Forecasting Malaysian Stock Price using Artificial Neural Networks (ANN)

Abdulrazak F. Shahatha Al-Mashhadani¹, Sanil S Hishan², Hapini Awang³, Kamal Ali Ahmed Alezabi⁴

¹² Azman Hashim International Business School, Universiti Teknologi Malaysia
³ School of Information Technology, UCSI University, Kuala Lumpur, Malaysia
⁴ School of Computing, Universiti Utara Malaysia, Sintok, Malaysia

Emails: razak.shahatha@utm.my¹, hishanssanil@gmail.com, hapini.awang@uum.edu.my³, kamal@ucsiuniversity.edu.my⁴

Abstract— Predicting a stock price is a very difficult task because it is complex and involves many factors. This has led to drop in the investment level in the Malaysian stock market. It is difficult to predict the stock market because its environments are unstable and dynamic. Recently, the demand for neural network in the business arena is on the increase. It is need to analyze vast data in order to search for information and knowledge that do not exist by using traditional methods. This included stock market prediction that is a very significant research in business area. In regard to Bursa Malaysia, Artificial Neural Network (ANNs). ANNs was only used to predict main index, i.e. Kuala Lumpur Composite Index (KLCI), but no attempt to predict share price and in particular banking sector. Since ANN has potential to predict non-linear behavior, this research attempts the use of ANNs to predict banking sector stock price in FTSE Bursa Saham Malaysia Kuala Lumpur Composite Index (FBM KLCI). One of the interesting topics of stock-market research is stock market prediction. Precise stock forecasting becomes the greatest challenge in the investment industry because stock data distribution changes over time. This paper investigates the use of ANN to predict Malaysian stock price, in particular Maybank Berhad stock price. The feedforward neural back-propagation network with Training Function Gradient Decent Training Algorithm is used in this study. The outcome of selected stocks, namely Maybank, are modeled and simulated and the results show that ANN offers a very accurate stock model and also generates competitive systems using all four trading strategies. The results also show that, neural network is a good tool to predict stock price movement with accuracy higher than 95%. Closing price is a good input for neural network model for stock price prediction.

Keywords— Artificial Neural Networks (ANNs), Stock Market, Prediction, Stock Exchange, Backpropagation, FeedForward.
1. INTRODUCTION

The researchers in the field of computational computing and machine learning are attracted to stock price forecasting because it has a significant role in contributing towards the economic growth of Malaysia. The economy of Malaysia increases when the confidence of the investors toward the stock market is increased. An artificial Neural Network is a machine learning and soft computing technique that offers solutions in share price forecasting. Neural Network offers multiple approaches to support the decision making and the problem solving in the area of finance where it has the ability to model nonlinear dynamic system. Nowadays, among fields that attract application of ANNs include biological, physical science, industry, and finance [1].

Bursa Malaysia Stock Exchange is the Malaysian stock market where the public trading of Malaysian’s shares commenced. The first Malaysian stock exchange, namely The Malayan Stock Exchange, emerged in 1960. Currently, it is known as Bursa Malaysia which included a Main Board market and ACE market. The market consists of 823 listed companies. The Kuala Lumpur Composite Index (KLCI) is the name of Malaysian main index. It is a capitalization-weighted stock market index that was introduced in 1986. On Monday, July 6, 2009 Bursa Malaysia adopted FTSE Bursa Malaysia Index methodology to enhance the Malaysian Index. It is now known as FTSE Bursa Malaysia KLCI (FBM-KLCI). The FBM-KLCI includes the top 30 companies in Malaysian Main Board. Bursa Malaysia Main Board are grouped into several sectors namely manufacturing, banking, hotel, industrial products, customer product, Infrastructure, Plantation, Property, Technology and Trading/Services.

This study focuses on predicting stock price of Finance sector, in particular banking sector in Malaysia. This study attempts to use neural network technology to build a neural network trading system specifically adopting feed forward backpropagation learning algorithm to predict stock price of Bursa Malaysia. This study will focus on banking sector and select Maybank Berhad as the stock listed in FBM KLCI.

The stock market reflects economic well-being. It is the major economic predictor and indicator of the economy's aggregate performance due to its ability to fall before economic downturn and to rise before the economy is covered [2]. According to [3], the history of stock market started when Verenide Oost-Indische Company (VOC) which is the Dutch East Indian Company started to issue shares when it was founded on 1602, this lead to the existence of a secondary market in the seventeenth century.

Basically, prices of stock are not fixed. When the stocks are sold to the general public, prices are rising and falling based on independence of the free market. Such constantly changing market forces make it difficult to predict short-term stock market movements. And that is why short-term investment in the stock market is risky. The claim that the behavior of market's stock prices is unpredictable and following the random direction was strongly supported [4].

The basis of this work is to develop ANNs to predict stock price of Bursa Malaysia’s banking sector. To achieve the aim, the following objectives will be achieved.
1) To identify input attributes of ANNs to predict banking sector stock price
2) To identify the suitable network model to predict banking sector stock price
3) To develop ANNs model for predicting banking sector stock price
4) To predict banking stock price using ANNs.

2. BACKGROUND STUDY

A. Forecasting Stock Market Price

Stock market price prediction has attracted much attention from academia as well as business. The objective of investment regardless of the type of investment is trading a known dollar amount today for some expected future stream that is expected to be greater than the current outlay causing it to make it feel a lot like legalized gambling. Thus, the precise prediction of the stock price is important in assisting the investors to improve the stock returns.

In the previous two decades forecasting the share price is becoming a significant field in the research area. The initial attraction to Artificial Intelligence technology grew out of the desire of technical analysts to improve the efficiency of what the investors are already doing, that is developing and testing the systems. As claimed by the authors in [1], along the way it is discovered that besides holding the promise of speeding up development time, Artificial Intelligence technology also has the powerful potential (both real and imagined) to create immensely profitable trading systems, without a great many of the human stumbling blocks that usually occur along the way. Therefore, although the environment of the stock market are very hard to predict, dynamic and complicated. It is still an important research topic in business [6]; [7]; [8]; [9] and [10].

The previous study shows that the success of using the neural network methodology has better forecast accuracy when compared to other method. [11] reported that neural network prediction models surpassed all statistical and other financial time series machine learning models to achieve forecast accuracy of up to 58 percent. In [12], the authors made a evaluation of neural network with regression and other competing statistical models and concluded the relative advantages of neural network are data point in neural network do not have to follow a certain distribution because the sample size effect is not as critical as in regression analysis despite the fact that larger sample size is still beneficial, and lesser problem, although elimination of highly correlated input variables can make the model more robust.

In time series forecasting, the authors in [13] adopted a four layered network with nine input parameters selected from studies of Fortune 500 and Business Week’s top 1,000 firms, resulting in 90 percent correct classification of well-performing firms and 65 percent correct classification of poor-performing firms. Only 65 percent correct classification generated from multiple discriminant analysis technique.
B. Time Series Patterns

The most common methods of forecasting that make estimates of the future based on the past patterns, past relationships, or predictions of subjective. These three methods for assessing the future are used to classify prediction as univariate (using past patterns), multivariate (using past relationships between multiple variables) and qualitative (using subjective judgments and methods). Univariate methods are most widely used. These patterns include outlier, trendy, cyclical, promotional, random and mixing up of the these patterns.

In time series, Data requirement used are the past data to make future prediction and type of forecast are divided into passive forecast and active forecast. In this study active forecast were used because the factors are flexible and are subject to changes.

C. Banking Sector

A bank is a financial unit functioning as a financial distributor. It embraces customers’ deposits and canals them into lending, either directly through lending or indirectly by way of capital markets. A bank connects two parties, namely capital deficits clients and capital surplus customers. Because of its significance within the financial system, Banks have a high regulation in many countries. Generally these banks must meet minimum capital requirement that are based on Intenational capital standard. In Bursa Malaysia, banking companies are listed in Financial Sector. Financial Sector has its own index namely Financial Index. Among the big banking companies listed in Bursa Malaysia Financial Sector are CIMB Bank, Maybank, Public Bank, Hong Leong Bank and RHB Bank. Overall performance of companies in Financial Sector can be observed by analysing financial stock price.

For bank or any such investment based financial institution it is important to have a well-coordinate scientific based understanding of all financial Indexes. Stock market forecasting is an act in which a stock or other financial instrument traded on a financial exchange attempts to assess its future value [14]. A successful prediction of the future price of a stock could generate considerable profit. The efficient hypothesis in the market suggests that the random hypothesis governs stock price movements and thus is unpredictable. Some people disagree and have a multitude of techniques and methods that supposedly allow them to obtain future information on prices. Artificial Neural Network (ANN) is one of such many techniques that is able to extrapolate a correlation, however non-linear, between stock price and some random indicators such as previous day closing price, volume, trading etc, to give an accurate prediction of future stock. There by allowing the bank to maximize its profits on stock prediction.

The Importance of predicting stock price for banking sector realized by [15] by founding that most stock market indicators are closely related to the development of the banking sector, whereas most well-developed stock markets have well-developed banking sectors. Furthermore, [16] show that companies are expanding more quickly than individual corporate

\[ y_i = f_i \left( \sum_{j=1}^{m} w_{ij}x_j - \theta_i \right) \]
characteristics in countries which have better functioning banks and capital markets. Moreover, there is an important role in providing a deeper understanding of banking vulnerability and the stock price mechanism on the real markets in the causal link between stock prices and bank loaning activities [17]. Additionally, the inherent frailty of the banks in the crisis has resulted in the collapse of many banking sector, especially stock markets.

D. Artificial Neural Network (ANN)

Artificial neural networks is a smart data mining tool used in many challenging processes of pattern recognition, including stock market prediction. As artificial neural networks (ANNs) have become very effective method for stock market forecasting due to their ability to deal with unpredictable, flirty or inadequate data, which vary rapids within very short periods of time. Numerous studies and implementations of the ANNs for the resolution of business problems in comparison to traditional methods without artificial intelligence have shown their benefit. According to [18], Production / operations (53.5%) and financial activities (25.4%) are the most common areas of Ann's applications in the past 10 years. Finance ANNs have their most common stock efficiency applications and stock selection forecasts. However, having said that, for a general machine analysis, no systematic procedure exists to establish an optimal neural network. This article predicts that a company holds an stock price based on its stock value history of two forms of neural networks, one Multi-layered Perceptron (MLP) and one Recurrent Elman (backpropagation) network.

In neural network, the relationship and rules are essentially developed and are typically reasonably broad for accurate interpolation in high-dimensional spaces. The success of neural network design depends on its capability to perform against unknown data or the data that the system has never seen during the design stage. The biggest drawback of neural network as specified by [19] is neural network cannot explain results. An ANN comprises of a series of processing components which are interconnected, also known as neurons or nodes [20] and [21]. It can be represented as a directed graphs where each node $i$ execute a transfer function $f_i$ of the form.

Neural network is known to be an effective method or stock market prediction [22]. It is functioning like human brain where it can learn from experience and handle non-linear data as well as dealing with partial application domains like a stock market behaviors. The neural network performance is better than other linear method because the accuracy of performing the forecasting is higher and consumes shorter time. In training Neural Network resembles the brain in two respects:

1. Knowledge is obtained through the learning process by the network from its context.
2. For storing acquired knowledge, interneuron link strengths, called synaptic weights are used.
E. Neural Network in trading system

Generally speaking, neural networks are trained by comparing the output it produces with an expected result and feeding the error back into the network so it can correct itself slightly. The next time the neural network is tested the output is slightly closer to the expected result. Technical indicators can be incorporated in a new and potentially advantageous way to reach positive results in a neural network based trading system [23]. There are three stages of the trading system: preprocessing of data, return rate forecast and post processing. For ANNs in order to determine accurate prediction of a dynamic trading system a pre calculated, with corresponding results of a data preprocessing is added. This helps the ANN to learn and adapt to the trading system. A deliberate error is also added so that the ANN algorithm knows how to correct itself. It is through this adapting process that an ANN is able to learn about the dynamics of the trading system and give a far better performance or prediction than any comparing traditional linear statistical approach.

Specifically looking at the suitability of neural network technology in trading systems revealed an impressive number of positive outcomes even when compared to the renowned statistical approach. The key attractions are firstly its ability to detect patterns unrecognized by traders because it is too subtle. Secondly is adaptive nature, where neural network could learn from experience, and even adapt to markets as condition changes.

F. Positive Achievements of Trading Systems Developed

According to [24], the capability of neural network in modeling linear or non-linear relationship among variables derived from technical, fundamental and macro factors is a significant strength that leads to improve predictive accuracy and show significant market timing ability.

The work in [25] proved that neural networks with technical indicators generate significant profits and better results than machine learning and statistical models.

ANNs is economical promising and has achieved best performance compared to the other techniques [26]. According to [27], ANNs is a method that is well tested for analyzing the financial areas of the stock market. There have several research done to test the effectiveness of the stock market prediction using neural methods such as detected by [28] where the study detected stock market forecasting using the ANNS kernel approach and the method of recursive error forecasting. The authors concluded that ANNs perform better than conventional methods. In general, many studies shows that ANNs able to forecast the stock market precisely compared to other methods [28].

Further remarkable outcomes are pointed out. For example, the authors in [29] reported developing a neural networks stocking selection system that has exceeded the SP500 performance. The mentioned system attempts to autonomously discover current market trends as well as rediscover new trends as the market evolves to new states. The advantage highlighted is that the system does not contain rules or patterns discovered by humans. Their decisions are based entirely on patterns it discovers. The authors in [30] have used improved
sine cosine algorithm (ISCA) with ANN to validate the predictability of stock price direction which outperforms other methods.

G. The way ANN learn and process
Neural Network can learn and improve its performance through learning from its environment is of primary significance. As shows in Figure 1, According to [31], the improvement in performance through minimizing classification error, takes place over time via an iterative procedure of adjusting its weights synaptic and levels of bias and also with some decided measures.

3. THE LEARNING PROCESS
Learning is a mechanism through which a neural network's free parameters are modified by simulating the world in which the network is embedded [29]. Learning paradigm refers to an environmental model in which neural network operates and it can be classified into neural network learning paradigm as follows:

1) Supervised Learning
Supervised learning uses a number of inputs that know the relevant (desired) outputs. To calculate the connection with the weight of the neural network, the difference between the desired and actual output is used.

2) Unsupervised Learning
During unsupervised learning, the network displays only input stimuli. The network is self-oriented, internally structured so that every hidden processing factor strategically responds to different input stimuli (or stimuli groups). Unsupervised learning is useful for clustering analysis.

3) Backpropagation
This approach has been highly successful in multi-layered neural network training. The network will not only strengthen the way it performs a task, but also filter information about errors back through the system and adjust the relationship between layers, thus improving performance. This algorithm also tries to reduce the error among the actual desired output.
and the output of ANN. After comparing both by calculating it then the feedback through the network causing a synaptic weights that to change into a minimize error. The process is repeated up to a minimum value of the error. The parameter of the study, or the rate of learning, controls the rate of weight changes.

![Figure 2: The Architecture of Backpropagation Network](image)

**A. Feed Forward Neural Networks**

There are layers in a typical neural network. There is a source-node input layer and a neuron output layer in a single layered network. Moreover, one or more layers of hidden neurons are covered by a multi-layer network. Both network types are shown in Figure 3. The ability to extract higher order statistics from (input) data increases the ability of additionally hidden neurons. A network is also said to be fully connected if each node of each network layer in the adjacent forward layer is connected with each other.

![Figure 3: Feed-in network with a single neuron output layer (a) and fully connected feed-in network with a single hidden layer and a single output layer (b).](image)

**4. ARCHITECTURE**

This section explain the methodology used in the study, which include data, data preparation, pre-processing and neural network model.

**H. Type of Neural Network Model**

For this study, Feed Forward Back Propagation model is used as the neural network type, Training Function Gradient Decent With Momentum As a training function (TRAINGDM), the Learning Function Gradient is used Decent With Momentum (LEARNNGDM), the Mean Squared Normalated Error (MSE), the Transfer Function (TANSIG) is used and there are two layers including the hides and the output layer. All the network type and functions are available in a standard Matlab package. The Table 1 below shows the detail of the network.
Table 1: Neural Network Setting

<table>
<thead>
<tr>
<th>Network Type</th>
<th>Feed Forward Propagation</th>
<th>Back Propagation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Function</td>
<td>Training Function Gradient Decent With Momentum (TRAINGDM)</td>
<td></td>
</tr>
<tr>
<td>Adaption Learning Function</td>
<td>Learning Function Gradient Decent With Momentum (LEARNGDM)</td>
<td></td>
</tr>
<tr>
<td>Performance Function</td>
<td>Mean Squared Normalized Error (MSE)</td>
<td></td>
</tr>
<tr>
<td>Transfer Function</td>
<td>Hyperbolic Tangent Sigmoid (TANSIG)</td>
<td></td>
</tr>
<tr>
<td>Number of Layers</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

1. Data
The data used in this study is obtained from Yahoo Finance, where it provides historical data of all Malaysia stocks. For this study, the historical data collected from the Yahoo financing of Maybank Berhad, which includes opening prices, lowest prices, highest price, volumes and closing prices. For this study, 2 years historical data from January 2016 to December 2017 is used.

2. Identifying Attributes and Selection of Input Variables
For the input variable we used the data available from Yahoo Finance, which include opening price, lowest price, highest price, volume and closing price. The items are all important and may have effect in determining share price movement. Thus we group four set of input data with certain of attributes. Detail of input data set are provided in Table 2.

Table 2: Description of Data Set

<table>
<thead>
<tr>
<th>Input Data Set</th>
<th>Description</th>
<th>Number of Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set 1</td>
<td>5 days of closing price</td>
<td>5</td>
</tr>
<tr>
<td>Data Set 2</td>
<td>5 days of closing price and volume</td>
<td>10</td>
</tr>
</tbody>
</table>
Data Set 3 | 5 days of closing price, and different of daily lowest price and highest price. | 10

Data Set 4 | 5 days of closing price, volume, and different of daily lowest price and highest price. | 15

K. Data Preparation

i. Clean Data

Data was cleaned by removing all prices during close or suspension period. This is because during the closing and suspension no transaction is allowed. Hence the volume would be zero and the closing price carry the previous day price. After data cleaning, there is about 712 input data is available for use in neural network.

ii. Training and Testing data

Training dataset is used to develop the model, whereas the testing dataset is used to measure how well the model interpolates over the training dataset. Selection of the training and testing datasets is particularly important to the development of effective models [1]. From the 712 input data 612 input data were used for training while 100 input data were used for testing. In summary, Table 3 show the data used in the experiment. Several setting of epoch were used during the training. Each network model was trained for 1000, 3000 and 5000 epoch.

<table>
<thead>
<tr>
<th>Data Sets</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Set</td>
<td>612 Input Data</td>
</tr>
<tr>
<td>Testing Set</td>
<td>100 Input Data</td>
</tr>
<tr>
<td>Total</td>
<td>712 Input Data</td>
</tr>
</tbody>
</table>

iii. Data preprocessing

In this phase the data have been preprocess prior to feeding into the network. According to [32] processing has several advantages: it extracts the information that has real value to the network and it can consolidate several variables to one input data that give more meaningful information. We do simple pre-processing to obtain the different between the lowest daily price and highest daily price, where the different is highest daily price minus lowest daily price.
I. Define Network Structure and Selecting Learning Algorithm

i. Network Structure

The network structure for forecasting stock price of KLCI banking sector using artificial neural network backpropagation technique is a multilayer feedforward network, meaning the neuron outputs feed forward to subsequent layers are as follows.

![Network Structure Design](image)

Figure 4: Network Structure Design

ii. Select a learning algorithm

The learning algorithm is in accordance to the claim made by [1] stating that the most commonly used neural network technique in financial modeling is the backpropagation learning algorithm.

![Backpropagation Learning Algorithm](image)

Figure 5: Backpropagation Learning Algorithm

5. SIMULATION APPROACH

The design of the neural network simulator mimics the following recommendation by [33], with two major modules:

i. Learning phase

The appropriate architecture selected in this study is the multilayer feedforward network. In the learning phase, the simulator is capable of reading subsets of instances which used to train and test the network by means of the backpropagation learning algorithm.

ii. Generalization phase

Subsequently the simulator is capable of reading prediction records illustrated to the network without revealing the anticipated output. The performance of the network is then assessed by output comparison generated by the network simulator with the output anticipated.
6. RESULT ANALYSIS

Matlab’s Neural Network package was used to conduct the experiment. Matlab is used because it is a reliable and established package that has been used by many researchers. Figure 6 to 9 are screen capture of some of the experiments.

![Figure 6: Performance (MSE)](image)

![Figure 7: Neural Network Training](image)
Figure 8: Training, that Gradient = 0.012114 at epoch 5000. Validation checks = six, at epoch 5000.

Figure 9: Regression (plot regression), Regression = 0.94035

Table 4 shows the percentage of errors for each network model (with different input data set and number of learning) when tested with testing data. All network model with 1000 epoch of training achieved more errors than network model with 3000 and 5000 epoch. All network model with 5000 epoch of training achieved more errors than network model with 3000 epoch. This indicates that training with 1000 epoch suffer from insufficient training, while training with 1000 epoch suffer from over training. For all network model, training with 3000
epoch achieved the best result. This indicates 3000 epoch is the optimal iteration for training neural network in predicting Maybank Berhad share price.

Table 4: Percentage of Errors of Different Network Model

<table>
<thead>
<tr>
<th></th>
<th>1000 epoch</th>
<th>3000 epoch</th>
<th>5000 epoch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set 1</td>
<td>2.001986%</td>
<td>1.332703%</td>
<td>1.625854%</td>
</tr>
<tr>
<td>Data Set 2</td>
<td>2.521365%</td>
<td>1.335073%</td>
<td>1.547953%</td>
</tr>
<tr>
<td>Data Set 3</td>
<td>1.523605%</td>
<td>1.518211%</td>
<td>1.571076%</td>
</tr>
<tr>
<td>Data Set 4</td>
<td>1.678359%</td>
<td>1.533301%</td>
<td>1.543805%</td>
</tr>
</tbody>
</table>

Among the all network model trained with 3000 epoch, network model trained with Data Set 1 achieved the least error that is 1.332703%, network model trained with Data Set 2 achieved second best performance that is 1.335073%, network model trained with Data Set 3 achieved third best performance that is 1.518211% and network model trained with Data Set 4 achieved worst performance that is 1.533301%. This indicates closing price alone can be a good input to neural network for predicting share price. Too many inputs such as Data Set 4 that contains 15 input unable to produce accurate share price prediction. As evident from the result, volume of the transaction is a better input than the different of lowest daily price and highest daily price.

7. CONCLUSION
Predicting stock price market is becoming famous and challenging nowadays. Neural Networks is a good tool for predicting stock price. As evident from the experiment, prediction using neural network can produce high accuracy of stock price. Table 5 below show the accuracy of share price prediction, where the best prediction achieved 98.67% accuracy, and the worse prediction achieved 97.48% accuracy.

Table 5: Percentage of Accuracy of Different Network Model

<table>
<thead>
<tr>
<th></th>
<th>1000 epoch</th>
<th>3000 epoch</th>
<th>5000 epoch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Set 1</td>
<td>97.9980%</td>
<td>98.6673%</td>
<td>98.3741%</td>
</tr>
<tr>
<td>Data Set 2</td>
<td>97.4786%</td>
<td>98.6649%</td>
<td>98.4520%</td>
</tr>
<tr>
<td>Data Set 3</td>
<td>98.4764%</td>
<td>98.4818%</td>
<td>98.4289%</td>
</tr>
</tbody>
</table>
Data Set 4

Also it is interesting to note that closing price as a sole input type can be a very good input for neural network in predicting stock price. However, when it is combined with other input types such as volume of transaction and different of price between daily lowest and daily highest price the performance may drop.

Nevertheless, this is only a preliminary study and obviously there is ample room to increase the neural forecasting power before reliable implementation is carried out to yield positive return on investment through accurately and adaptively selecting winning stocks.

8. REFERENCES


