advantage of their ability to handle more complicated fitness functions.

Keywords: Artificial Neural Networks, Classification, Evolution Programs, Fitness Functions, Genetic Algorithms, Heart Disease.

Introduction

In medical classification problems, it is often the case that a more precise classification is required on one of the classes involved, even at the expense of a poorer classification on another. This could arise because the consequences of misclassifying one class, such as an illness, could be far more serious, leading to a subject not receiving adequate and timely treatment. Often the results reported using a classification technique tend to present the overall classification correctness rather than the individual classification correctness for each class. This can be deceptive. If the number of subjects that fall into one class is greater than those in another, an apparently good classifier might perform poorly on the smaller class.

In this work, several different strategies designed to solve this problem using NN are proposed and applied to a heart disease dataset, originally compiled at the Cleveland Clinic Foundation and supplied by Robert Detrano, M.D.,Ph.D. of the V.A. Medical Center. This dataset was previously used in the "StatLog" project under the ESPRIT programme of the European Community [1], which compared diverse classification algorithms on a variety of datasets. In its present form, the dataset contains 270 cases of 13 attributes each, of which 120 show the presence of heart disease and 150 do not (the original dataset contained a larger set of examples, attributes and classes). The 13 attributes consist of the age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise induced angina, ST depression induced by exercise, the slope of the peak exercise ST segment, number of major vessels colored by flouroscopy, and thal.

Methods

A fixed feedforward NN architecture was used throughout the trials. A major reason for choosing this classification technique was the fact that the strategies applied could be directly incorporated into the learning process, and not left to the final testing stage as is the case of most statistical procedures. The NN consisted of 13 input units, a hidden layer of 13 units, and a single output unit. Threshold units on the first and hidden layers were also included, and a standard sigmoidal activation function was applied. Training of the NN was conducted using the same 216 (80%) randomly selected cases. The remaining 54 cases were used for testing. The 216 cases included 89 presenting heart disease, and 31 for the test set. Preliminary results obtained showed an improvement on normalising the data and this was done throughout the trials.

The first strategy modified BP [2] using fixed "cost factors" to weigh the errors more heavily for those cases presenting heart disease. The objective was to obtain a better classification of this class. Cost factors of 2, 4, and 10 were used and the results are presented in table 1 of the next section (the factor of 1 corresponds to the normal application of BP). The number of cases presenting heart disease in the training set was also duplicated and the 305 cases were trained with BP to consider the effect on classification. This result is shown in the last row of table 1.

The second strategy consisted in changing the error function to the sum of the misclassified cases. A cost factor was again included to penalise the number of misclassified cases presenting heart disease. To avoid the difficulties of using BP with such functions, an EP was applied to minimize the error function. Genetic algorithms and EPs have been used successfully in NN training [3], and their versatility in handling a variety of difficult optimisation problems with non-continuous fitness functions has been shown [4]. An EP was chosen over a Genetic Algorithm because of the naturalness of coding the NN as a chromosome, and an apparently more efficient convergence rate. The EP simulations conducted were realised using SUGAL [5]. After considerable experimentation, the following combination of parameters were found to produce the best results. A population of 10 chromosomes was used, each chromosome consisting of 196 real numbers (the number of weights required for the NN architecture). The genes were initialised using a gaussian distribution with 0 mean and a variance of 1. Selection was accomplished using the roulette method. 2-point crossover was applied to pairs of selected chromosomes, and gaussian mutation of decreasing amplitude to all the genes. An elitist strategy maintained the best individual in the population. The EP was applied for 300 generations.

Results

Table 1 and 2 show the results obtained using a NN trained with the variations to BP and an EP, respectively, as explained in the Methods section.

Factor	Train (%Good)			Test (%Good)			Epochs
	Pres	Abs	Tot	Pres	Abs	Tot	
1	91.9	88.2	89.4	77.4	91.3	83.3	100
2	93.3	74.8	82.4	90.3	87.0	88.9	100
4	91.0	83.5	86.6	83.9	91.3	87.0	100
10	93.3	72.4	81.0	90.3	78.3	85.2	100
Dup	93.3	81.9	88.5	80.7	26.1	57.4	500

TABLE 1. Results of the NN trained with Backpropagation

Factor	Train (%Good)			Test (%Good)			Gens
	Pres	Abs	Tot	Pres	Abs	Tot	
1	78.7	96.1	88.9	67.4	95.7	79.6	300
2	89.9	83.5	86.1	77.4	87.0	81.5	300
4	95.5	74.8	83.3	87.1	78.3	83.3	300
10	98.9	64.6	78.8	93.6	56.5	77.8	300

TABLE 2. Results of the NN trained with an EP

The first four rows give the results for the different cost factors. The last row of table 1 presents the results obtained by duplicating the cases presenting heart disease in the training set. The figures represent the percentage of cases correctly classified according to their class (Pres means heart disease was present, Abs that it was absent). The total percentage correctly classified is given for the different trials in the Tot column.

Conclusions

We have proposed different alternatives to improve the classification results for a particular class by incorporating a cost factor into the learning process of a NN. As one can conclude from the results given in the tables, the NN training with the EP consistently improves the classification results of the desired class on increasing the cost factor. However, as can be expected this is at the expense of a decrease in the correct classification on the other class. A reasonable cost factor seems to lie within the 2-4 range. The NN training with BP produced uneven results, although improved classification was usually obtained. Duplicating the cases resulted in poor performance on the test set, due to the over-fitting of the model on the training set.

The generality of the NN training with the EP, and their ability to handle different fitness functions, opens up promising avenues for future work.

Acknowledgement: This work was partly conducted using the Sugal Genetic Algorithm package, written by Dr. Andrew Hunter at the University of Sunderland, England.

References

[1] Michie, D., Spiegelhalter, D.J., Taylor, C.C.(Eds),"Machine Learning, Neural and Statistical Classification", Ellis Horwood, 1994.

[2] Hassoun, M.H.,"Fundamentals of Artificial Neural Networks", The MIT Press, 1995.

[3] Montana, D.J.,"Neural Network Weight Selection Using Genetic Algorithms", in Intelligent Hybrid Systems, Wiley, 1995, Goonatilake and Khebbal (Eds).

[4] Michalewicz, Z.,"Genetic Algorithms+Data Structures = Evolution Programs", Springer-Verlag, 1992.

[5] SUGAL User Manual V2.1, http://osiris.sund.ac.uk/ahu/sugal/home.html