

A FRAMEWORK FOR MULTI-BACKPROPAGATION

WAN HUSSAIN WAN ISHAK
FADZILAH SIRAJ
ABU TALIB OTHMAN
*Faculty of Information Technology
Universiti Utara Malaysia*

ABSTRACT

Backpropagation algorithm is one of the most popular learning algorithms in the Neural Network. It has been successfully implemented in many applications. However, training Neural Networks involve a large amount of data. Therefore, training the network is time consuming as each training session requires several epochs, which usually takes several seconds or even minutes. This paper proposes a multi-backpropagation approach to minimize the complexity of the network. The approach does not require an alteration of the algorithm. Instead, the large network is split into several smaller networks. An integrating network is then constructed to integrate the output from the smaller networks.

Key Words: *Neural Network, Backpropagation Network, Multi Backpropagation Network.*

ABSTRAK

Algoritma Rambatan Balik merupakan salah satu algoritma pembelajaran yang popular dalam Rangkaian Neural. Ianya telah diimplementasikan dengan jayanya dalam pelbagai aplikasi. Walau bagaimanapun, pembelajaran dalam Rangkaian Neural melibatkan jumlah data yang banyak. Oleh itu, proses latihan rangkaian memerlukan masa yang lama. Hal ini kerana setiap sesi latihan melibatkan beberapa pengulangan yang mengambil masa beberapa saat atau minit. Kertas kerja ini mencadangkan pendekatan berbilang rangkaian bagi meminimumkan kompleksiti rangkaian. Pendekatan ini tidak melibatkan pengubahsuaian terhadap algoritma. Sebaliknya, rangkaian yang besar dipecahkan kepada beberapa rangkaian yang kecil. Rangkaian

surface. Therefore to obtain convergence, the learning rate must be slowly reduced. In addition, study shows the existence of a relationship between gain, learning rate and weights in backpropagation networks (Thimm *et al.*, 1996). This is followed by the implications of this relationship for variations of the backpropagation algorithm.

However, a large volume of data is involved in training the network, so NN can be too complex and difficult to train. Therefore, this paper presents a multi-network approach to minimize the NN complexity. Using this approach, more data or rules could be inserted into the system without affecting its performance.

MULTI-BACKPROPAGATION FRAMEWORK

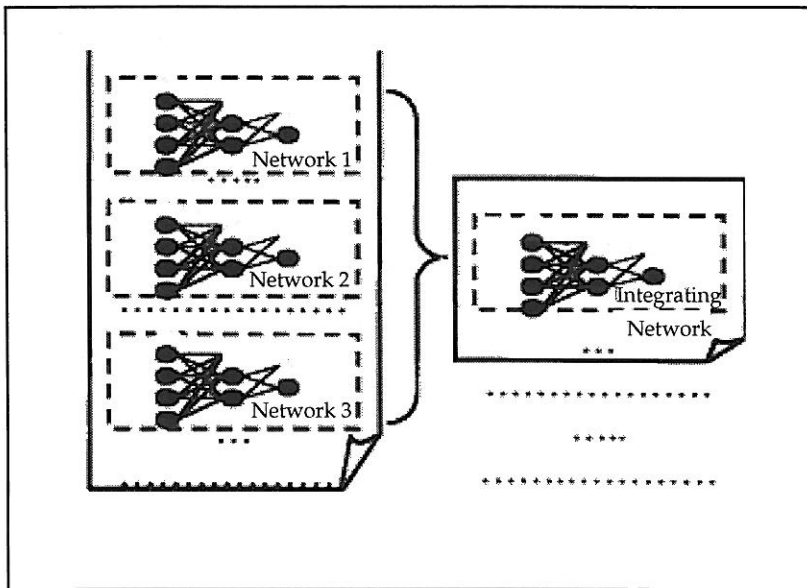
Backpropagation network is able to deal with various types of data, and model a complex decision system. Backpropagation network with hidden layer (or so called multi layer network) is able to process and model more complex problems. However, some problem domains might involve a large amount of data. Backpropagation network with four input units and two hidden units, for example, require several epochs, which create a complex model. More input units or hidden units could increase the complexity of the model and increase its computational complexity. In other words, an addition to the input unit or hidden unit could increase the model complexity and increase training time. This is because a larger network is more difficult to train. Like human learning, a complex problem requires certain period of time to establish learning.

In Figure 1 we illustrate the problem of multiple logical operation **AND**, **OR** and **XOR** that is $(A \text{ AND } B) \text{ AND } (C \text{ OR } D) \text{ OR } (E \text{ XOR } F)$ as the theoretical framework of the multi-network approach. In NN this problem is presented as a set of inputs into the network. The relation between each input is considered to be understood by the network. As we have six inputs, the total combination would be 64. Hence, training the network to learn all 64 data sets is time consuming where for each epoch the nets have to learn 64 different patterns. Each pattern is fed into the network one at a time and its error information term is calculated.

Basically this technique combines several operations and some of the operations are repeated. For example, $(A \text{ AND } B)$ and $(A \text{ AND } B) \text{ AND } (C \text{ OR } D)$ are an **AND** problem. Solving both problems require

Minimizing the complexity means reducing the complexity of each pattern by normalizing its attributes. Normalization referred to in this study is the same as that applied in relational databases where attributes are grouped into several categories to minimize the relationship between attributes. This technique could reduce the redundancy of data. Originally, the idea of multi backpropagation network is similar to the concept of bottom-up hierarchical neural network (see for example Ohno-Machado, 1996). Several specialized networks are constructed to represent a certain component of the problem and another network integrates the outputs to produce the final result (Figure 2).

Figure 2
Framework for Multi-Backpropagation Representation



EXPERIMENT AND DISCUSSION

This study uses the Myocardial Infarction problem as the research domain. The Myocardial Infarction problem consists of 26 variables. Each variable is represented in Boolean, either true or false. In this study, the variables are divided into five different groups, they are **COMPLICATIONS**, **ECG**, **INVESTIGATION**, **MEDICATION** and **RISK FACTORS** group (Figure 3).

Figure 4
Integrating Networks

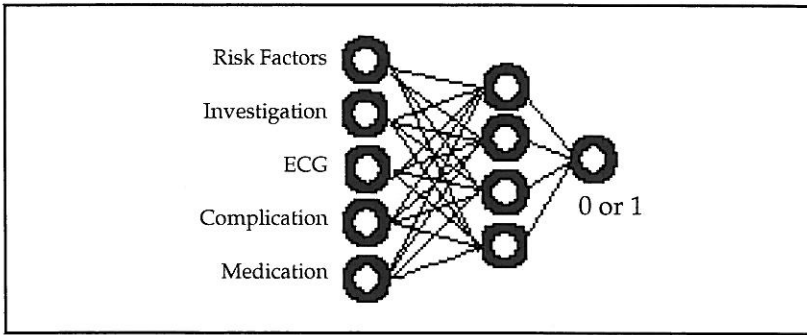


Figure 5
Predicting the Presence of Myocardial Infarction

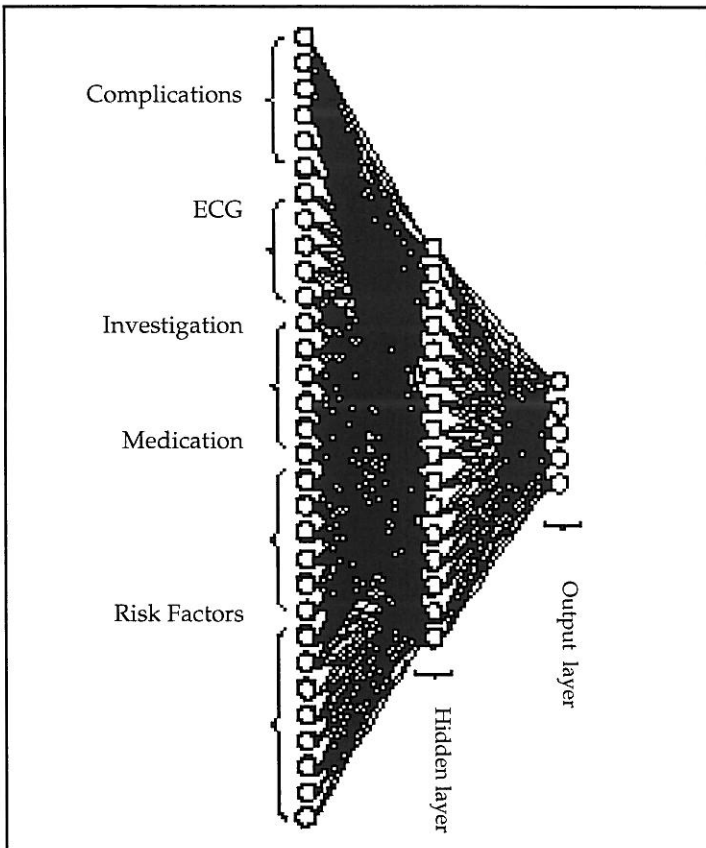


Table 2
Results Average

Network	Time (Ms)	Epochs	Results
Risk Factor	281	2	100
Medication	197	5	100
Investigation	32	1	100
ECG	440	34	100
Complication	83	3	100
Integrating network	22	1	100
Average	175.833	7.66667	100
Total	1055	46	600

CONCLUSION

The original data sets are too large and too complex to learn. Training the network using the single network approach does not cover all data sets. Even though the network generalizes approximately in two 115,421 milliseconds and with 40 epochs, it only represents a small portion of the whole data sets. Other data is not involved in training and the network may not recognize some of these patterns. In addition, the single network approach is estimated to take approximately 1,037,472,836 milliseconds and 359,544 epochs to generalize.

In the multi-network approach, the large network is divided into several smaller networks. Each network is trained separately. Another network called the integrated network was constructed to compile smaller networks into one network. Although many networks had to be constructed and trained separately, the multi-network approach has reduced the complexity of the network with large data sets and overcome the limitation of a single network approach. This is because the networks represent all possible combinations of data and train them respectively. In other words, in the multi-network approach, all data sets are used in training. The knowledge produced by the network can be applied for all possible data sets. The experiment reveals that the multi-network approach takes an average 175.833 milliseconds to complete the learning. In total, 1055 milliseconds and 46 epochs were taken for all networks to generalize.