A New Neural Network Structure for Detection of Coronary Heart Disease

Z. Shen, M. Clarke, R. Jones and T. Alberti
Department of Electrical Engineering, Brunel University, Middlesex, UK

We investigate the application of neural networks for the detection of Coronary Heart Disease (CHD). We have used a Neural Network (NN) on data from a self-applied questionnaire to implement a decision system designed to seek out high risk individuals in a large population. A Multi-Layered Perceptron (MLP) was trained with risk factors to distinguish CHD. We also describe a modification to the architecture of the neural network in which an extra layer of neurons is added at the input. We present possible interpretations of the weights of these neurons, and show how they can be used as a selection criteria for which questions to use as inputs. The technique is compared against other statistical methods. We go on to demonstrate the system's capability for detecting both the symptomatic and asymptomatic patient.

Keywords: Backpropagation; Coronary heart disease; Diagnosis; Neural networks

1. Introduction

Despite a recent fall in CHD mortality, it remains the most common cause of death in the UK and other developed countries. Although the pathophysiology is better understood and many risk factors have been identified, its exact cause remains unclear. Medical and surgical care of established disease has improved prognosis, but people continue to be denied such treatment because of sudden death caused by CHD. The accepted strategy is, therefore, to prevent CHD by both reducing risk factors in society as a whole through health education, and recognising the at-risk individuals and reducing their level of risk by medical treatment and education. Difficulties with the strategy abound, and in particular, the yield of at-risk individuals is relatively small, with a significant rate of false positive findings.

We seek to develop an accurate method of recognising the at-risk individual in a large and apparently asymptomatic population. When trained on self-applied questionnaire data, neural networks recognise people with CHD risk. They form a relatively high-risk population in which further cardiac investigation will operate with greater accuracy. Our work is novel in using data obtained directly from the patient, rather than from the clinician's history and examination. These are used to validate results and not, as is often the case, as training inputs. Neural networks are attractive for this work – their robustness compensating for the noisy and incomplete data that are the answers returned in many self-applied questionnaires.

2. The Medical Data Collection and Input Selection

2.1. The Questionnaire

A questionnaire containing 300 questions about health status was completed by 102 patients aged 45 to 65 years, and involved in a long-term study of CHD in primary health care. Earlier in the study
they had undergone: exercise ECG testing, blood analysis and 72 hour ambulatory ECG monitoring [1], and were subject to continuous clinical monitoring. There were 81 questionnaires returned, and of these 16 were from patients with known or subsequently proven CHD. A further 16 questionnaires from people who had neither overt symptoms nor a risk of CHD were a control group deemed to have No Heart Disease (NHD), and these matched the disease group in age and sex.

2.2. Selection of Inputs by Statistical Calculation

We carried out pre-selection of the inputs to the neural network because of the limited data set available, and as not all of the questions were likely to enhance the diagnosis of CHD – the questionnaire was designed to assess health status in general rather than CHD alone. We have used a Root Mean Square Error (RMSE) metric (1) to determine the discriminatory power of each question:

\[
RMSE_i = \sqrt{\left( \sum_{n=1}^{N} (TP_n - FP_n)^2 \right)}
\]

where

- \( TP_n = \frac{T_{nCHD}}{T_{CHD}} \) is the proportion of CHD patients giving answer \( n \) to question \( i \);
- \( FP_n = \frac{T_{nNHD}}{T_{NHD}} \) is the proportion of NHD patients giving answer \( n \) to question \( i \);
- \( RMSE_i \) is the value of RMSE for question \( i \);
- \( N \) is the number of possible answers to question \( i \);
- \( T_{nCHD} \) is the number of CHD patients who gave answer \( n \) to question \( i \);
- \( T_{nNHD} \) is the number of patients without CHD who gave answer \( n \) to question \( i \);
- \( T_{CHD} (T_{NHD}) \) is the total number of patients with (without) CHD.

The questions were ranked, and those with the largest RMSE value were chosen as inputs for the network.

2.3. Selecting Inputs using a Double Input Layered MLP

The RMSE is performing a simple form of cluster analysis with no transformation of the data space, and is only able to rank questions on their individual merit; it does not consider the multi-factorial contribution of each factor. That is, any single input may appear unimportant on its own, but may be highly significant in combination with other factors. To overcome this limitation, we investigated modifying the structure of the MLP to include an extra hidden input layer (Fig. 1). With the second input layer, a set of weights between these two single connected input layers is obtained after training, and can be used to determine the contribution of each input. Those at the top of the rank order of the weights have the greatest contribution, and are therefore chosen as inputs.

Figure 2 shows the relationship between the rank order of the questions produced by the two methods, RMSE and weights between the double input layers, for the 24 questions in the questionnaire related to stress. The rank order for both methods is generally the same (a straight line if exactly the same), but there are notable differences. However, it is the questions selected according to double input layer weights that gives the best performance (Table 1) and this is discussed in more detail in Section 4.2.

3. The Double Input Layered MLP

3.1. The Structure

Figure 1 shows the structure of the double input layered MLP. It consists of a traditional three layered fully connected feedforward network, and an additional input layer, the number of nodes in this layer being the same as the number of inputs. During training the extra input layer has its weights adapted according to the error term \( \delta_i \) from the layer immediately above it. During testing this layer only multiplies its input by a constant, the weight. However, we have separated an input weight which can be used to determine the relative significance of each input.

The principle of error backpropagation and weight adaptation are the same as the traditional backpropagation learning algorithm [2]. We only give the equations for calculating the values for the additional layer

\[
out^p = f_\theta(w_{i^p}input^p)
\]

\[
\delta_i^p = \sum_h \delta_h^p w_{ih} f_\theta(out^p)
\]

\[
\Delta w_{i new} = \beta \sum_p \delta_i^p out^p + \alpha \Delta w_{i old},
\]

\[w_{i new} = w_{i old} + \Delta w_{i new}\]

The letter \( p \) indicates the \( p \)th training pattern,