PREDICTION AT KUALA LUMPUR STOCK EXCHANGE (KLSE) USING ARTIFICIAL NEURAL NETWORK

MUHAMMAD HAFIZ BIN SAMSURI

BACHELOR OF COMPUTER SCIENCE (SOFTWARE ENGINEERING)

UNIVERSITI MALAYSIA PAHANG (UMP)

2014
ABSTRACT

It is known that Artificial Neural Network (ANN) technique that can make prediction. Stock market prices are tough to predict because there are many factors involve thus making it hard to make prediction. Three counters from KLSE, AIRASIA, AMBANK, and ASTRO are chosen to compare the prediction between those counters. Six month's of data from each of those counter are used for the predictions and compared for getting the results. Two values of predictions for today and yesterday are obtained by using the prototype. By comparing both values, the increase and decrease of the stock prices can be revealed. The results from the prototype shows that ANN can be used to predict increase or decrease of the stock prices for each of the related counters.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER 1</th>
<th>INTRODUCTION .............................................................................................................................</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1 INTRODUCTION ..............................................................................................................</td>
</tr>
<tr>
<td></td>
<td>1.2 PROBLEM STATEMENT .................................................................................................</td>
</tr>
<tr>
<td></td>
<td>1.3 OBJECTIVES .....................................................................................................................</td>
</tr>
<tr>
<td></td>
<td>1.4 SCOPE ................................................................................................................................</td>
</tr>
<tr>
<td></td>
<td>1.5 THESIS ORGANIZATION ................................................................................................</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>LITERATURE REVIEW .....................................................................................................................</td>
</tr>
<tr>
<td></td>
<td>2.1 INTRODUCTION ..............................................................................................................</td>
</tr>
<tr>
<td></td>
<td>2.2 FUZZY LOGIC ...................................................................................................................</td>
</tr>
<tr>
<td></td>
<td>2.3 ARTIFICIAL NEURAL NETWORK ................................................................................</td>
</tr>
<tr>
<td></td>
<td>2.4 RULE INDUCTION .........................................................................................................</td>
</tr>
<tr>
<td>CHAPTER 3</td>
<td>METHODOLOGY ...........................................................................................................................</td>
</tr>
<tr>
<td></td>
<td>3.1 INTRODUCTION ............................................................................................................</td>
</tr>
<tr>
<td></td>
<td>3.2 METHODOLOGY .............................................................................................................</td>
</tr>
<tr>
<td></td>
<td>3.3 HARDWARE AND SOFTWARE .............................................................................................</td>
</tr>
<tr>
<td></td>
<td>3.4 GANTT CHART ...............................................................................................................</td>
</tr>
<tr>
<td>CHAPTER 4</td>
<td>DESIGN AND IMPLEMENTATION .................................................................................................</td>
</tr>
<tr>
<td></td>
<td>4.1 INTRODUCTION ............................................................................................................</td>
</tr>
<tr>
<td></td>
<td>4.2 DEFINE ALGORITHMS .....................................................................................................</td>
</tr>
<tr>
<td></td>
<td>4.3 DATA GATHERING ..........................................................................................................</td>
</tr>
<tr>
<td></td>
<td>4.3.1 AIRASIA ...................................................................................................................</td>
</tr>
<tr>
<td></td>
<td>4.3.2 AMBANK ..................................................................................................................</td>
</tr>
</tbody>
</table>

---

---
4.3.3 ASTRO

4.3.4 The Usage of Data

4.4 The Interface Design for the Prototype

4.4.1 Interface1.java

4.4.2 Interface2.java

4.4.3 Interface3.java

CHAPTER 5

RESULTS AND DISCUSSION

5.1 INTRODUCTION

5.2 RESULT ANALYSIS

5.3 COMPARISON BETWEEN PREDICTION AND REAL DATA

5.3.1 Comparison of AIRASIA Data

5.3.2 Comparison of AMBANK Data

5.3.3 Comparison of ASTRO Data

5.4 RESEARCH CONSTRAINS

5.4.1 Development Constraints

5.4.2 System Constraints

CHAPTER 6

CONCLUSION

REFERENCES

APPENDIX A

Calculation1.java

Calculation2.java
LIST OF FIGURES

Figure 1: Flowchart of the thesis ................................................................. 12
Figure 2: Artificial Neuron Structure ......................................................... 16
Figure 3: Sigmoid Curve ........................................................................... 17
Figure 4: Interface 1.java ......................................................................... 31
Figure 5: Interface2.java .......................................................................... 32
Figure 6: Interface3.java .......................................................................... 33
LIST OF TABLES

Table 2: Hardware and software used in the system .......................................................... 13
Table 3: Six month of AIRASIA data ............................................................................... 22
Table 4: Six month of AMBANK data ........................................................................ 26
Table 5: Six month of ASTRO data ............................................................................... 30
Table 6: Comparison of AIRASIA real and predicted data .......................................... 40
Table 7: Comparison of MAYBANK real and predicted data ..................................... 44
Table 8: Comparison of ASTRO real and predicted data ........................................... 47
LIST OF ABBREVIATION

KLSE - Kuala Lumpur Stock Exchange
FRET - fuzzy rule extraction tool
FAM - Fuzzy Associative Memories
ANN - Artificial Neural Network
A stock market or equity market is a public entity for the trading of company stock (shares) and derivatives at an agreed price. The stocks are listed and traded on stock exchanges which are entities of a corporation or mutual organization specialized in the business of bringing buyers and sellers of the organizations to a listing of stocks and securities together. Sometimes, the market seems to react irrationally to economic or financial news, even if that news is likely to have no real effect on the fundamental value of securities itself. So, it is quite hard to really predict the ups and downs of stock price. Luckily, there exist techniques in Artificial Intelligence (AI) especially for making prediction. Fuzzy logic, Artificial Neural Network, and Rule Induction are three of the techniques that are proven effective in making prediction. Three (3) counters will be compared from Kuala Lumpur Stock Exchange (KLSE) as references to prove the technique used in this thesis.
1.2 PROBLEM STATEMENT

Stock market is a place that has high risk because stock price changes frequently and hard to predict. The prices will depend on many factors such as internal developments of the company, world event, inflation and interest rate, exchange rates, and hype. For internal developments, stock price of a company will increase. World event such as war has a lot of effect on stock price too. The periodic adjustment of inflation and interest rate, also exchange rates for money affect stock price. Hype involves the release of new product or promotion of a company that directly involves the stock price of that company. It is problematic to make decision about the stock even for an experienced person. They must have perfect timing either want to sell, buy, or hold the stock in their portfolio. One wrong move will cause great losses of money and time. Also, human mind cannot analyze data as fast as computer when it comes to gigantic amount of data. There is simply not enough time to analyze all that data as fast as possible especially for a human mind.
1.3 OBJECTIVE

The objectives of the research are to:

1) To build a prototype that can make prediction for stock market prices at Kuala Lumpur Stock Exchange (KLSE).
2) To apply the use of Artificial Neural Network, an Artificial Intelligence technique used for prediction.
3) To test the ability of the prototype to predict the increase or decrease of the stock market prices at KLSE.
1.4 SCOPE

1) Make use of 6 months' worth of data from Kuala Lumpur Stock Exchange (KLSE) to make prediction.
2) Compare 3 related counters chosen from Kuala Lumpur Stock Exchange (KLSE)
3) Compare 3 different Artificial Intelligence techniques to be used for prediction, Fuzzy Logic, Artificial Neural Network, and Rule Induction.
1.5 THESIS ORGANIZATION

This thesis consists of six (6) chapters:

Chapter 1: This chapter will discuss on introduction to research on the Prediction of Stock Market Using Artificial Neural Network. Problem statement is also discussed and also the scopes and objectives of this thesis.

Chapter 2: This chapter is literature review, will explain more about the selected project and also previous work related to it. This is where information related with regards to the project will be explained.

Chapter 3: This chapter proposed the methodology that discusses the overall approach and framework of research. It should cover method, technique or approach to be used.

Chapter 4: This chapter is design and implementation. It is to develop the framework and model through flow work, continuously design the research which includes any planning of data analysis.

Chapter 5: This chapter is results and discussion. Its purpose is about finding the results from data analysis. The important content is Result Analysis, and Research Constrains.

Chapter 6: This chapter is for conclusion. Conclusion for all the researches that have been done throughout the chapters must be included.
CHAPTER 2
LITERATURE REVIEW

2.1 INTRODUCTION

Stock market prediction is an area that has been gaining long interest across the fields of finance and many research communities. The difficulty of predictability of the financial markets has stood for a long time while many techniques are invented in order to reduce it. Technical analysis believes that historical prices and other indicators can reveal correlations and patterns of stock price movements and leads to the predictability of future prices. In recent years, more advanced computational techniques such as artificial neural networks, genetic algorithms and support vector machines make the analysis and learning of such patterns more popular. These machine learning techniques have been seen successfully application on text based task such as information retrieval, text categorization and clustering. Holding the belief that stock prices are highly correlated and sensitive to unseen news and events that bring new information for the market to absorb, approaches utilizing learning scheme to find out such correlation and impact from past in order to predict future reaction on news are devised.

In general the approaches to predict stock price could be roughly categorized into two kinds, fundamental analysis and technical analysis. Fundamental analysis is based on macroeconomic data, such as export and imports, money supply, interest rates, inflationary rates[8], foreign exchange rates, unemployment figures, and the basic financial status of companies such as dividend yields, earnings yield, cash flow yield, book to market ration, price-earning ratio, lagged returns, and size.
Technical analysis is based on the rationale that history will repeat itself and that
the correlation between price and volume reveals market behavior. Prediction is made
by exploiting implications hidden in past trading activities, and by analyzing patterns
and trends shown in price and volume charts.

Basically the test of weak form efficient market is to test whether there exist
excess return by using technical analysis. There have been some researches claiming the
existence of the weak form of efficient market [9]. Also some claiming that the weak
form efficient market does not exist, and so far the research remains inconclusive.

One of the most commonly used methods in technical analysis is the moving
average filter rule. The criterion is that the buying signal happens when the short term
moving average line breaks through the long term moving average line from down, and
the selling signal happens when the short term moving average line breaks through the
long term moving average line from up. The logic behind this rule is to identify periods
when expected returns deviate from unconditional means [10]. Although concluded that
the filter rules are not useful, the non trivial ability to predict the price change by using
the filter rules.

Similar to the filter rules, KD technical rules proposed by Lane (1957), is trying
to capture the period when expected returns deviate from unconditional means by using
K and D indexes instead of the moving averages. Essentially K and D indexes with the
advantages of momentum, relative strength, and moving average, and with the
consideration of the highest and the lowest prices, are expected to be capable of
capturing the short term variance. However, the KD filter rules could be too simple to
be effective. Besides, the parameters for the rules are arbitrary.

However, this thesis is all about making prediction for stock market at KLSE by
using certain techniques that can be use for prediction. Several peoples agreed that there
are three AI techniques identified for predicting stock market; Artificial Neural
Network, Fuzzy Logic, and Rule Induction [1], [2], [3].
2.2 FUZZY LOGIC

Fuzzy logic are trading models designed usually consist of a few fuzzy rules expressing the relationship between inputs and desired output of the market. In these models, inputs are fuzzified, membership functions are created, association between inputs and outputs are defined in a fuzzy rule base, and fuzzy outputs are rest at the crisp values [4]. Fuzzy logic is already in used a lot of time by other people.

Benachenhou (1994) developed a fuzzy rule extraction tool (FRET) that extracts fuzzy rules from input-output data by FAM method, and then uses them in a fuzzy decision support system. A fuzzy rule set derived from sample data is then used as a fuzzy expert system for trading. The model achieves the ratio of winning versus losing trades equal to 4.6.

Assisting the traders in arriving at purchase decisions, Man and Bolloju (1995) implemented a prototype of a fuzzy rule based decision support system. To extract and transfer dealers' expertise, they employed unstructured interviews with some experienced dealers. Fuzzy rules representing the dealers' decision making process are quite close to the terminology used by the dealers and the rules are easily interpretable by the dealers. Authors believe that use of fuzzy logic for knowledge representation has facilitated a high level of abstraction of the experts' knowledge. Moreover, the flexible relationship represented by membership functions and fuzzy rules, between the variables in the model have provided a robust model of the decision making process.

Fuzzy Associative Memories (FAM) a method proposed by Kosko (1992) was used to determine market rules. In Kosko's method, the weight vector of a network trained by input-output data is considered as the membership function of input-output space. The model is trained by daily data of six months, and the rate of correct prediction is found to be 74%. The duration in which this performance is achieved is not mentioned in the literature
Ye and Gu (1994) developed a hybrid neuro-fuzzy model in which fuzzy logic enhances a neural trading system. The model takes the closing price, and Shanghai Stock Indicator of individual stocks as inputs, and outputs three values indicating whether a trend is ascending, descending or stationary.

2.3 ARTIFICIAL NEURAL NETWORK

Neural networks are usually used for pattern recognition or classification, are a connected set of simple processing elements or nodes, where a weight is associated to each connection between nodes. Weights are initialized randomly at the beginning, and as the network begins to learn, the weights change [5]. Like fuzzy logic, neural network also have seen to be used for prediction.

Several systems based on neural networks have produced promising results. Most of them only use indicators and historical market data, such as moving average, or closing price. A study has been done to predict stock prices on short-term, day to day type of prediction. It is tested by using three German stocks randomly chosen to compare the effectiveness of neural networks and BACK-PROPAGATION. The results for neural networks were outstanding within 10 days time span, an accuracy of up to 90% was achieved from the prediction between those three German stocks. By comparing both technique, it led to the expectation that neural network could improve prognosis of stock prices in the future [18].

James Hall (1994) built a stock selection system using neural networks. Unlike many stock selection methods, this system does not use expert rules. Instead, the system discovers some patterns in the market and selects attractive stocks based on them.

Gencay (1998) used a total of 90 years of daily Dow Jones Industrial Average Index from 1897 to 1988; to examine the predictability of the stock market. The market is modeled by single layer feed for neural networks. The technical trading used in this study is very simple and popular, like the moving average. The simplest version of this rule suggests that traders sell, whenever the price climbs above its moving average, and
buy when it drops below. The results report strong evidence of nonlinear predictability in the stock market returns by using the past buy and sell signals of the moving average miles. They also suggest that it is worth investigating more elaborate rules aid the profitability of these rules after accounting for transaction costs.

Lawrence (1997) stated that neural networks are not perfect in their prediction, but their performances are better than all other methods and some day make dynamic, chaotic systems such as the stock market can be fully understood [19].

2.4 RULE INDUCTION

Rule induction or decision tree methods are capable of culling through a set of predictors by successively splitting a data set into subgroups on the basis of the relationships between predictors and the output field.[6]

All the techniques mentioned are all used already for many existing systems and proven to be effective. However, only the best technique is going to be used for the system in this thesis and it is neural networks. Schöneburg (1990) has done research using neural networks and stated that the prediction using that technique gave 90% accuracy than other techniques. It is also simple and easy to implement in the system to make prediction fast and efficiently.

The three techniques are compared and neural networks proven to be the best among all three techniques.
CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

The project is done by doing research about stock market at KLSE. Trying to enter the stock market is difficult for inexperienced brokers, but it is not impossible. By using certain techniques in Artificial Intelligence, buying or selling are much easier even for beginners. First, the research is done by researching other theses about the prediction of stock market. During research, three AI techniques are found applicable for prediction in stock market. The three techniques are fuzzy logic, neural network, and rule induction.

By more researching, it is found that by comparing the three techniques, fuzzy logic is the best choice for predicting stock market. All the other techniques are good but among the theses researched, it is proven that fuzzy logic are better.

Next is detail research about the stocks at KLSE. Research is done thoroughly and three counters are chosen from KLSE, which are AIRASIA, AMBANK, and ASTRO.
3.2 METHODOLOGY

The development of this thesis involves a few steps that must be done to accomplish all the objectives. The following flowchart will show all the important steps that are vital to this thesis.

![Flowchart of the thesis](image)

Figure 1: Flowchart of the thesis
### 3.3 HARDWARE AND SOFTWARE

<table>
<thead>
<tr>
<th>Hardware &amp; Software</th>
<th>Descriptions</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Toshiba Qosmio F60 Laptop</td>
<td>The hardware use to keep all the software and project.</td>
<td>All information are stored here</td>
</tr>
<tr>
<td>2) IDM 6.15</td>
<td>Software used to download information.</td>
<td>Much faster download that saves time.</td>
</tr>
<tr>
<td>3) Netbeans IDE 7.3</td>
<td>Software use to actually make the program for predicting stock market.</td>
<td>Main software for programming.</td>
</tr>
<tr>
<td>4) Chartnexus 4.2</td>
<td>Software that shows all stock markets at KLSE for the last 3 years.</td>
<td>Shows all information about the stock market.</td>
</tr>
<tr>
<td>5) Microsoft Word</td>
<td>Software used to make documentation.</td>
<td>Make it easier with many tools for making document.</td>
</tr>
<tr>
<td>6) Microsoft Project</td>
<td>Software use to make Gantt chart</td>
<td>Make schedule more effective and precise</td>
</tr>
</tbody>
</table>

Table 1: Hardware and software used in the system
3.4 GANTT CHART
CHAPTER 4

DESIGN AND IMPLEMENTATION

4.1 INTRODUCTION

The purpose of this chapter is to develop the framework and model through flow work and continuously designing the research which include any planning of data analysis.

This thesis will discuss the process and data gathering for research purposes and sketching the work flow and model using special software such as Microsoft Visio, Rational Rose or other. Microsoft Words are not allowed.

It will be explain how the data/model/process/hardware been implemented into selected algorithm.

4.2 DEFINE ALGORITHMS

Artificial neural networks are inspired from biological neurons in the brain. The brain has been studied by scientist to understand their complexity and only small portion knows by scientists. Signals are transferred from one neuron to another by synapses using a complex chemical process released from the sending side of the junction. It is this characteristic that the artificial neural network attempt to reproduce
by making artificial ways by copying the brain functions. The neuron model shown in Figure 2 is the one widely used in artificial neural networks for making prediction.

The artificial neuron shown in this figure has $N$ input, represented by $x_1, x_2, x_3, \ldots, x_n$. Each line connected to the inputs is assigned a weight which represented by $w_{1j}, w_{2j}, w_{3j}, \ldots, w_{nj}$. Weights in the artificial neural networks have connections in biological neurons. For the prediction of stock market, there are eight inputs that includes weight are the open, close, highest and lowest stock price for yesterday and today. The input and the weight will be used to find the value of $x$ by using the following formula.

Transfer Function:

$$ x = \sum_{j=1}^{N} w_{ij} x_i $$

The weight, $w$ can be the open stock price of today and close stock price of yesterday, and input, $x$ is the highest and lowest stock price during the day. Four inputs are needed for yesterday and today for making the prediction of the closing price for the stock market.
After finding the value of $x$, activation function needs to be calculated and one formula for the activation function is chosen. The formula is the sigmoid activation function, a mathematical function having an "S" shape of sigmoid curve.

Sigmoid Activation Function:

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

![Figure 3: Sigmoid Curve](image)

The value of $x$ in this sigmoid activation function is used from the output calculated in the transfer function.

The last step is to use formula for prediction by using all the calculation done in transfer function and activation function. The formula used is shown below.

Prediction Formula:

$$Prediction = Range \times (output + low \text{ price})$$

The value for range is the highest price minus the lowest price. The value for output is the value calculated by using the sigmoid function. The value of low price is the lowest stock price during the day.

Two value of prediction will be available that are prediction value of yesterday and today stock price. Lastly, both the values are compared to show prediction of the stock price of today will increase or decrease. This is a Feedforward ANN technique where connections between unit do not form a cycle and move in one direction, forward from input nodes, through hidden nodes to outputs.