RULE EXTRACTION FROM MULTI-LAYER PERCEPTRON NEURAL NETWORK USING DECISION TREE FOR CURRENCY EXCHANGE RATES FORECASTING

SOLEH ARDIANSYAH

The thesis submitted in fulfilment of the requirements
for the award of the degree of Masters of Science (Computer)

Faculty of Computer Systems & Software Engineering
UNIVERSITY MALAYSIA PAHANG

JUNE 2015

ABSTRACT

Neural network can be used in acquiring hidden knowledge in datasets. However, knowledge acquired by neural network was presented in its topology, the weights on the connections and by the activation functions of the hidden and output nodes. These representations are not easily understandable since neural networks act as a black box. The black box problem can be solved by extracting rule from trained neural network. Thus, the aim of this study was to extract valuable information (rule) from trained multi-layer perceptron (MLP) neural networks using decision tree. The main process in extracting rules from MLP using decision tree for currency exchange rate forecasting can be divided into two stages. In the first stage, the MLP network was built based on the parameter that was defined in the previous chapter. We also perform training and testing process experimentally and then the performance was evaluated in order to obtain the network with the best performance. The MLP network which provides the best prediction performance will be extracted by decision tree in the second stage by mapping input-output of the network directly. The forecasting result have shown that MLP network of EUR/USD produced a significant results compared to MLP network of GBP/USD and USD/JPY in term of MSE, RMSE, MAPE, and DS. It is quite evident that as the number of hidden neurons increases, MSE and MAPE decrease. In addition, the number of iterations for each model continues to increase along with the increasing number of hidden neurons. The results on decision tree induction show that C4.5 algorithm induction produced a significant result in term of accuracy 84.07% - 86.34%, precision and recall 93.17% and 81.97% respectively. This study has shown how rule can be extracted from MLP network by decision tree without making any assumptions about the networks activations function or having initial knowledge about the problem domain. The extracted rule can be used to explain the process of the neural network systems and also can be used in other systems like expert systems.

ABSTRAK

Rangkaian neural boleh digunakan dalam memperoleh pengetahuan yang tersembunyi di dalam dataset. Walau bagaimanapun, pengetahuan yang diperoleh oleh rangkaian neural dibentangkan dalam topologi, berat pada sambungan dan fungsi pengaktifan yang tersembunyi, dan nod output. Perwakilan ini tidak mudah difahami kerana rangkaian neural bertindak sebagai kotak hitam. Masalah kotak hitam boleh diselesaikan dengan mengeluarkan peraturan dari rangkaian neural yang terlatih. Oleh itu, tujuan kajian ini adalah untuk mendapatkan maklumat berharga (peraturan) dari terlatih pelbagai lapisan perceptron rangkaian neural menggunakan pokok keputusan. Proses utama boleh dibahagikan kepada dua peringkat untuk mengeluarkan peraturan dari pelbagai lapisan perceptron (MLP) menggunakan pokok keputusan untuk ramalan kadar pertukaran mata wang. Dalam peringkat pertama, rangkaian MLP ini dibina berdasarkan parameter yang ditakrifkan dalam bab sebelumnya. Kami juga melaksanakan latihan dan proses ujian uji kaji dan kemudian prestasi akan dinilai untuk mendapatkan rangkaian dengan prestasi yang terbaik. Rangkaian MLP yang menyediakan prestasi ramalan yang terbaik akan diambil melalui pokok keputusan di peringkat seterusnya dengan pemetaan input-output rangkaian secara langsung. Hasil ramalan menunjukkan bahawa rangkaian MLP sebanyak EUR / USD melakukan yang terbaik rangkaian MLP berbanding GBP / USD dan USD / JPY dari segi MSE, RMSE, MAPE, dan DS. Ia agak jelas bahawa apabila bilangan neuron tersembunyi meningkat, MSE dan MAPE berkurangan. Di samping itu, bilangan lelaran bagi setiap model terus meningkat seiring dengan peningkatan bilangan neuron tersembunyi. Keputusan pada induksi pokok keputusan menunjukkan bahawa induksi algoritma C4.5 menghasilkan hasil yang ketara dari segi ketepatan 84,07% - 86,34%, ketepatan dan ingat 93,17% dan 81,97% masing-masing. Kajian ini telah menunjukkan bagaimana peraturan boleh diekstrak dari rangkaian MLP dengan pokok keputusan tanpa membuat sebarang andaian mengenai fungsi rangkaian pengaktifan atau mempunyai pengetahuan awal tentang domain masalah. Peraturan dikeluarkan boleh digunakan untuk menjelaskan proses sistem rangkaian neural dan juga boleh digunakan dalam sistem lain seperti sistem pakar.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

FOREX	Foreign exchange / currency exchange rates
ANN	Artificial Neural networks
EUR	European Euro
USD	United State Dollar
GBP	Great Britain Pound sterling
JPY	Japanese Yen
RSI	Relative strength index
STOCH	Stochastic oscillator
DT	Decision tree
MSE	Mean squared error
RMSE	Root mean squared error
MAE	Mean absolute error
MAPE	Mean absolute percentage error
DS	Directional symmetry
OHLC	Open, high, low, and close prices of currency exchange rates
MLP	Multi-layer perceptron

CHAPTER 1

INTRODUCTION

1.1 RESEARCH BACKGROUND

Artificial neural networks (ANN) are widely used in many applications such as a prediction technique for its nonlinear structure which is able to detect the relationship among numbers of different variables and even with the most complex problems, and that is the reason why ANN becomes a very promising prediction technique. A large number of studies have been reported in the literature with reference that ANN forecasting models provided the better performance (Seliem, 2006; Yao & Tan, 2000; Eng et al., 2008; El-Bakry & Mastorakis, 2010; Zhang & Hu, 1998; Kayal, 2010; Guresen et al., 2011; Chang, 2011).

ANN are parallel adaptive networks of simple nonlinear computing element called neurons which are intended to model some of the functionality of the human nervous systems and attempt to partially capture some of its computational strength (Kabeer & Narayanan, 2007; Kohonen, 1988). In the other word, ANN mimic the process which found in biological neural networks (human brain) and learning from a given dataset.

However, there are several limitations of ANN. ANN act as a *black box* (Tzeng & Ma, 2005; Sjoberg et al., 1995; Chandra et al., 2007) which information stored in a set of connections and weights that provide no insights as how a task is performed. ANN does not provide a solution that easily to understand since the relationship between input-output is not clear. The parameter selection of ANN is tedious and often done in a trial-error process even when ANN performs very similar tasks. The appropriate choice of network parameters can be varied (Setiono, 2013).

On the other hand, Decision Tree (DT) is able to generate and represent both classifiers and regression models in tree forms, class labels represented by leaves and conjunctions of features represented by branches that connected leaves to another leaves. Its learning algorithm is easy to understand in contrast to ANN's *black box*. DT provides a solution that always convergent to classify the training set of the data correctly (King & Rughooputh, 2002). Decision tree also can be easily converted to *if-then* rules. Moreover, rules are easy to understand with a natural language and they are also useful to explore the data by finding the hidden relationship between the numbers of input variables to a target variable. Since DT combines data exploration and modelling, DT is very good as a first step in the modelling process and even when used as a final step of the other technique (Hand et al., 2001). The best way of extracting the knowledge is by extracting if-then rules that is presented in neural networks. (Boz, 1995; Kamruzzaman & Islam, 2010).

There are many problems in real-life which future events can be predicted by the past history. For example, predicting the behaviour of currency exchange rates. A currency exchange rate (Forex) becomes a very profitable market with a daily transaction of more than US$ 3.0 trillion. This situation attracts many people to trade Forex. However, many inexperienced traders and new traders that do not have enough information about trading knowledge they are keep trying their luck in Forex trading. This behaviour may cause some losses risk. Many traders have concluded that the opportunities outweigh the risk. Therefore, forecasting of Forex has become a challenge for many years and raised many theories and experiments. Over the years, linear techniques have been used by researchers and analysts since they are very simple and easy to apply. However, this linear indicator has always worked well on a linear movement but stops helplessly when dealing with nonlinear behaviour of the market (Seliem, 2006).

In order to understand how ANN accomplishes the forecasting task and what is the information stored in trained ANN, this study aims to extract valuable information in the form of *if-then* rules from trained multi-layer perceptron neural networks by using decision tree. First, neural network will perform forecasting task of foreign exchange datasets. This is followed by an identification of the suitable architecture of ANN in order to perform the forecasting task. The neural network models which perform the best will be extracted using decision tree by mapping input output of the network directly. The extracted rule described the knowledge acquired by ANN while learning from given datasets.

1.2 PROBLEM STATEMENT

While neural network attain higher predictive accuracy than the other method or human experts. Generally, it is difficult to understand how ANN provides a particular conclusion due to its architecture complexity. ANN acts as a *black box* which the information stored in a set of weights and connections that provide no clear vision as how a task is performed. Furthermore, ANN does not provide a solution that is easy to understand since the relationship between input-output is not clear. The parameter selection of ANN is tedious and often done in a trial-error process even when ANN performs very similar tasks and the appropriate choice of network parameters can be varied. The black box problem can be solved by extracting the rule from neural network. Therefore, it is desirable to have a set of rules to explain how ANN solves a problem based on given datasets. General statement of problem in this study can be formulated as follows:

- How to extract rule from trained multi-layer perceptron neural networks by using decision tree?
- How to build multi-layer perceptron neural network in order to train the network for currency exchange rates forecasting?

1.3 RESEARCH OBJECTIVES

As described previously, neural networks become very promising technique since it can provide better performance instead of its black box limitation. In order to understand the stored knowledge inside multi-layer perceptron neural network, this research embarks on the following objectives:

- To extract valuable information in the form of *if-then* rules from multi-layer perceptron neural networks by using decision tree.
- To develop multi-layer perceptron neural networks forecasting model and train the network using currency exchange rate dataset.
- To evaluate the performance of multi-layer perceptron neural networks and determine the prediction accuracy of decision tree result.

1.4 RESEARCH SCOPE

The scopes of this research are as follows:

i. The design of multi-layer perceptron neural networks forecasting model is based on experiment on its architecture.

ii. The architecture of multi-layer perceptron neural network with the best result will be used to extract the rule.

iii. Decision tree is used for rule extraction process from multi-layer perceptron neural networks by mapping input-output of the network directly.

iv. This study using Levenberg Marquardt algorithm to perform forecasting task and decision tree C4.5 algorithm to extract the multi-layer perceptron neural networks.

v. This study focused on forecasting task by neural networks and knowledge extraction of trained neural network by decision tree for currency exchange dataset.

1.5 THESIS CONTRIBUTION

The work carried out in this study aims to produce a key contribution in combining the artificial neural networks and decision tree for rule extraction. The focus is to explore the influence of the neural networks architecture in currency exchange rates forecasting and extracted the rule that stored in multi-layer perceptron neural networks. Furthermore, a contribution can be made to the literature on the comparative benefits of the models from knowledge gain through the empirical studies. As far as we know, there is no other documented work which extract rule from multi-layer perceptron neural network by using decision tree for currency exchange rates forecasting as presented in this thesis.

1.6 THESIS OUTLINE

This thesis is divided into five chapters and organized as follows:

Chapter 2 presents a literature review related to neural network extraction process and overviewed of the importance of rule extraction. This chapter also provides a briefly preview of algorithm to construct artificial neural network model, optimization and step by step on designing ANN, decision tree algorithm and forecasting currency exchange rates.

Chapter 3 the research methodology, is devoted to the brief description of the techniques that used to build the multi-layer perceptron neural networks model and how to enhance these models for the best performance. This chapter also summarise the research designs of the models, data collection, as well as data analysis and pre-processing, and the performance measurement.

Chapter 4 explains several experiments have been conducted in order to test the performance the multi-layer perceptron as forecasting model and rule extraction process from multi-layer perceptron neural network with the best result. This chapter focuses on analysing the result from the experiments and the performances of the multi-layer perceptron and the rule extraction result are evaluated.

Chapter 5 is the conclusion of the work undertaken. The conclusion explains the answer to the question in the problem statement, the objective, research methodology, the result and discussion of the thesis. The study end with concluding remark and present the future work of this thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter reviews the existing literature that related to the artificial neural networks, the architecture of multi-layer perceptron, design, learning paradigm, training algorithm, optimization, and advantages and disadvantages of neural network.

Next, discussion of research studies of rule extraction from neural networks. The reviews on relevant studies that are presented indicating a research gap into rule extraction from multi-layer perceptron neural network. Then, a brief discussion of the importance of rule extraction, decision tree algorithm, the theory and relevant research studies of currency exchange rates are reviewed. The chapter ends with summary of the literature review.

2.2 NEURAL NETWORKS

The human brain is formidably complex structure, comprising a network of roughly a hundred billion neurons. Each neuron has about ten thousand synapses on an average, and therefore, the total number of connections is phenomenal one thousand trillion or 10^{15} (Kumar, 2004). A cubic millimetre of brain contain on an average nearly a billion synapses. The information-processing cells of the brain are the neurons. Neurons come in various shape and size: some are small as a red blood cell; others half as large as the entire organism. As shown in Figure 2.1, each neuron has a *soma* or cell body which contains the cell's nucleus and other vital components called *organelles* which perform specialized tasks (Kumar, 2004) and its main communication links:

- A set of *dendrites* which form a tree-like structure that spreads out from the cell. The neuron receives its input electrical signal along these.
- A single *axon* which is a tubular extension from cell soma that carries an electrical signal away from the soma to another neuron for processing.

The dendrites and axon together are sometimes called processes of the cell. The real neurons integrate hundreds or thousands of temporal signals through their dendrites. These signals modify their internal potential in a complex way that depends on the inhibitory or excitatory nature of the synapses at which the signal impinge on the cell.

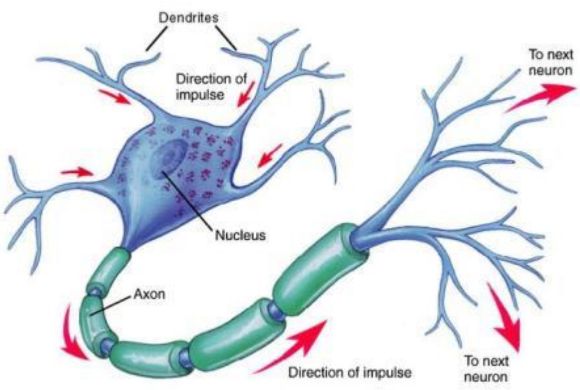

Figure 2.1: Basic diagram of a neuron, the unit of the nervous system. The axon is what the action potential is propagated along

Source: http://www.biocog.com/nervoussystem.htm

Artificial neural networks (ANN) are massively parallel adaptive networks of simple nonlinear computing element called neuron which is intended to abstract and model some of the functionality of the human nervous system in the attempt to partially capture some of its computational strengths (Kabeer & Narayanan, 2007; Kohonen, 1988). In other words, ANN mimic the process which found in biological neural networks and learning from a given datasets.

2.2.1 Architecture of Neural Network

The biological neural systems have inspired the creation of the computation process that has performance is identical with the human nervous system. As with the biological neural networks, mathematical model of neural network connecting a number of inputs and outputs from an adaptive system which is organized in layers of processing elements interconnected (Kumar, 2004). The basic structure of artificial neural networks is a neuron (see Figure 2.2).

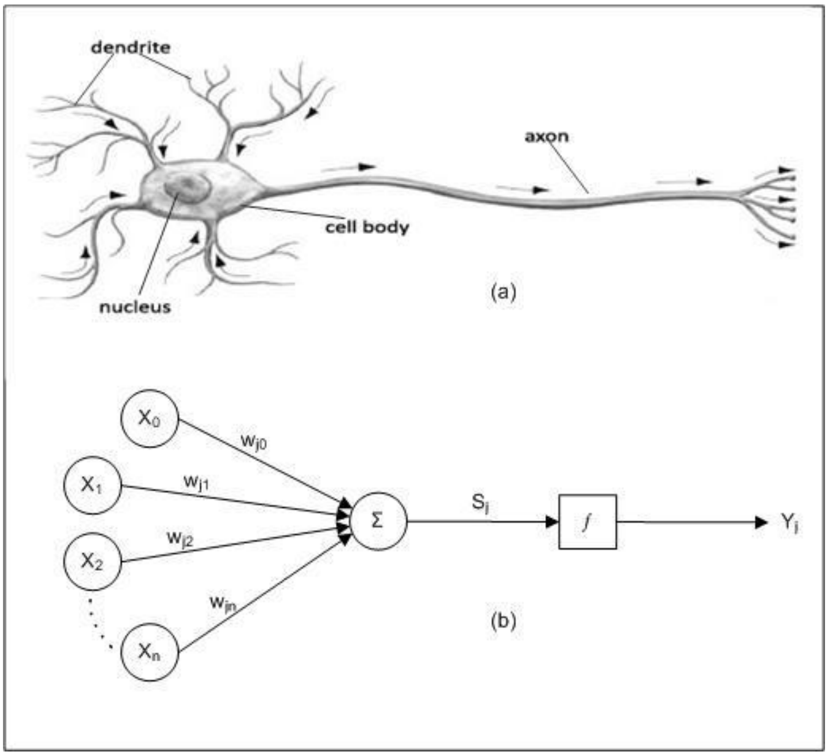

Figure 2.2: The analogy of biological cell and artificial neuron

As shown in Figure 2.2 (b), the neural model usually consists of:

i. Input (x) to receive the signal
ii. Connection weights (w_{ji}) to store the information
iii. Threshold (w_0) to set the bias value
iv. Processing element ($\sum$) and activation function (f)
v. Output (Y_j) present the results of information processing to the next cell

The neuron as shown in Figure 2.2 can be modelled by the following mathematical equations:

$$S_j(x) = \sum_{i=1}^{n} w_{ji}x_i + w_0x_0 \tag{2.1}$$

$$Y_j(x) = f\left(S_j(x)\right) \tag{2.2}$$

If $S_j > 0$, then $Y_j = 1$; if $S_j < 0$, then $Y_j = 0$; where Y_j is the output of the processing element, w_i is the input connection weight and x_i is the number of input neuron.

This can be of three layers: input, hidden, and output. Input neurons are designated to receive external stimuli that have been presented to the network. Outputs from the network are generated as a signal of output neuron. Hidden neurons compute the intermediate functions and their states are not accessible to the external environment. A layer is connected to the other weight through weight-labelled link. The output of the neuron is typically altered by a transfer function. The most common functions are binary threshold, linear threshold, and sigmoid function and they may differ from neuron to neuron within the network. Neural network is trained by using time series data in order to capture the nonlinear characteristic of the data (Yu et al., 2005).

ANN with a single layer has a limitation in pattern recognition. This limitation can be overcome by adding one or more hidden layers between the input and output layer. Although the use of more than one hidden layer has more benefits for some cases, however it takes a long time to perform the training process. So generally people try a network with one hidden layer first (Siang, 2012).

Back propagation trained the network in order to obtain a balance between the ability of the network to recognize the patterns on the training process as well as the network's ability to provide the correct response to the similar input pattern (but not identical) with the pattern that used during training process. Figure 2.3 shows the architecture of multi-layer perceptron neural network with *n* inputs (plus bias), a hidden layer that consists of *p* units (plus bias), and *m* output units. v_{ji} is the connection weight from the input unit x_i to the hidden layer unit z_j. v_{j0} is the weight that connecting the bias from input unit to a hidden layer unit z_j). w_{kj} is the connection weight from the hidden layer unit z_j to the output y_k (w_{k0} is the weight that connecting the bias from hidden layer to the output z_k).

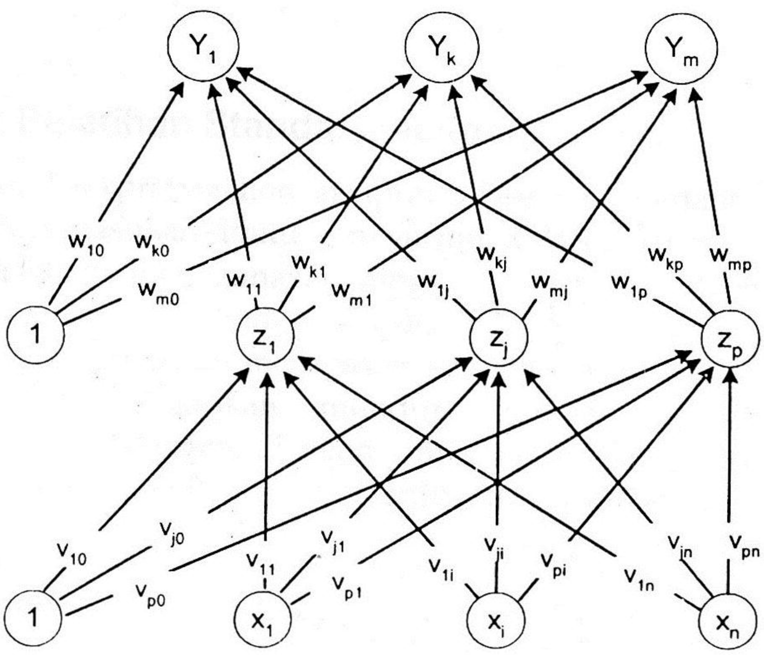

Figure 2.3: The architecture of multilayer perceptron neural network

Source: Siang (2012)

Activation function used in back propagation must meet several requirements, namely: continuous function, differentiable, and monotonic. The activation functions commonly used is the Log-sigmoid activation function with range (0, 1) (see Figure 2.4).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.3}$$

$$f'(x) = f(x)\big(1 - f(x)\big) \tag{2.4}$$

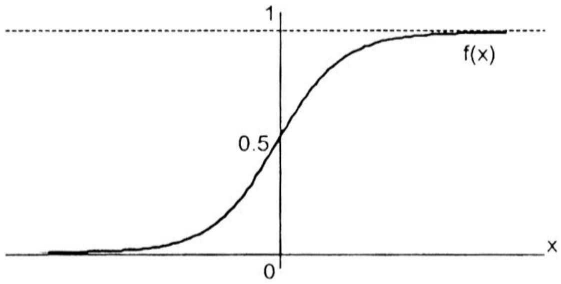

Figure 2.4: Log-sigmoid activation function

Source: Siang (2012)

Another function that commonly used is the Hyperbolic-Tangent Sigmoid function whose similar to the log-sigmoid function, but with the range [-1, 1].

$$f(x) = \frac{2}{1 + e^{-x}} - 1 \tag{2.5}$$

$$f'(x) = \frac{\big(1 + f(x)\big)\big(1 - f(x)\big)}{2} \tag{2.6}$$

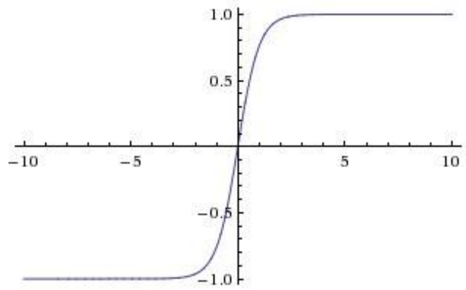

Figure 2.5: Hyperbolic tangent sigmoid activation function

Source: http://cs231n.github.io/neural-networks-1

The graph function is shown in Figure 2.5. Log-sigmoid function has a maximum value = 1. Then for the targets > 1, the inputs and outputs must be transformed first so all patterns have the same range as the activation function that used.

The alternative way is the sigmoid activation function only used in the layer which is not the output layer. The activation function that used in the output layer is identity function: $f(x) = x$.

2.2.2 Designing Neural Network as Forecasting Model

The common misconception about neural network is that they represent the kind of artificial intelligence which is not only capable of replacing the human brain, but which also possesses some magical power that takes inputs and find a way of making money by itself. This concept can be dangerous for trader. Therefore it's important to choose the right parameters in order to get the ANN forecasting model with the best performance (Vonko, 2006).

Nelson and Illingworth (1991) provided an overview of the eight steps on designing ANN as forecasting model. The eight steps on designing ANN as forecasting model are presented in Figure 2.6.

Eight steps on designing ANNs as forecasting model:

1. Variable Selection
2. Data collection
3. Data processing
4. Training, testing and validation set
5. Neural network paradigms :
   - Number of hidden layers
   - Number of hidden neurons
   - Number of output neurons
   - Transfer functions
6. Evaluation Criteria
7. Neural Network training
   - Number of training iteration
   - Learning rate and momentum
8. Implementation

Figure 2.6: Step by step on designing ANN

The first step is variable selection. In this step the researcher must determine the appropriate input variable that will be used. There are two kinds of data which is available in foreign exchange market, namely technical data and fundamental data. Technical data is a past price of forex or indicator that calculated from past price. Fundamental data is intrinsic value of forex including economy and industry condition. Technical data is commonly used by researcher as inputs since it's easy to obtain. Intermarket data can also be used as input variables since they have close relationship. This study used technical data as the input variables.

The next step is data collection and it is critical aspect of any type of research study. The inaccurate data collection can lead to invalid result of the study. The data collection, data pre-processing and data divisions will be describing briefly in subchapter 3.3.

The fifth step is neural network architecture. In this steps, the researcher must defined ANN's architectures including the number of hidden layer, number of hidden neuron, number of output neurons, transfer function, etc. the parameter of designing the architecture in this study are describe as follows:

i. Number of hidden layer; ANN with a single layer has a limitation in pattern recognition. This limitation can be overcome by adding one or more hidden layers between the input and output layer. Although the use of more than one hidden layer has more benefits for some cases, however it takes a long time to perform the training process. So generally people try a network with one hidden layer first. Panchal et al. (2011) outlined the behaviour analysis of multilayer perceptron; if the accuracy of the result is a critical factor for an application then the multiple hidden layers should be used.

ii. Number of hidden neuron; Deciding the number of neurons in the hidden layer is a very important part of ANN architecture since there is no exact number of the hidden neuron (Tsai & Wang, 2009). Using too few neurons in the hidden layer will result in under-fitting and over-fitting will occur when using too many neurons in the hidden layer. Therefore, this problem depends on experimental method to

CHAPTER 1

INTRODUCTION

1.1 RESEARCH BACKGROUND

Artificial neural networks (ANN) are widely used in many applications such as a prediction technique for its nonlinear structure which is able to detect the relationship among numbers of different variables and even with the most complex problems, and that is the reason why ANN becomes a very promising prediction technique. A large number of studies have been reported in the literature with reference that ANN forecasting models provided the better performance (Seliem, 2006; Yao & Tan, 2000; Eng et al., 2008; El-Bakry & Mastorakis, 2010; Zhang & Hu, 1998; Kayal, 2010; Guresen et al., 2011; Chang, 2011).

ANN are parallel adaptive networks of simple nonlinear computing element called neurons which are intended to model some of the functionality of the human nervous systems and attempt to partially capture some of its computational strength (Kabeer & Narayanan, 2007; Kohonen, 1988). In the other word, ANN mimic the process which found in biological neural networks (human brain) and learning from a given dataset.

However, there are several limitations of ANN. ANN act as a *black box* (Tzeng & Ma, 2005; Sjoberg et al., 1995; Chandra et al., 2007) which information stored in a set of connections and weights that provide no insights as how a task is performed. ANN does not provide a solution that easily to understand since the relationship between input-output is not clear. The parameter selection of ANN is tedious and often done in a trial-error process even when ANN performs very similar tasks. The appropriate choice of network parameters can be varied (Setiono, 2013).

On the other hand, Decision Tree (DT) is able to generate and represent both classifiers and regression models in tree forms, class labels represented by leaves and conjunctions of features represented by branches that connected leaves to another leaves. Its learning algorithm is easy to understand in contrast to ANN's *black box*. DT provides a solution that always convergent to classify the training set of the data correctly (King & Rughooputh, 2002). Decision tree also can be easily converted to *if-then* rules. Moreover, rules are easy to understand with a natural language and they are also useful to explore the data by finding the hidden relationship between the numbers of input variables to a target variable. Since DT combines data exploration and modelling, DT is very good as a first step in the modelling process and even when used as a final step of the other technique (Hand et al., 2001). The best way of extracting the knowledge is by extracting if-then rules that is presented in neural networks. (Boz, 1995; Kamruzzaman & Islam, 2010).

There are many problems in real-life which future events can be predicted by the past history. For example, predicting the behaviour of currency exchange rates. A currency exchange rate (Forex) becomes a very profitable market with a daily transaction of more than US$ 3.0 trillion. This situation attracts many people to trade Forex. However, many inexperienced traders and new traders that do not have enough information about trading knowledge they are keep trying their luck in Forex trading. This behaviour may cause some losses risk. Many traders have concluded that the opportunities outweigh the risk. Therefore, forecasting of Forex has become a challenge for many years and raised many theories and experiments. Over the years, linear techniques have been used by researchers and analysts since they are very simple and easy to apply. However, this linear indicator has always worked well on a linear movement but stops helplessly when dealing with nonlinear behaviour of the market (Seliem, 2006).

In order to understand how ANN accomplishes the forecasting task and what is the information stored in trained ANN, this study aims to extract valuable information in the form of *if-then* rules from trained multi-layer perceptron neural networks by using decision tree. First, neural network will perform forecasting task of foreign exchange datasets. This is followed by an identification of the suitable architecture of ANN in order to perform the forecasting task. The neural network models which perform the best will be extracted using decision tree by mapping input output of the network directly. The extracted rule described the knowledge acquired by ANN while learning from given datasets.

1.2 PROBLEM STATEMENT

While neural network attain higher predictive accuracy than the other method or human experts. Generally, it is difficult to understand how ANN provides a particular conclusion due to its architecture complexity. ANN acts as a *black box* which the information stored in a set of weights and connections that provide no clear vision as how a task is performed. Furthermore, ANN does not provide a solution that is easy to understand since the relationship between input-output is not clear. The parameter selection of ANN is tedious and often done in a trial-error process even when ANN performs very similar tasks and the appropriate choice of network parameters can be varied. The black box problem can be solved by extracting the rule from neural network. Therefore, it is desirable to have a set of rules to explain how ANN solves a problem based on given datasets. General statement of problem in this study can be formulated as follows:

- How to extract rule from trained multi-layer perceptron neural networks by using decision tree?
- How to build multi-layer perceptron neural network in order to train the network for currency exchange rates forecasting?

1.3 RESEARCH OBJECTIVES

As described previously, neural networks become very promising technique since it can provide better performance instead of its black box limitation. In order to understand the stored knowledge inside multi-layer perceptron neural network, this research embarks on the following objectives:

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

In this chapter, the research methodology used for each case studies is briefly discussed in the following sequence: the research framework, multi-layer perceptron neural network forecasting model, rule extraction from multi-layer perceptron neural network by using decision tree, data collection, data analysis and pre-processing, and performance measurement. The conclusions that can be made as a result of using this methodology are discussed at the end of the chapter.

3.2 RESEARCH FRAMEWORKS

In this research, there are several phases which are literature study, logical design, implementation, training and testing, and analysis. The result on every phase in this methodology will be divided into several steps that can be achieved in a suitable time frame. During the literature study phase, this study review on foreign exchange rate, current issue or enhancement of artificial neural network and decision tree, the combination model, data collection and analysis.

The framework of the forecasting model will be designed during the logical design phase. After that, we will build the multi-layer perceptron (MLP) neural network forecasting model based on the framework design. During the implementation phase, the algorithm will be applied into programs. All the developments will use appropriate programming techniques and development tools. Next, the model will be trained and tested using foreign exchange datasets until it can be well function. The result will be analysed and the report can be produced. If not, the logical design will be revised and follow with the implementation. The research framework is summarized in flowchart, see Figure 3.1.

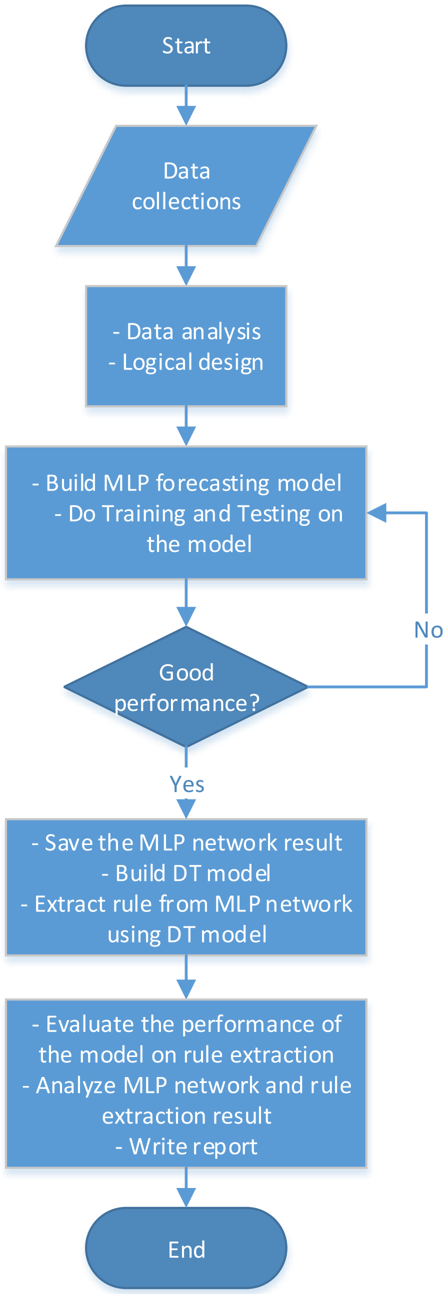

Figure 3.1: Research frameworks

3.3 DATASET DESCRIPTION

3.3.1 Data Collection

When collecting the data for chosen variable, we must consider the cost and availability of the data. The technical data is readily available from many vendors at a reasonable cost whereas fundamental data is more difficult to obtain. This study used technical data of foreign currency exchange rates which obtain from *Meta Trader* software. *Meta Trader* (latest version is MT5) is an electronic trading platform widely used by online retail foreign exchange speculative trader. It was developed by *Meta Quotes Software* and licensed to foreign exchange brokers who provide the software to their clients. The software consists of both a client and server component. The server component is run by the broker and the client software is provided to the broker's customers, which use it to see live streaming prices, charts and to place orders as well as manage their account.

Even the data provider have a reputation of providing high quality data, all the data must be checked for error by examining day to day changes, range, logical consistency, and missing observation. This study used daily time series data of foreign currency exchange rates of European Euro (EUR), British Pound sterling (GBP), and Japanese Yen (JPY) against US Dollars (USD). The goal of forecasting model is to predict the next day of the currency exchange rates. The period of the data that being used is from January 2000 to December 2012 daily exchange rates.

3.3.2 Data Pre-processing

Data pre-processing refers to analysing and transforming the input and output variables to minimize the noise, highlight important relationships, and detect the trend of the data. Since ANN uses pattern matching, the representation of the data is critical in designing a successful network. The input and the output variables for which the data was collected are rarely fed into the network in raw form. The raw data must be scaled into the transfer function ranges which usually between 0 to 1 or -1 to 1, starting from