CHAPTER 1

1. INTRODUCTION

1.1. Background

In the last years of the previous century, the United Arab Emirates became one of the fastest developing countries in the Middle East and South Asia. Dubai, which is considered the commercial capital and the center of international business in the country, has taken the leadership in developing and modernizing both governmental and private sectors with state-of-the-art strategies, policies, trends, technologies and infrastructure. One of the fastest and most powerful growing sectors is finance.

On March 26th, 2000, Dubai Financial Market commenced operation with listing shares of seven companies and ten joined brokers. The mission of the market was to create a fair, efficient, liquid and transparent marketplace that provides choices through the best utilization of available resources in order to serve all stakeholders [5]. DFM has grown rapidly and has scored magnificent records in terms of trading volume and market values. Today, with more than 50 listed companies, DFM is considered a leading financial market in the Gulf area and Middle East.

1.2. Problem statement

Forecasting market indices and stock prices is an essential topic in finance and has always been a major challenge facing investors. The key factor in predicting price movements is to discover the patterns and relationships in the stock market. It is known that stock prices are usually affected by economical, political and sentimental factors. These factors interact with each other in a very complex manner. In fact, one of the known financial hypothesis is efficient market hypothesis (EMH), which declares that stock price movements in an efficient market are random. In other words, the stock price movements are unpredictable, which in turn means that dramatic monetary profits from stock price movements are rare occurrences, if not impossible. [4] [6] [7] [1]

However, many market professionals have re-evaluated the efficient market hypothesis, as it's proven that certain patterns in price movement take place occasionally. These professionals believe that mechanisms governing such patterns can be extracted and modeled. There have been many methods and techniques examined and applied to predict stock price movements. Mainly, these are classified under fundamental analysis, technical analysis or time series forecasting. [4] [7]

Fundamental analysis is based on examining the financial statement and financial ratios of a company to determine financial strength, future growth and profitability prospects in order to estimate whether the stock's price is undervalued or overvalued. By analyzing the company's operations and the market in which it is operating, it is possible to determine the company's intrinsic values and expected returns. The resulting information is used to forecast future earnings; therefore, stock prices can be predicted eventually. Apparently, with such massive calculations required, fundamental analysis can only be used with long-term investments. [12]

Technical analysis deals with historical prices and volume information on the assumption that history repeats itself and price movement patterns can be extracted from historical price data. Quantitative indicators (such as strength index, moving average, etc) and charting patterns (such as head-and-shoulders, flag, etc) are variables used for price prediction. Technical analysis is commonly used among traders and mainly used for short and medium-term investments. [4] [11]

Time series forecasting techniques such as multivariate regression and autoregressive integrated moving average have been used to model historical price data as non-linear function. Pesaran and Timmermann (1994) presented a good example of using multivariate regression in predicting the S&P 500 index and the Dow Jones Industrial Average. Other developed methods and models are linear auto-regressive models, principal component analysis (PCA), genetic algorithms (GAs) and artificial neural networks (ANNs). In recent years, ANNs have been broadly used in prediction or forecasting studies in all functional areas of business, including accounting, economics, finance, management information systems, marketing and production management. As a result, stock prediction was another key application for ANNs. [6]

1.3. Research objective

Major studies applied ANNs to predict stock price movement in mature markets like the United States, Europe and Japan. There have been much fewer researches on emerging markets like Taiwan, China and the Middle East.

The purpose of this study is to demonstrate the accuracy of ANNs in predicting stock price movements for companies traded in DFM. As a further demonstration of ANNs accuracy, a comparative analysis with another new methodology, Polynomial Classifiers, will be presented.

This research contains:

1. Detailed review of the latest methodologies and models in market forecasting.
2. Development of prediction model using ANN.
3. Development of a comparative prediction model using polynomial classifier.

4. Analysis and discussion of the results obtained.
5. Recommendations on how this model can be enhanced for trading purposes.

1.4. Thesis organization

This research is organized into seven main chapters. Chapter one introduces the research background, the statement of the problem, the research objectives, as well as the structure of content. Chapter two will detail the literature on artificial neural network and polynomial classifiers; in addition to the latest research, work was done in stocks prediction.

Chapter three is a theoretical background on artificial systems in general, with in-depth view of artificial neural network and polynomial classifier in terms of characteristics and design prospects.

Chapter four will illustrate and describe the methodology used in developing the prediction model, how the input data has been set up to properly feed the model, and how outputs were presented. Chapter five shows how the prediction model was implemented and what prediction modes have been used in both techniques (NN and PCs). Chapter six contains the analysis, discussion and comparison of the results obtained by each technique.

Finally, chapter seven lists several recommendations on how this model can be enhanced and used in trading applications; chapter seven also tailors the overall conclusion of this research.

CHAPTER 2

2. LITERATURE RIVIEW

Although Efficient Market Hypothesis (EMH) states that stock prices follow random walk and, hence, are unpredictable, many researchers and practitioners questioned this theory. Engle, 1982, used the ARCH(p) (Autoregressive Conditional Heteroscedasticity) to model the volatility clustering and fail tail characteristics of time series. Due to the large increase of the time lag (p) caused by the ARCH (p) model; Bollerslev, 1986, developed a generalized ARCH (p) as GARCH(p,q). Looking at GARCH(p,q), the leverage effect (a negative effect has bigger influence than a similar positive effect), made GARCH model to be extended to an EGARCH model (Exponential GARCH) developed by Nelson, 1991. [2]. Lo and MachKingley, 1988, applied variance estimators to show the illogic of the random walk model [9]. One of the earliest studies using artificial intelligence in stock market prediction techniques is Kimoto et al., 1990, where several learning algorithms were used for developing a Tokyo Stock Exchange prices index prediction system [1]. Pesaran and Timmermann, 1995, concluded that stock price returns are predictable when market volatility is high. Geneacy R., 1998, illustrated that technical trading rules such as moving average are more successful in predicting exchange rates than models that follow random walk theory. Darrat and Zhong, 2000, conducted a study on the Chinese stock market that showed that the market doesn't follow a random walk theory model [9].

The above mentioned studies and other studies and models require strict assumption about distributions of time series, so it is difficult to model market variables caused by many noises in market conditions and environments. Therefore, the concept of Artificial Neural Network (ANN) has been applied to complex financial markets. Neural Networks are information processing paradigms that are structured in the same way biological nervous systems, such as the human brain, process information. The network is composed of a large number of interconnected processing elements and neurons operating in parallel with a certain function to solve a specific problem. Like the human brain, ANNs learn by example and have the capability of relating the input and output parameters without requiring a prior knowledge of the relationships of the process parameters.

The concept of using ANNs is not new. Hu., 1964, was the first to apply ANNs in his study when he used the Window's adaptive linear network in weather forecasting; the research was limited due to the shortage of training algorithms at that time. Using ANNs for forecasting developed further when Rumelhart et al., 1986, introduced the Backpropagation algorithm. Lapedes and Farber, 1987, conducted a study and illustrated that ANNs can be used for modeling and forecasting nonlinear time series. Werbos, 1988, found that ANNs trained with back-propagation outperformed the traditional statistical methods like regression and Box-Jenkins approaches. [3]

Forecasting nonlinear time series was one of the earliest applications of ANNs, as in Lapedes and Farber, 1987, 1988. A major area using ANNs is in analyzing and predicting deterministic chaotic time series (which occur mostly in engineering and physical science) with and without noise, as in Jones et al., 1990; Lowe and Webb, 1990; Deppisch et al., 1991; Ginzburg and Horn, 1991; Rosen, 1993; Poli and Jones, 1994. Applications of ANNs are vast and cover many disciplines like airborne pollen (Arizmendi et al, 1996), environmental temperature (Balestrino et al., 1994), rainfall (Chang et al., 1991), total industrial production (Aiken et al., 1995), wind pressure profile (Turkkan and Srivastava, 1995) and electric load consumption (Park and Sandberg, 1991; Bacha and Meyer, 1992; El-Sharkawi et al., 1991; Ricardo et al., 1995). [3]

A principle and main application for ANNs is found in the financial domain. Neural Networks have been used in several financial applications like economics, accounting, management information systems, marketing and production management. Alici Y., 1996, showed that ANN performed better bankruptcy prediction for UK companies than the conventional statistical methods such as Discriminant Analysis and Logistic Regression. Robles and Naylor, 1996, applied neural network to commodity trading and showed that ANN outperformed the traditional weighted moving average rule and a buy-and-hold strategy [15].

In stocks price prediction, Artificial Neural Networks have demonstrated an outstanding performance. Kimoto and Asakawa, 1990, showed that excellent profits are achieved when modular neural networks are used in predicting the timing of buying and selling for the Tokyo Stock Exchange. Saad et al., 1998, compared different types of neural networks such as PNN (probabilistic neural network), RNN (recurrent neural network) and TDNN (time-delay neural network) in predicting daily closing prices in stock markets. The results showed that all networks tended to be equally feasible [13]. More recent study conducted by A.-S. Chen et al., 2003, showed that neural network models are useful in predicting the direction of index returns based on the study applied on Taiwan Stock Index [7]. Thawornwong and Enke, 2004, illustrated that neural network models could successfully generate higher returns and lower risks in predicting the directions of future excess stock return than the buy-and-hold strategy, conventional linear regression and the random walk models [1]. Q. Cao et al., 2004, developed a prediction model on the Shanghai stock market using neural network and proved that neural networks offer an opportunity for investors to enhance prediction power in selecting stocks [9]. Tsang, P.M., et al., 2006, designed a prediction system using neural networks to predict short-term price movement directions in the Hong Kong stock market. Their system scored 74% accuracy without the use of extensive market data or knowledge. [4]

Baba and Kozaki, 1992, presented a back-propagation neural network combined with a random optimization technique to predict stock markets in Japan. Results proved that the proposed approach was of significant help in forecasting stock prices. Takanashi et al., 1998, proposed a neural network that embodied multiple line-segments regression techniques to predict stock prices. The proposed model performed well in prediction. Leigh et al., 2002, combined pattern recognition with neural networks to predict the New York Exchange Composite Index. The results gave confidence in the developed model [13].

From all the above mentioned studies and researches, it is strongly proven that ANNs have an outstanding ability in financial forecasting in general, and stock price movements in particular.

In recent years, Polynomial Classifiers started to have more dominance in artificial intelligence applications. Polynomial Classifiers (PC) are discriminative models of neural classifiers. Whether used to model the manifolds of each class or to discriminate the patterns of different classes, neural classifiers can be divided into relative density models and discriminative models. Examples of relative density models include mixture linear models and auto-association networks. Whereas, Discriminative neural classifiers include the multi-layer perceptron (MLP), the radial basis function (RBF) net, and the polynomial classifier (PC). [8]

PCs can be described as higher-order neural networks which consist of a single-layer network with the polynomial terms of patterns feature as inputs [8]. The polynomial classifiers are learning algorithms proposed and adopted in recent years for regression, classification and recognition with significant properties and generalization capability [10].

Due to their need for less training examples and far less computational requirements, PCs have shown superior performance to multilayer neural networks. One of the most used applications of polynomial classifiers is recognition and identification. K.T. Assaleh and W.M. Campbell have applied polynomial classifiers to speaker identification and speaker recognition [17, 18]. They reported excellent results and achieved higher accuracy compared to other traditional methods. Another recent study conducted by K. Assaleh and M. Al-Rousan (2004) on recognition of Arabic sign language alphabet using polynomial classifiers delivered superior recognition results [16].

During this research preparation, it is worth mentioning here that there were no dedicated researches that have studied the polynomial classifiers into the prediction of stock price movements. In addition, there were no researches found that reported using neural network to predict stock price movements in Dubai Financial Market.

This thesis aims to develop and analyze a prediction model using two different techniques – neural network and polynomial classifiers – to forecast the stock price movements in Dubai Financial Market.

CHAPTER 3

3. INTELLIGENT SYSTEMS

3.1 Artificial Intelligence Overview

Known as “the science and engineering of making intelligent machines”, Artificial Intelligence refers to automating tasks that demonstrate intelligent behavior. Examples include control, planning and scheduling, handwriting, natural language, speech and facial recognition. The applications of Artificial Intelligence in general can be grouped into two types: Classifiers and Controllers. In concept, classifiers are based on pattern recognition, and can be seen as functions that can be formed based on trials or examples. These examples are known as observations or patterns. The defining feature of intelligence is the capability of learning from past experience and solving problems when important information is missing in order to be able to handle complex situations and to react correctly to new ones. The classifiers performance depends greatly on the characteristics of the data to be classified.

During the nineties of the previous century, Artificial Intelligence has become a rich area for researches and studies. Many models and systems were developed and the most widely used ones were the neural network , support vector machines, k-nearest neighbors, Gaussian mixture model, Gaussian, naive Bayes, decision trees, radial basis functions, and polynomial classifiers.

In this research, the focus will be on Neural Networks (NN) and Polynomial Classifiers (PCs). The following two sections will examine both systems more deeply.

3.2 Neural Network

Neural Networks can be described as interconnected network of processing elements known as artificial neurons, with different weights assigned to each connection. The network can approximate any function that maps between the inputs and outputs, provided that proper topology and suitable weights have been used.

There are several types of neural networks. The most common type is the Feedforward Neural Network. Other types include Recurrent Neural Network (RNN), Probabilistic Neural Network (PNN), Time-Delay Neural Network (TDNN), Radial Basis Function Network (RBF), and other types.

The feedforward neural network computes input to output mapping based on calculations that occur in interconnected nodes. These nodes are known as hidden nodes and are arranged in layers. Calculations in each of the hidden nodes are done as sigmoidal function of the weighted sum of inputs from the input layers.

The back propagation principle allows the network to learn the weights or the connections between the nodes through data training, aiming to result in a minimum vale of the least square error between the actual values and the estimated values as output of the neural network.

The basic structure of a feedforward back propagation network consists of:

- X input nodes.
- K hidden nodes
- Y output nodes.

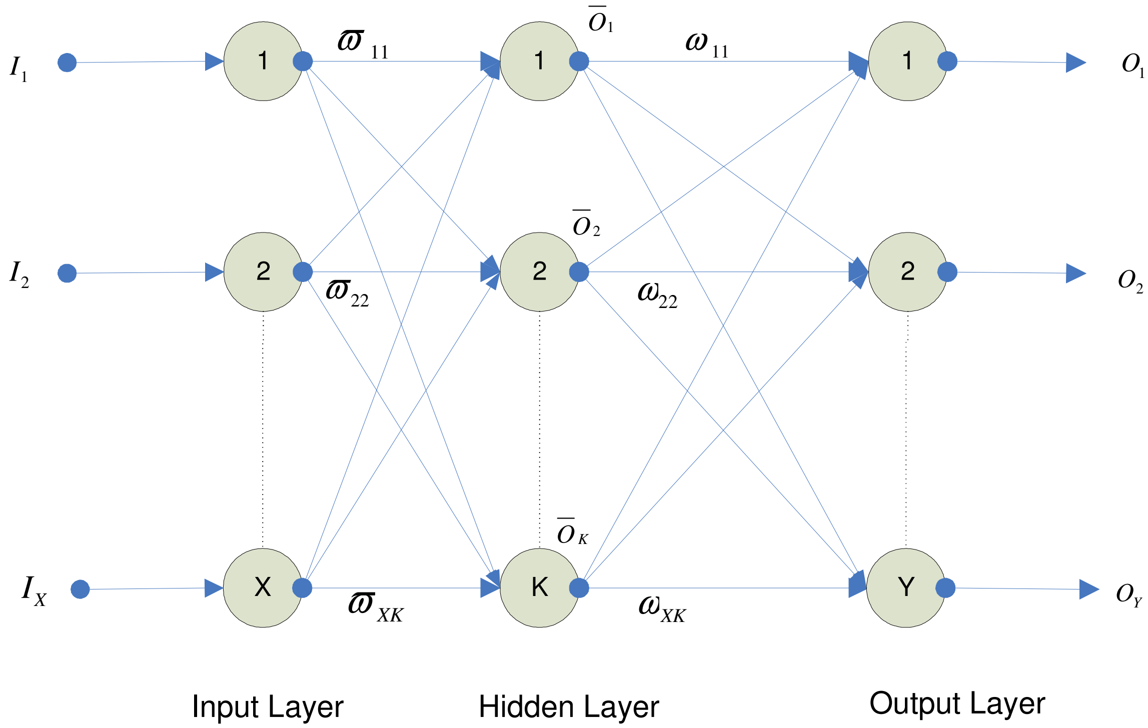


Figure 3.1 feed forward neural network structure

To simplify the algorithm of Back Propagation, it can be summarized in three phases:

Let,

O_y : The output value of the output layer Y

$\overline{O}_k$: The output of the hidden layer K

ω_{ky} : The network weight for the hidden layer and the output node

ϖ_{xk} : The network weight for the input node and the hidden node

I_x : The input value for the input layer X

T_y : The target value for the output layer

Δ : The different in current and new value for the next iteration

μ : The learning rate

α : The momentum factor

i : The number of iteration or epochs

E : Threshold error

1) Initializing Phase: all network weights are initialized. The learning rate, the momentum factor, the threshold error and the number of iteration are all set. Usually, the learning rate and the momentum factor are assigned small positive values (0.05-0.1). The number of iteration is in few hundreds (250-500), whereas the threshold error is set to a very small positive value.

2) Forward Pass Phase: Inputs are assigned from the training data using certain patterns. The outputs of the hidden layer and the output layer are calculated as follows:

$$\overline{O}_k(i) = f\left(\sum_{k=1}^{K} \varpi_{xk}(i) I_x\right) \qquad (1)$$

and

$$O_y(i) = f\left(\sum_{k=1}^{K} \omega_{ky}(i) \overline{O}_k(i)\right) \qquad (2)$$

where

$$f(x) = \frac{1}{1+e^{-x}} \qquad (3)$$

The desired targets T_y will be used to calculate the sum of squared system error E(i) for al inputs I_x as follows:

$$E(i) = \frac{1}{2}\sum_{y=1}^{Y}\left[t_y - O_y(i)\right]^2 \qquad (4)$$

If $E(i) \le E$, that indicates the algorithm is complete and the network has converged, for as long as $E(i) > E$, then the iterations will continue and proceed to the next phase (Backward Pass) in order to recalculate the network weights.

3) Backward Pass Phase: In this phase, the changes of the network weights will be calculated in order to be used in the following iteration (i+1).

$$\Delta \omega_{ky}(i+1) = \mu \delta_y(i) \overline{O}_k(i) + \alpha \Delta \omega_{ky}(i) \tag{5}$$

$$\Delta \overline{\omega}_{xk}(i+1) = \mu \overline{\delta}_k(i) I_x + \alpha \Delta \overline{\omega}_{xk}(i) \tag{6}$$

where

$$\delta_y(i) = \left(t_y - O_y(i)\right) O_y(i) \left(1 - O_y(i)\right) \tag{7}$$

$$\overline{\delta}_k(i) = O_k(i) \left(1 - \overline{O}_k(i)\right) \sum_{y=1}^{Y} \delta_y(i) \omega_{ky}(i) \tag{8}$$

After the changes of weights are calculated, the weights will be updated for the next iteration by:

$$\omega_{ky}(i+1) = \omega_{ky}(i) + \Delta \omega_{ky}(i+1) \tag{9}$$

$$\overline{\omega}_{xk}(i+1) = \overline{\omega}_{xk}(i) + \Delta \overline{\omega}_{xk}(i+1) \tag{10}$$

3.3 Polynomial Classifiers

Being known as a higher order neural network, the polynomial classifiers structure is very similar to ANNs. However, the concept of the hidden layer in polynomial classifiers is replaced with polynomial expansion.

The structure of the polynomial classifiers in the training stage can be shown as follows:

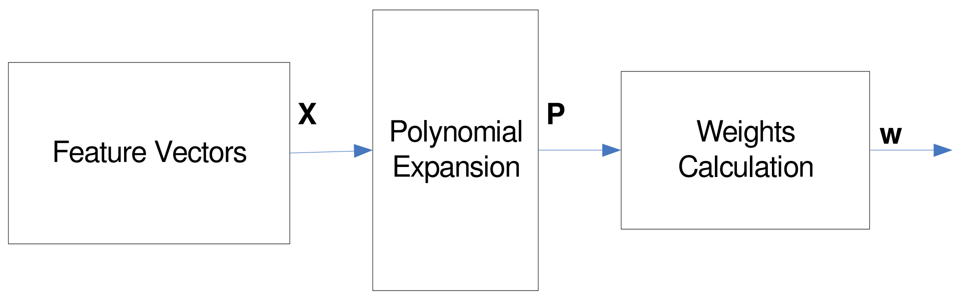


Figure 3.2 polynomial classifiers structure-training stage

When validating the polynomial classifiers, the desired outputs are multiplied by the calculated weights to produce the predicted outputs as shown in the figure below:

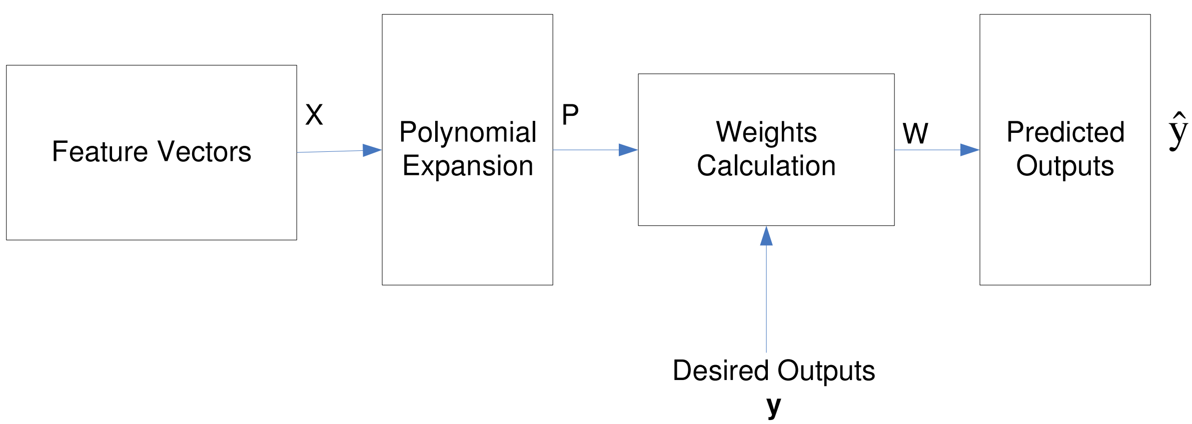


Figure 3.3 polynomial classifiers structure-validation stage

In order to illustrate the above structure, let's take the case of stock prediction. The feature vectors will be included in an input matrix formed for the historical prices of the previous five days for example. The desired outputs are the next day price.

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ x_2 & x_3 & x_4 & x_5 & x_6 \\ \cdots & & & & \\ & & & & \\ x_{n-5} & x_{n-4} & x_{n-3} & x_{n-2} & x_{n-1} \end{bmatrix} \Rightarrow \begin{bmatrix} x_6 \\ x_7 \\ .. \\ \\ x_n \end{bmatrix}$$

X-Matrix **y**-Vector

Therefore, the structure of the polynomial classifier will be as follows:

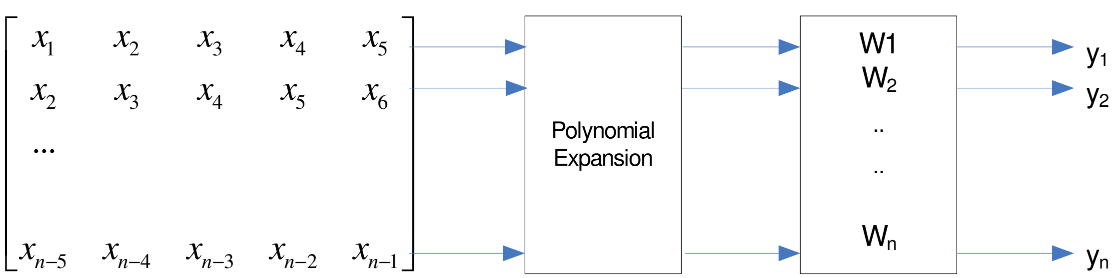


Figure 3.4 polynomial classifier structure for stock prediction

The polynomial expansion is represented by P(x), which is a vector of polynomial basis function up to order K. The form of P(x) used here is x_{i1}, x_{i2}, … $x_{i\ n}$, and therefore, a two element (x_1,x_2) second order feature vector **p**(x) would have the following expansion:

$$\mathbf{p}(\mathrm{x}) = \begin{bmatrix} 1 & x_1 & x_2 & {x_1}^2 & {x_2}^2 & x_1 x_2 \end{bmatrix} \tag{11}$$

And a three elements second order **p**(x) would be:

$$\mathbf{p}(\mathrm{x}) = \begin{bmatrix} 1 & x_1 & x_2 & x_3 & x_1 x_2 & x_1 x_3 & x_2 x_3 & {x_1}^2 & {x_2}^2 & {x_3}^2 \end{bmatrix} \tag{12}$$

So the feature vectors in the input matrix **X** will be expanded as:

$$\mathbf{P_i} = \begin{bmatrix} p(x_{i,1}) & p(x_{i,2}) & \ldots & p(x_{i,N_i}) \end{bmatrix} \tag{13}$$

Now, the output vector **Y** can be expressed as:

$$\mathbf{P.W = Y} \tag{14}$$

In order to map the expanded input $\mathbf{P_i}$, to the desired output **y**, the weights associated with polynomial classifier are calculated to produce the minimum error as follows:

$$w_i^{opt} = \underset{w_i}{\arg\min} \| p.w_i - y \| \tag{15}$$

This equation can be solved by method of normal equation [x,y], as follows.

For formula (14), multiply both sides by the transpose matrix of P:

$$\mathbf{P^T\, PW = P^T\, Y} \tag{16}$$

Let $\mathbf{R} = \mathbf{P}^{\mathbf{T}}\,\mathbf{P} \;\rightarrow\; \mathbf{RW} = \mathbf{Y}$ (17)

Multiply both sides by the inverse matrix of **R**:

$\mathbf{R}^{-1}\,\mathbf{RW} = \mathbf{R}^{-1}\mathbf{P}^{\mathbf{T}}\mathbf{Y} \;\rightarrow\; \mathbf{IW} = \mathbf{R}^{-1}\mathbf{P}^{\mathbf{T}}\mathbf{Y}$ (18)

Where **I** is the identity matrix. Therefore,

$\mathbf{W} = \mathbf{R}^{-1}\mathbf{P}^{\mathbf{T}}\mathbf{Y}$ (19)

From this definition of **w**, processing of a new testing feature vectors will be simplified to multiply the expanded form of these vectors by the already defined W, to obtain the desired output.

CHAPTER 4

4. METHODOLOGY

4.1 Data Source:

This study is based on historical prices of stocks listed in Dubai Financial Market. The historical prices were obtained from the Dubai Financial Market official website. DFM maintains and updates the records after each trading day.

4.2 Selected Securities:

Dubai Financial Market is an emerging market. It started with six listed companies, and in the first quarter of 2007, the company number fifty joined the market. The selection of securities to be used for the prediction models should comply with the following:

- Early listing date: The Company in selection should be listed at earlier stage to guarantee sufficient historical data.

- Active trading history: By convention, companies in the stock market are divided into two groups; active and inactive; according to the number of deals made on their stock in each trading day. Active securities assures higher number of deals made compared to inactive companies who only incur few number of deals in each trading day, and sometimes no deals would be made on these inactive companies. For active companies, the number of daily recorded deals started from at least couple of hundreds per day, up to couple of thousands in some extreme intensive trading days.

- Different sectors: In order to make sure that the models will not be data dependant and to verify the results are obtained at each step, it was of great importance to select companies that belong to different sectors like banking, investments, and real estate.

For this research, the selected securities are for:

- Emmar Properties
- Dubai Islamic Bank
- Dubai Investments

4.2.1 Emaar Properties:

Emaar Properties, the Dubai-based Public Joint Stock Company, was established in 23rd July, 1997 as one of the first companies to enter the new properties market in Dubai. Emaar's vision is to be one of the most valuable lifestyle developers in the world beyond real estate development. The company developed several real estate projects in its primary market of Dubai including Dubai Marina, Arabian Ranches, Emirates Hills, The Meadows, The Springs, The Greens, The Lakes, The Views and lately, its most ambitious project within the UAE, the AED 73 billion (US$20 billion) Downtown Burj Dubai development, which comprises Burj Dubai - stated to be the world's tallest tower when completed in 2008; along with The Dubai Mall - the world's largest entertainment and shopping mall. Emaar has internationally expanded and has joint ventures and projects across the region covering India, Egypt, Turkey, Morocco, Syria, Pakistan, Tunisia and Saudi Arabia.

Emaar was listed in Dubai Financial Market in 26th March, 2000 under real estate and construction sector. The authorized capital is 6,096,328,000.00 AED and the number of issued shares is 6,096,328,000.00 shares with 1.00

AED Par Value per Share. The current market closing price for Emaar is around 11.85 AED.

- The data used from Emaar stock historical prices, covers the period from 01-April-2000 to 16-March-2006 of daily closing prices (except Fridays).
- The total number of data points is 2176 point.

4.2.2 Dubai Islamic Bank

Dubai Islamic Bank was established in 12th March, 1975 and has the unique distinction of being the world's first fully-fledged Islamic bank, a pioneering institution that has combined the best of traditional Islamic values with the technology and innovation that characterize the best of modern banking. The bank has won so many awards locally and regionally for its outstanding performance and records in banking and finance. Although Islamic banking has become commonly dominant among local and international banks in UAE; Dubai Islamic Bank still leads the way, remaining true to its roots as a customer-centered organization where close personal service and understanding form the basis of all its relationships.

Dubai Islamic Bank was listed in Dubai Financial Market in 26th March, 2000 under banking sector. The authorized capital is 3,000,000,000.00 AED and the number of issued shares is 2,800,000,000.00 shares with 1.00 AED Par Value per Share. The current market closing price for Dubai Islamic Bank is around 7.07 AED.

- The data used from Dubai Islamic Bank stock historical prices, covers the period from 09-December-2000 to 23-November-2006 of daily closing prices (except Fridays).

- The total number of data points is 2176 point.

4.2.3 Dubai Investments

Dubai Investments (DI) is a world-class company that invests in viable and profitable entities in several business fields such as agriculture, telecommunications, finance, and real estate. The company has a very successful track record stretching back over 12 years since its establishment in 16th July, 1995, and has shown leadership in all fields of investment activities in the United Arab Emirates and the Middle East. With over 25,000 shareholders, and paid-up capital of DH 1.8 billion, Dubai Investments is the largest investment company listed on the UAE stock exchange. The company has grown rapidly and increased the number of its subsidiaries to 37 companies.

Dubai Investments was listed in Dubai Financial Market in 26th march, 2006 under the Investment and Financial Services sector. The authorized capital is 2,574,000,000.00 AED and the number of shares issued is 1,973,400,000.00 with 1.00 AED Par Value per Share. The current market closing price for Dubai Investments is around 4.29 AED.

- The data used from Dubai Investments stock historical prices, covers the period from 01-April-2000 to 16-March-2006 of daily closing prices (except Fridays).
- The total number of data points is 2176 point.

4.3 Data Setup:

The daily closing prices were selected as inputs to the prediction models. The data was arranged to begin with Saturday and end with Thursday. For Fridays, zeros were filled to adjust for the sequential order of days. In case of public holidays or unusual holidays, the previous day closing price was used.

It is important to point out that data was adjusted to accommodate shares split. For instance, Emaar share was split in the ratio of ten shares for every one share as agreed on the company extraordinary meeting held on 22nd June, 2004. Change of par value from AED 10.00 to AED 1.00 and accordingly, the number of issued shares becomes 2,650,000,000.

Apparently, the data sequence will incur a seemingly illogic drop after 22nd June, 2004. According to that, the data sequence will look like:

-	
-	
27-June-04	55.9
28-June-04	57.25
29-June-04	59.2
30-June-04	59.2
01-July-04	5.9
-	
03-July-04	5.73
04-July-04	5.57
-	
-	

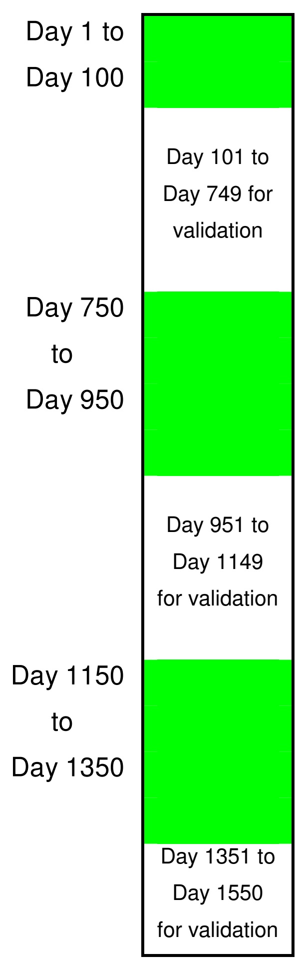
In order to adjust for such action, all previous daily closing prices were divided by 10, and the adjusted data vector sequence look like this:

-	
-	
27-June-04	5.59
28-June-04	5.72
29-June-04	5.92
30-June-04	5.92
01-July-04	5.9
-	
03-July-04	5.73
04-July-04	5.57
-	
-	

4.4 Learning Methods:

In order to monitor the learning progress and prediction accuracy, the neural network and polynomial classifiers prediction model would be trained on different training or learning methods. These training methods have been divided into three stages: one third of data, half of data, and two third of data. This means that in the first stage, one third of data will be used to train the model, and the remaining two third of data will be used for validation. In the second stage, half of data will be used for training and half for validation. The last stage of learning will take two third of data for training and the remaining third for validation.

The portion of data used for training in the three methods (third, half and two third) will consist of data points captured from different periods throughout the available past prices history. For instance, if presumably five hundred data points to be used for training, then these points will be seen as one block. Instead, several blocks from different periods will form the total five hundred points, by taking the first hundred from early data stream, the next two hundred from some middle data, and the last two hundred will be somewhere around most recent data, as shown in the below figure.



This technique was important to be used to assure that both prediction models will learn different market situations of price volatility, like bear market (period of decline), stable market (period of neither growth nor decline), and bull market (period of growth).

4.5 Prediction Modes:

Three modes will be used for predicting stock prices in this study.

1) Predicting the next day giving the five preceding days:

   In this mode, the inputs to the model are the closing prices of the previous five trading days, and the desired output is closing price for the following trading day.

$$
\begin{array}{llllll}
d_1 & d_2 & d_3 & d_4 & d_5 & \rightarrow d_6 \\
d_2 & d_3 & d_4 & d_5 & d_6 & \rightarrow d_7 \\
d_3 & d_4 & d_5 & d_6 & d_7 & \rightarrow d_8 \\
\ldots & & & & & \\
\ldots & & & & & \\
\ldots & & & & & \\
d_{n-5} & d_{n-4} & d_{n-3} & d_{n-2} & d_{n-1} & \rightarrow d_n
\end{array}
$$

2) Predicting the next three days giving the six preceding days:

   In this mode, the inputs to the model are the closing prices of the previous six trading days, and the desired output is closing prices for the following three trading days.

d_1 d_2 d_3 d_4 d_5 $d_6 \rightarrow d_7$ d_8 d_9

d_2 d_3 d_4 d_5 d_6 $d_7 \rightarrow d_8$ d_9 d_{10}

d_3 d_4 d_5 d_6 d_7 $d_8 \rightarrow d_9$ d_{10} d_{11}

...

...

...

d_{n-8} d_{n-7} d_{n-6} d_{n-5} d_{n-4} $d_{n-3} \rightarrow d_{n-2}$ d_{n-1} d_n

3) Predicting the next three days giving the twelve preceding days:

Similar to the previous mode, however in this mode, the inputs are doubled in size to cover the previous twelve trading days, while the desired output still be the closing prices for the following three trading days.

d_1 d_2 d_3 d_4 d_5 d_6 d_7 d_8 d_9 d_{10} d_{11} d_{12} $\rightarrow$ d_{13} d_{14} d_{15}

d_2 d_3 d_4 d_5 d_6 d_7 d_8 d_9 d_{10} d_{11} d_{12} d_{13} $\rightarrow$ d_{14} d_{15} d_{16}

d_3 d_4 d_5 d_6 d_7 d_8 d_9 d_{10} d_{11} d_{12} d_{13} d_{14} $\rightarrow$ d_{15} d_{16} d_{17}

…

…

d_{n-14} d_{n-13} d_{n-12} d_{n-11} …. d_{n-5} d_{n-4} d_{n-3} $\rightarrow$ d_{n-2} d_{n-1} d_n

The selection of these three modes was based on the trading week. The first mode was developed initially to predict the last day of the trading week, Thursday; given the previous five days of the week. Later on, this mode was generalized to predict any next day, given the previous five days regardless if the day to be predicted is Thursday or not. Similarly, mode 2 was for providing a complete week of trading to predict the first three days, Saturday to Monday; from the next week. It was modified later on to predict any subsequent three days, given the previous six days. Similarly for mode 3.

CHAPTER 5

5. PREDICTION MODEL IMPLEMENTATION

In this study, two techniques were developed for predicting stock prices. The first technique is Neural Networks model and the second technique is Polynomial Classifiers. The inputs in both techniques are similar, which are the historical stocks prices. Similarly, the desired output of each technique is the future prices. Apparently, the characteristics of the historical prices in the input vary according to what prediction mode is being used as explained in chapter 4. For example previous 6 days vs. previous 12 days form the input for prediction modes 2 and 3 sequentially.

5.1 Neural Network Prediction Technique:

As illustrated in chapter 3, the network topology is Backpropagation Feedforward with single hidden layer. MATLAB® software version 7.0.0 was used for neural networks system in order to construct the prediction model. The network error was set to zero, and the network maximum number of epochs for training was set to 500. The network momentum constant was set to a very small value. The model was used in all the three prediction modes. In each mode, the model was implemented in three different stages based on the amount of data for training. The start was third of available historical data to be used for training and the remaining two third to be used for validation. After that, the training data was increased to cover half of the available historical data while the remaining half was kept for validating the model. Finally, two third of data was utilized for training the model, and the remaining one third for validation. This way of implementing the model using three different amounts of data for training was applied in all the prediction modes as explained below.

5.1.1 Mode 1: Predicting the next day given the previous five days

As explained in chapter 4, the inputs to the prediction model will be arranged following the below pattern:

$$
\begin{array}{cccccccc}
d_1 & d_2 & d_3 & d_4 & d_5 & \rightarrow & d_6 \\
: & : & : & : & : & & : \\
d_{n-5} & d_{n-4} & d_{n-3} & d_{n-2} & d_{n-1} & \rightarrow & d_n
\end{array}
$$

This model will be used in three different methods according to the following:

5.1.1.1 Using one third of data for training and two third for validation

For all stocks used in this research (namely, Emaar, Dubai Islamic Bank and Dubai Investment), the amount of data points is equal and equal to 2,176 points of historical prices.

One third of data is around 679 points. As explained in chapter 4, the training data was taken from different sequences (like early days, few years back, and recent days) to make sure that the model was trained on several price movement schemes.

Therefore, the 679 points were selected as follows:

- Day 1 to day 280 (total 280 points).
- Day 841 to day 980 (total 140 points).
- Day 1261 to day 1400 (total 140 points).
- Day 1821 to day 1939 (total 119 points).

The remaining two third of data (around 1,421 points) is kept for validation.

5.1.1.2 Using half of data for training and half for validation

Half of data is around 1064 points, and will constitute of:

- Day 1 to day 350 (total 280 points).
- Day 806 to day 1050 (total 245 points).
- Day 1401 to day 1939 (total 539 points).

The remaining half of third of data (around 1,112 points) is kept for validation.

5.1.1.3 Using two third of data for training and one third for validation

Two third of data is around 1379 points, and will constitute of:

- Day 1 to day 280 (total 280 points).
- Day 421 to day 700 (total 280 points).
- Day 841 to day 1120 (total 280 points).
- Day 1261 to day 1540 (total 280 points).
- Day 1681 to day 1939 (total 259 points).

The remaining one third of data (797 points) is kept for validation.

The results of validation will be illustrated and discussed in the next chapter.

5.1.2 Mode 2: Predicting the next three days given the previous six days

Similarly, the inputs to the prediction model in this mode will be arranged following according to the below pattern:

$$\begin{array}{cccccccccc} d_1 & d_2 & d_3 & d_4 & d_5 & d_6 & \rightarrow & d_7 & d_8 & d_9 \\ : & : & : & : & : & : & & : & : & : \\ d_{n-8} & d_{n-7} & d_{n-6} & d_{n-5} & d_{n-4} & d_{n-3} & \rightarrow & d_{n-2} & d_{n-1} & d_n \end{array}$$

The same three training stages will be used, with exact data points input distribution explained in 5.1.1.1, 5.1.1.2, and 5.1.1.3. The results of validating this model will be shown in the next chapter.

5.1.3 Mode 3: Predicting the next three days given the previous twelve days

For this mode, more memory was provided to the data input vector to cover the previous 12 trading days as follows:

$$\begin{array}{ccccccccccc} d_1 & d_2 & d_3 & d_4 & \ldots & d_{10} & d_{11} & d_{12} & \rightarrow & d_{13} & d_{14} & d_{15} \\ : & : & : & : & & : & : & : & & : & : & : \\ d_{n-14} & d_{n-13} & d_{n-12} & d_{n-11} & \ldots & d_{n-5} & d_{n-4} & d_{n-3} & \rightarrow & d_{n-2} & d_{n-1} & d_n \end{array}$$

Again, the same three methods used in Mode 1 and Mode 2 will be applied here, to obtain a fair comparison among the three prediction modes.

5.2 Polynomial Classifier Prediction Technique

The example given in chapter 3 about polynomial classifiers structure for stock prediction was used in prediction mode 1, for predicting the next day given the previous five days. The same structure was developed and expanded to be used in mode 2 and mode 3. The inputs to the polynomial classifier technique developed here were the same used in the neural network technique, with the same structure, same training methods through out the three prediction modes.

One difference about the experiment using polynomial classifier technique compared to neural networks is that the model was used in two different configurations, one with first order classifier and the other with second order classifier. The results obtained from each classifier are recorded and compared to the results obtained by neural network technique, as will be shown in the next chapter.

CHAPTER 6

6. EXPERIMENTAL RESULTS AND DISCUSSION

6.1 Prediction Model:Using Neural Networks

The start was with Emaar stock. The neural network model was implemented as demonstrated in 5.1, for all the three prediction modes, and through all the three training methods. After showing the results obtained on Emaar stock, the same procedure was followed in both Dubai Islamic Bank stock and Dubai Investment stock.

6.1.1 Mode 1: Predicting the next trading day given the previous five days.

6.1.1.1 Emaar stock:

When implementing the neural network model on Emaar stock for predicting the closing price of the next trading day; the three training stages were used to show how the network adapted more training data. The criteria used to measure the prediction accuracy here (and in all the results obtained in this study) was the average error between the predicted closing prices and the actual closing prices, as follows:

Let,

P_n : Actual closing price for day n

$\hat{P}_n$: Predicted closing price for day n

N : Number of days in validation

ε_n : Prediction error for day n

Then the prediction error is:

$$\varepsilon_n = P_n - \widehat{P}_n \tag{20}$$

And the prediction error percentage is:

$$\varepsilon_n \% = \frac{\varepsilon_n}{p_n} x100 \tag{21}$$

But in practice, the predicted price could be lower or higher than the actual price. In order to overcome the negative sign, it's either the difference is in absolute term or squared. It was chosen in this study to absolute the error as follows:

$$|\bar{\varepsilon}| = \frac{1}{N}\sum_{n=1}^{N}|\varepsilon_n| \tag{22}$$

And

Average Absolute Error Percentage =

$$|\bar{\varepsilon}\%| = \frac{1}{N}\sum_{n=1}^{N}|\varepsilon_n \%| \tag{23}$$

For example, if the predicted price was 5.06 AED and the actual price was 4.94 AED, then the error is $4.94 - 5.06 = -0.12 AED$

And the error percentage

$$= \frac{-0.12}{4.94} x100 = -2.43\%$$

This error is actual error. The absolute error and the absolute percentage are being used in representing the average error of all the predicted days in validating the model performance.

The results obtained for predicting the next day giving the previous five days in all the three training methods are as follows:

Validation Criteria	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.066 AED (1.86%)	0.080 AED (1.56%)	0.152 AED (1.40%)
Error Standard Deviation	0.1144	0.1825	0.2987

Table 6.1 results of mode 1 - neural network prediction model on Emaar stock

From table 6.1, it is shown that the neural network could achieve a very small error averages in the three training stages. It is important to point out here that the average error obtained from validating the network when trained on only one third of data was not bad at all. It is also clear that the neural network could improve the prediction power by almost 25% when two third of data was used for training the network, compared to one third.

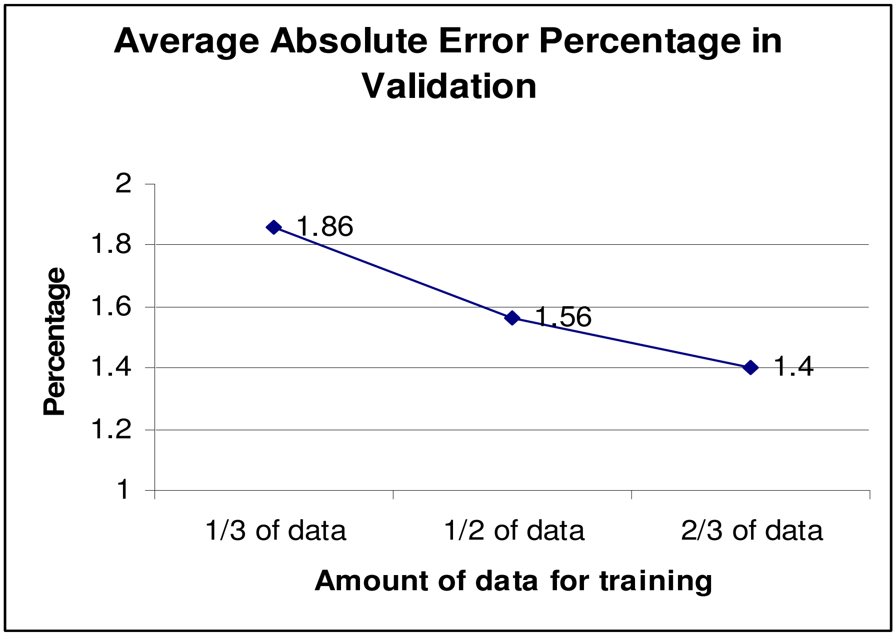


Figure 6.1: the improvement of neural network prediction-mode 1 in terms of average absolute error over the three training methods on Emaar stock.

When looking at the actual prediction error through all the validating days of the network that was trained on one third of data, the error found to have a normal distribution as shown below.

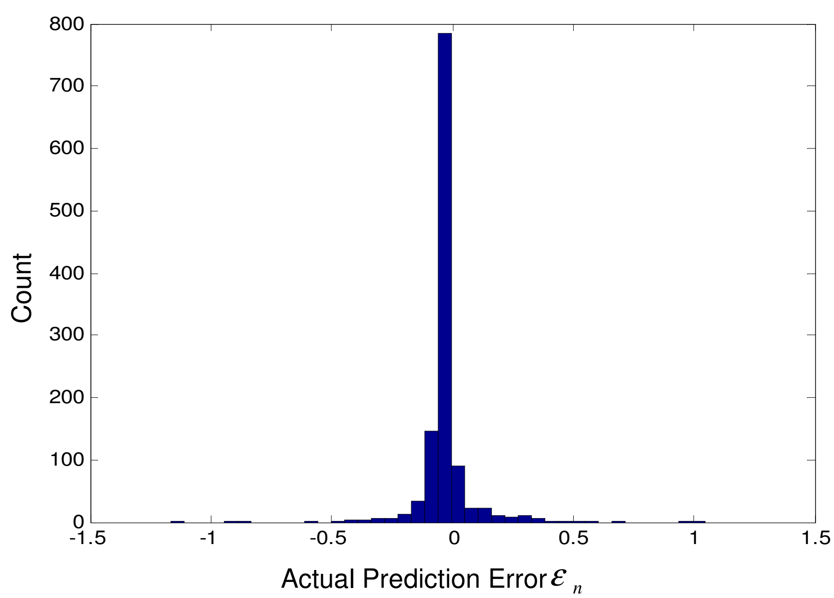


Figure 6.2: actual prediction error distribution of neural network trained on 1/3 of data – mode 1 on Emaar stock

Seeing that the actual prediction error tends to have normal distribution; another criteria was proposed to measure the prediction accuracy of the neural network. Three different error percentage intervals were defined as 1%, 5% and 10%, where the 1% interval contains all the prediction error percentages ($\varepsilon_n\%$) fall within - 0.01 to 0.01 from the actual price. Similarly, the 5% interval has all the prediction error percentages fall within - 0.05 to 0.05 from the actual price, and so on for the 10% error interval.

The results of three error intervals were as follows:

Error Interval	Percentage of actual errors included in the interval
1% (-0.01 to 0.01)	21.52%
5% (-0.05 to 0.05)	97.13%
10% (-0.10 to 0.10)	99.92%

Table 6.2: error intervals of prediction error of neural network trained on 1/3 of data – mode 1 on Emaar stock

Table 6.2 shows that more than fifth of the predicted prices fell within 1% from the actual prices, where almost all the predicted prices were within 10% from the actual prices.

Similarly, when using half of data for training the neural network; the actual error between predicted and real values was found to have normal distribution as shown below

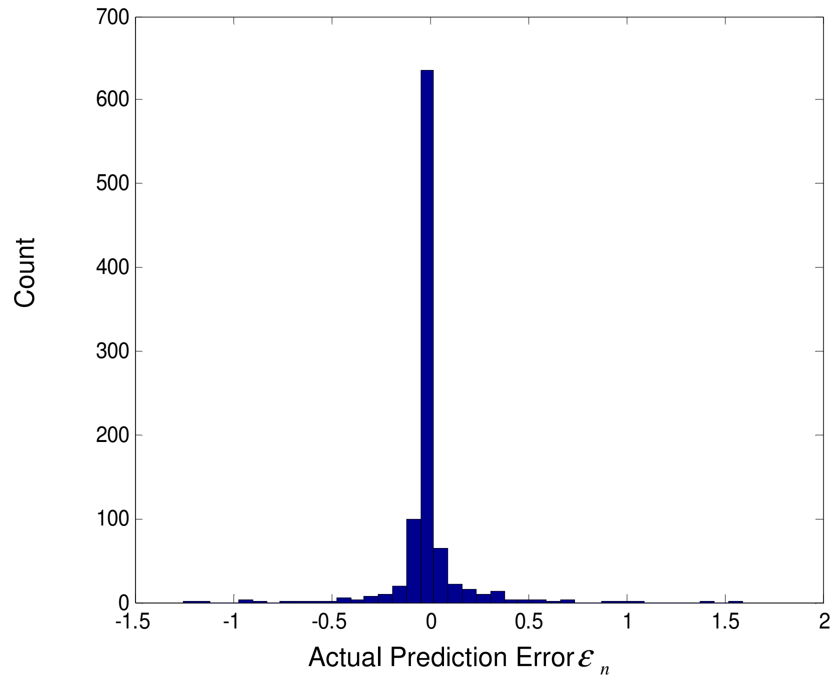


Figure 6.3: actual prediction error distribution of neural network trained on 1/2 of data – mode1 on Emaar stock

Here, more days were predicted accurately compared to the previous stage, as shown in the results of the error percentage confidence intervals:

Error Interval	Percentage of actual errors included in the interval
1% (-0.01 to 0.01)	36.38%
5% (-0.05 to 0.05)	97.02%
10% (-0.10 to 0.10)	99.68%

Table 6.3: error intervals of prediction error of neural network trained on 1/2 of data – mode 1 on Emaar stock

Finally, the error distribution for stage three (where two third of data was used for training) was also normal.

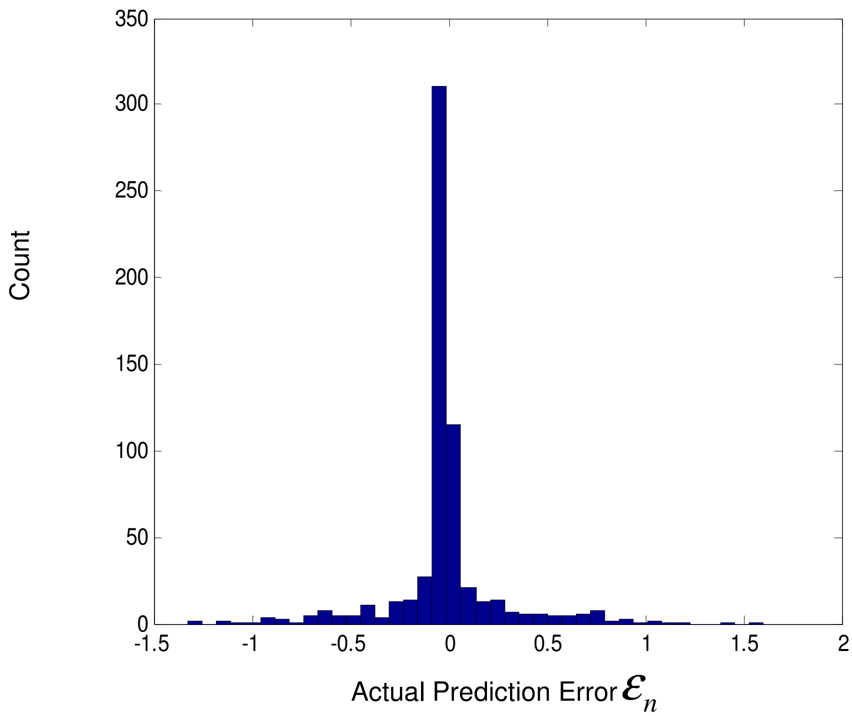


Figure 6.4: actual prediction error distribution of neural network trained on 2/3 of data – mode1 on Emaar stock

Training the network on two third of the data could enhance more the prediction accuracy, compared to the previous two stages as shown in the table below.

Error Interval	Percentage of actual errors included in the interval
1% (-0.01 to 0.01)	49.61%
5% (-0.05 to 0.05)	96.85%
10% (-0.10 to 0.10)	100.00%

Table 6.4: error intervals of prediction error of n on neural network trained on 2/3 of data – mode1 on Emaar stock

The results of all confidence intervals for the three training stages can be summarized as follows:

Error Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	21.52%	36.38%	49.61%
5% (-0.05 to 0.05)	97.13%	97.02%	96.85%
10% (-0.10 to 0.10)	99.92%	99.68%	100.00%

Table 6.5: error intervals of prediction error of neural network in all training methods – mode 1on Emaar stock

.

From all the above, its shown that whether the network was trained on one third, half or two third of the historical data; the actual error between predicted prices and the real prices tends to be normally distributed, whereas more days were predicted accurately when the network was trained on larger amount of data.

In order to verify the results obtained in this prediction mode of the neural network, the same analysis was applied on two more stocks, Dubai Islamic Bank and Dubai Investments.

6.1.1.2 Dubai Islamic Bank stock:

The results of applying the neural network model on Dubai Islamic Bank stock were as follows for the three training methods:

Validation Criteria	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.122 AED (1.25%)	0.113 AED (1.04%)	0.117 AED (1.18%)
Error Standard Deviation	0.3320	0.3431	0.2795

Table 6.6 results of mode 1 - neural network prediction model on DIB stock

Although the network didn't improve much when providing more data to be trained on, as shown in half and two third of data; however, the overall network performance on Dubai Islamic Bank was superior to its performance on Emaar stock. This was also proven when calculating the percentage error confidence intervals as shown in the below table.

Error Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	58.14%	72.59%	67.40%
5% (-0.05 to 0.05)	97.01%	97.04%	96.22%
10% (-0.10 to 0.10)	99.35%	99.26%	99.37%

Table 6.7: error intervals of prediction error of neural network in all training methods – mode1 on DIB stock

Although the 100% was not achieved in the last interval (90%CI); it was found that more than the 99% CI almost had 60% of the error percentages. Also the error distribution in all the three stages was found to be normally distributed, just as in Emaar stock.

6.1.1.3 Dubai Investments:

The results on applying the neural network model on Dubai Investments stock were as follows for the three training stages:

Validation Criteria	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.042 AED (1.31%)	0.039 AED (1.30%)	0.049 AED (1.34%)
Error Standard Deviation	0.0982	0.0903	0.1112

Table 6.8 results of mode 1 - neural network prediction model on Dubai Investments stock

Similar to Dubai Islamic Bank stock, the network performance was equivalent in all the training stages, but the overall performance was good as in the previous two stocks. The error was found normally distributed in the three training stages, and almost 60% of the predicted days were 1% more or less than the actual days, as shown in the below table.

Confidence Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	58.06%	57.46%	57.67%
5% (-0.05 to 0.05)	96.85%	97.25%	95.34%
10% (-0.10 to 0.10)	99.92%	99.89%	100.00%

Table 6.9: error intervals of prediction error of neural network in all training methods – mode1 on Dubai Investments stock

From all the above three stocks, Emaar, Dubai Islamic Bank, and Dubai Investment; it is shown that the network performed excellently regardless of the stock being tested. The results obtained in this mode were verified on the other two prediction modes included in this study (predicting the next three days given: previous 6 days & previous 12 days).

6.1.2 Mode 2: Predicting the next three trading day given the previous six days.

In this mode, the same network used in mode 1 was used to predict the closing prices of next three days. Instead of feeding the network with the previous five trading days, the network here was fed with the previous entire trading week (six days). The network was also trained and tested on the same three stages (third, half and two third of data) to detect if there has been any prediction improvements. The same error criteria used in 6.1.1 was used in this part to obtain equivalent analysis. The network will be tested on all the three stocks.

6.1.2.1 Emaar Stock

The results on applying the neural network model on Emaar stock for the next three days were as follows for the three training stages:

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.091 AED (2.14%)	0.171 AED (4.36%)	0.235 AED (6.34%)	0.080 AED (1.61%)	0.142 AED (3.17%)	0.189 AED (4.59%)	0.154 AED (1.41%)	0.240 AED (2.54%)	0.299 AED (3.50%)
Error Standard Deviation	0.1776	0.2922	0.3533	0.1823	0.2988	0.3637	0.3048	0.4867	0.5779

Table 6.10 results of mode 2 - neural network prediction model on Emaar stock

From the above table, it is shown that the average error of the third predicted day is almost 50% more than the second day and more than double the error in the first predicted day. The network performance was improved significantly when trained on more data, as seen in the average error of the third day (from 6.34% to 3.50%).

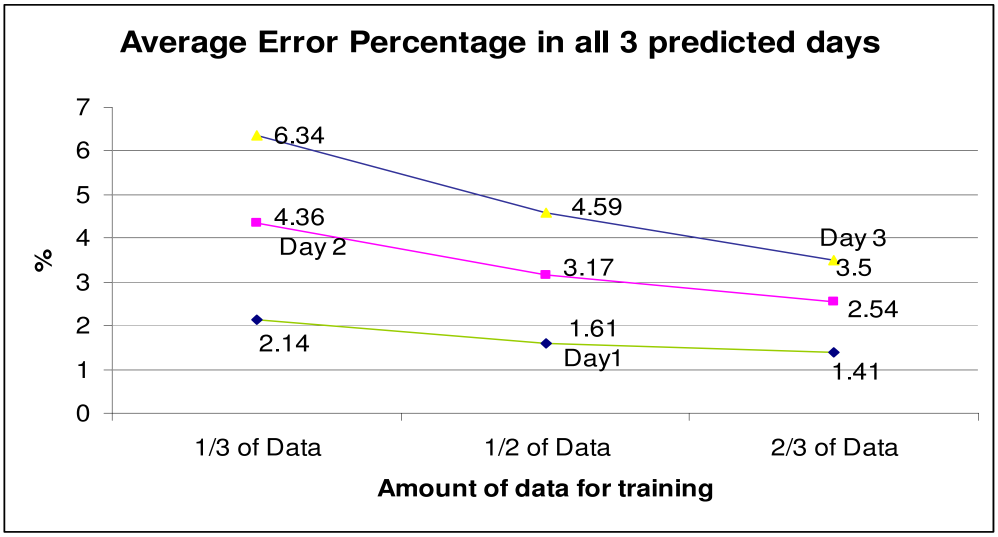


Figure 6.5: the improvement of neural network prediction accuracy-mode 2 in terms of average absolute error over the three training methods on Emaar stock.

When interpreting the results obtained in this mode for Emaar stock, it is seen that scoring 1.4% absolute average error for the first day, 2.5% for the second day and 3.5% for the third day, was certainly a very good achievement in terms of prediction accuracy. In other words, the neural network performance achieved in mode 1 seemed to continue in mode 2 for Emaar stock.

To have a better view of the network performance, the confidence intervals used in 6.1.1 were also used here for all the three days, as shown in the below table.

Confidence Interval	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	19.30%	7.82%	5.62%	33.97%	13.25%	8.44%	50.00%	23.06%	15.81%
5% (-0.05 to 0.05)	96.17%	65.47%	29.72%	97.01%	85.15%	62.82%	97.10%	89.03%	80.16%
10% (-0.10 to 0.10)	99.84%	97.80%	89.74%	99.89%	98.18%	95.62%	100.00%	98.39%	97.10%

Table 6.11: error intervals of prediction error of neural network in all training methods for the next three days – mode2 on Emaar stock

The table shows clearly how more days were predicted accurately when the network was trained on more data, as shown in the first day which was totally predicted within +/- 10% error, when the network was trained on two third of the data available. Error in all three predicted days was found to have a normal distribution.

Similar to mode 1; the results obtained in mode 2 on Emaar stock needed to be verified and checked on other stocks in order to conclude common statement of neural network performance.

6.1.2.2 Dubai Islamic Bank Stock

The results on applying the neural network model on Dubai Islamic Bank stock for the next three days were as follows for the three training stages.

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.123 AED (1.25%)	0.213 AED (2.22%)	0.286 AED (3.01%)	0.113 AED (1.04%)	0.190 AED (1.75%)	0.254 AED (2.33%)	0.118 AED (1.20%)	0.188 AED (1.91%)	0.242 AED (2.47%)
Error Standard Deviation	0.3339	0.5311	0.6966	0.3349	0.5341	0.6978	0.2818	0.4378	0.5565

Table 6.12 results of mode 2 - neural network prediction model on DIB stock

The network performed significantly better on Dubai Islamic Bank stock. Although the improvement among the three training stages were less noticeable, but the overall performance of the neural network was outstanding. That was also shown in the confidence intervals table below.

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	57.98%	32.00%	21.25%	72.01%	54.17%	39.53%	65.81%	45.85%	34.98%
5% (-0.05 to 0.05)	96.82%	90.80%	85.18%	97.01%	92.31%	88.35%	95.21%	90.73%	86.10%
10% (-0.10 to 0.10)	99.35%	98.05%	95.85%	99.15%	97.97%	96.15%	98.56%	96.96%	95.05%

Table 6.13: error intervals of prediction error of neural network in all training methods for the next three days – mode2 on DIB stock

The neural network applied to DIB stock in this mode could predict almost more than 60% of the first day and around 30% of the third day with +/- 0.01 error percentage.

6.1.2.3 Dubai Investments Stock

The results on applying the neural network model on Dubai Islamic Bank stock for the next three days were as follows for the three training stages:

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.042 AED (1.31%)	0.069 AED (2.22%)	0.088 AED (2.94%)	0.039 AED (1.31%)	0.065 AED (2.25%)	0.085 AED (3.04%)	0.049 AED (1.35%)	0.079 AED (2.20%)	0.101 AED (2.88%)
Error Standard Deviation	0.0991	0.1533	0.1910	0.0903	0.1393	0.1736	0.1123	0.1702	0.2106

Table 6.14 results of mode 2 - neural network prediction model on Dubai Investments stock

Similarly, the network here showed stability against the increased training data, but the performance was as good as the one obtained in 6.1.2.2 and 6.1.2.1, in both the average absolute error percentage and the confidence intervals as in the table below.

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	57.57%	24.92%	16.04%	55.45%	22.44%	14.64%	57.40%	27.00%	17.80%
5% (-0.05 to 0.05)	96.82%	92.10%	87.79%	97.22%	92.74%	88.46%	95.60%	89.80%	86.20%
10% (-0.10 to 0.10)	100.00%	98.70%	98.05%	99.89%	98.61%	97.65%	100.00%	98.60%	98.20%

Table 6.15: error intervals of prediction error of neural network in all training methods for the next three days – mode2 on Dubai Investment stock

6.1.3 Mode 3: Predicting the next three trading day given the previous twelve days.

In this mode, the same network used in mode 2 was used with double the amount of days being fed to the network. The purpose of expanding the input size was to explore the enhancement level in the network prediction power, if any. Similarly, the network was trained on the same three training stages, and the same analysis was applied on the three different stocks.

6.1.3.1 Emaar stock

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.094 AED (2.33%)	0.173 AED (4.68%)	0.232 AED (6.73%)	0.083 AED (1.73%)	0.145 AED (3.36%)	0.193 AED (4.81%)	0.161 AED (1.46%)	0.249 AED (2.61%)	0.308 AED (3.58%)
Error Standard Deviation	0.1797	0.2938	0.3546	0.1840	0.3002	0.3637	0.3165	0.5036	0.5967

Table 6.16: results of mode 3 - neural network prediction model on Emaar stock

From the above table, it's obvious that the network performed almost identically to the network used in mode 2, where six days were fed to the network. This was also shown clearly in confidence intervals table below:

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	17.61%	7.23%	5.98%	29.63%	12.09%	7.63%	46.10%	20.85%	15.59%
5% (-0.05 to 0.05)	95.10%	58.80%	28.41%	96.51%	83.55%	56.86%	96.95%	88.47%	78.47%
10% (-0.10 to 0.10)	99.75%	96.84%	84.72%	99.78%	98.04%	94.99%	99.83%	98.31%	96.78%

Table 6.17: error intervals of prediction error of neural network in all training methods for the next three days – mode3 on Emaar stock

The expectation before conducting this part was that the network will enhance the prediction accuracy when more days were provided as input. However, the obtained results show that no enhancement was achieved. The reason behind this could probably be that the network was saturated and there is no need for extra memory in the inputs.

When applying this technique on Dubai Islamic Bank stock and Dubai Investment stock, it was found that no enhancements were achieved on either stock. The results of Dubai Islamic Bank and Dubai Investment for this prediction mode (mode3) are attached in appendix (A).

6.2 Prediction Model: Using Polynomial Classifiers

As explained in 5.2, the Polynomial Classifiers model was applied for the stock price prediction. The inputs used in the neural network model were exactly employed to the polynomial classifiers model. The model was trained and tested on all the three stocks (Emaar, DIB and DI). The training stages were also implemented here to monitor any prediction enhancement of the polynomial classifiers.

It is important to mention that the polynomial classifiers model was used in two cases: first order polynomial classifier, and second order polynomial classifier. The reason for that is to study the performance of non-linear classifier compared to the linear one.

6.2.1 Mode 1: Predicting the next trading day given the previous five days.

6.2.1.1 Emaar Stock

1st order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.116 AED (1.37%)	0.113 AED (1.17%)	0.159 AED (1.28%)
Error Standard Deviation	0.2434	0.2615	0.3084

Table 6.18 results of mode 1 – first order polynomial classifier model on Emaar stock

The above table shows outstanding results obtained from the first order polynomial classifier. When analyzing the confidence intervals, it was found that first order polynomial classifier showed significant prediction accuracy as shown in the table below:

Error Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
99% (-0.01 to 0.01)	53.02%	61.74%	59.72%
95% (-0.05 to 0.05)	97.22%	97.59%	97.40%
90% (-0.10 to 0.10)	99.92%	100.00%	100.00%

Table 6.19: error intervals of prediction error of first order polynomial classifier in all training methods – mode1 on Emaar stock

2^{nd} order polynomial classifier:

Validation Criteria	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.155 AED (1.74%)	0.121 AED (1.32%)	0.168 AED (1.32%)
Error Standard Deviation	0.3008	0.2769	0.3251

Table 6.20: results of mode 1 – second order polynomial classifier model on Emaar stock

The second order polynomial classifier could also achieve good results. However, it was not as good as the results obtained from the first order polynomial classifier, especially when analyzing the confidence intervals.

Error Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	43.80%	52.81%	55.59%
5% (-0.05 to 0.05)	94.91%	97.25%	96.63%
10% (-0.10 to 0.10)	99.60%	99.89%	99.85%

Table 6.21: error intervals of prediction error of second order polynomial classifier in all training methods – mode1 on Emaar stock

The concept of introducing higher order polynomial classifier was to achieve higher prediction accuracy based on the fact that non-linear systems have better ability in capturing complex patterns such as stocks volatility. However, the results obtained didn't show any significant progress made in that aspect. It was shown that first order polynomial classifier could perform as good as the second order classifier. To verify this, both first order and second order polynomial classifiers of the same prediction mode were applied to another stock.

6.2.1.2 Dubai Islamic Bank Stock

The results on applying the polynomial classifiers model on Dubai Islamic Bank stock were as follows for the three stages:

1st order polynomial classifier:

Validation	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.121 AED (1.25%)	0.112 AED (1.04%)	0.117 AED (1.18%)
Error Standard Deviation	0.3320	0.3431	0.2795

Table 6.22: results of mode 1 – first order polynomial classifier model on DIB stock

Error Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	58.15%	72.59%	67.40%
5% (-0.05 to 0.05)	97.02%	97.04%	96.22%
10% (-0.10 to 0.10)	99.35%	99.26%	99.37%

Table 6.23: error intervals of prediction error of first order polynomial classifier in all training methods – mode1 on DIB stock

The results viewed in both tables are exactly identical to the results obtained on Dubai Islamic Bank stock using the neural network model. This is another evidence that first order polynomial classifier could perform well in stock price prediction.

2nd order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 2/3 for validation
Average Absolute Error	0.175 AED (1.38%)	0.168 AED (1.33%)	0.129 AED (1.24%)
Error Standard Deviation	0.6750	0.6509	0.3271

Table 6.24: results of mode 1 – second order polynomial classifier model on DIB stock

Error Interval	Percentage of actual errors included in the interval		
	1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	62.82%	67.62%	65.67%
5% (-0.05 to 0.05)	95.32%	95.45%	95.91%
10% (-0.10 to 0.10)	98.71%	98.62%	99.06%

Table 6.25: error intervals of prediction error of second order polynomial classifier in all training methods – mode1 on DIB stock

The performance of the second order polynomial classifier was within the same range, compared to the first order classifier.

The results of mode 1 of the neural network prediction model were confirmed when the results of mode 2 were presented. It was needed to verify the same thing with polynomial classifiers.

6.2.2 Mode 2: Predicting the next three trading days given the previous six days.

Moving further with the same analysis applied in the neural network model, both first order and second order polynomial classifiers were used in mode 2 as will be shown in this part.

6.2.2.1 Emaar stock

1st order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.119 AED (1.39%)	0.200 AED (2.51%)	0.251 AED (3.35%)	0.113 AED (1.15%)	0.184 AED (2.03%)	0.232 AED (2.78%)	0.162 AED (1.28%)	0.262 AED (2.18%)	0.326 AED (2.85%)
Error Standard Deviation	0.2570	0.4206	0.5081	0.2715	0.4395	0.5304	0.3231	0.5243	0.6337

Table 6.26: results of mode 2 – first order polynomial classifier prediction model on Emaar stock

From the results listed in the above table, it is shown that first order polynomial classifier confirmed its excellent performance in mode1. Although the prediction accuracy wasn't improved significantly when the classifier was trained on more amounts of data, but all the results obtained were better compared to the neural network results for the same prediction mode and stock. That is also confirmed in the confidence intervals table below.

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	52.81%	22.47%	13.64%	61.69%	33.45%	19.33%	59.09%	35.42%	20.69%
5% (-0.05 to 0.05)	96.31%	88.68%	81.54%	97.11%	92.82%	87.38%	96.55%	90.60%	85.27%
10% (-0.10 to 0.10)	99.76%	97.91%	95.83%	99.88%	98.73%	97.22%	99.84%	97.96%	95.92%

Table 6.27: error intervals of prediction error of first order polynomial classifier in all training methods for the next three days – mode2 on Emaar stock

2^{nd} order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.148 AED (1.68%)	0.285 AED (3.40%)	0.390 AED (4.68%)	0.119 AED (1.26%)	0.193 AED (2.28%)	0.241 AED (3.07%)	0.169 AED (1.31%)	0.275 AED (2.21%)	0.342 AED (2.77%)
Error Standard Deviation	0.2996	0.5407	0.7069	0.2823	0.4635	0.5546	0.3345	0.5510	0.6596

Table 6.28 results of mode 2 – second order polynomial classifier prediction model on Emaar stock

Apparently, no significant improvements have also been attained by using second order polynomial classifier compared to the first order classifier in either average absolute error percentage or error confidence intervals.

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	44.22%	24.96%	18.06%	56.37%	33.91%	18.98%	57.21%	42.63%	35.58%
5% (-0.05 to 0.05)	95.18%	77.21%	62.12%	97.34%	90.86%	82.52%	96.87%	88.71%	83.70%
10% (-0.10 to 0.10)	99.52%	95.59%	90.61%	99.77%	98.61%	96.64%	99.84%	97.96%	95.92%

Table 6.29: error intervals of prediction error of second order polynomial in all training methods for the next three days – mode2 on Emaar stock

6.2.2.2 Dubai Islamic Bank stock

The same procedure on examining the first order and second order polynomial classifiers was followed on Dubai Islamic Bank stock in order to verify the analysis of the previous parts.

1st order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.122 AED (1.25%)	0.213 AED (2.22%)	0.286 AED (3.01%)	0.113 AED (1.04%)	0.190 AED (1.75%)	0.254 AED (2.33%)	0.118 AED (1.20%)	0.188 AED (1.91%)	0.242 AED (2.47%)
Error Standard Deviation	0.3339	0.5311	0.6966	0.3349	0.5341	0.6978	0.2818	0.4378	0.5565

Table 6.30: results of mode 2 – first order polynomial classifier prediction model on DIB stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	57.98%	32.00%	21.25%	72.01%	54.17%	39.53%	65.81%	45.85%	34.98%
5% (-0.05 to 0.05)	96.82%	90.80%	85.18%	97.01%	92.31%	88.35%	95.21%	90.73%	86.10%
10% (-0.10 to 0.10)	99.35%	98.05%	95.85%	99.15%	97.97%	96.15%	98.56%	96.96%	95.05%

Table 6.31: error intervals of prediction error of first order polynomial classifier in all training methods for the next three days – mode2 on DIB stock

2nd order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.179 AED (1.41%)	0.321 AED (2.49%)	0.394 AED (3.14%)	0.169 AED (1.34%)	0.310 AED (2.38%)	0.387 AED (3.03%)	0.138 AED (1.32%)	0.222 AED (2.16%)	0.278 AED (2.74%)
Error Standard Deviation	0.7079	1.1869	1.3072	0.6907	1.1671	1.3256	0.3587	0.5914	0.7110

Table 6.32 results of mode 2 – second order polynomial classifier prediction model on DIB stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	62.54%	40.39%	29.48%	67.09%	42.74%	27.14%	63.23%	42.74%	34.03%
5% (-0.05 to 0.05)	95.44%	88.27%	84.12%	95.62%	88.78%	85.36%	96.13%	90.81%	86.13%
10% (-0.10 to 0.10)	98.70%	96.34%	94.54%	98.40%	96.37%	94.55%	98.71%	97.58%	96.29%

Table 6.33: error intervals of prediction error of second order polynomial classifier in all training methods for the next three days – mode2 on DIB stock

For the last prediction mode (mode 3), the same steps in 6.1.3 were applied on both first order and second order polynomial classifiers. The results for mode 3 also showed that first order polynomial classifier came up with almost the same outcomes of the neural network model, and the second order classifier performed good as well, but slightly less than first order classifier. For the sake of completion, the tables of error percentages and confidence intervals of mode 3 are attached in appendix B.

The analysis conducted on polynomial classifiers prediction model for Emaar stock in mode 1 and mode 2, and for Dubai Islamic Bank stock in mode 1 and mode 2, clearly showed that first order polynomial classifier could predict as good as the neural network (if not slightly better), while the second order classifier didn't demonstrate significant enhancement to the prediction accuracy.

6.3 Recent Updates.

By the time this thesis was completed, couple of recent experiments was made in order to verify the results obtained earlier. The idea was to use different securities and check on the results obtained.

ARAMEX and Union Properties stocks were selected. The trading history of ARAMEX was two years long (as the company was listed in DFM in 2005) while Union Properties had longer trading history that covers more than six years.

When applying mode 1 to ARAMEX stock using neural network, it was found out that the average absolute error was slightly higher than the error in the three selected securities, as shown in the table below:

Validation Criteria	Company	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 1/3 for validation
Average Absolute Error	ARAMEX	2.11%	2.01%	1.70%

Table 6.34 results of mode 1 – Neural network model on ARAMEX stock

Such results were expected, because of the short trading history available compared to the case of previous stocks. Therefore, the results of Union Properties stock were expected to be close to the results of Emaar, DIB, and Dubai Investments. The table below shows that clearly.

Validation Criteria	Company	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 1/3 for validation
Average Absolute Error	Union Properties	1.26%	1.09%	1.21%

Table 6.35 results of mode 1 – Neural network model on Union Properties stock

In order to have a broad overview of the results obtained in all the stocks in this study, the following table shows the results obtained on each stock using neural network for predicting the closing price of the next trading day (mode 1).

Mode 1: Predicting the next day given the previous five days, using Neural Networks				
Validation Criteria	Company	1/3 of data for training & 2/3 for validation	1/2 of data for training & 1/2 for validation	2/3 of data for training & 1/3 for validation
Average Absolute Error	Emaar	1.86%	1.56%	1.40%
	DIB	1.25%	1.04%	1.18%
	Dubai Investments	1.31%	1.30%	1.34%
	ARAMEX	2.11%	2.01%	1.70%
	Union Properties	1.26%	1.09%	1.21%

Table 6.36 results of mode 1 – Comparison of neural network model on the five stocks

The previous table had a comparison in terms of the average absolute error, while the below table will have a comparison with regards to the error intervals, as follows:

Error Interval	Company	Percentage of actual errors included in the interval		
		1/3 of data for training	1/2 of data for training	2/3 of data for training
1% (-0.01 to 0.01)	Emaar	21.52%	36.38%	49.61%
	DIB	58.14%	72.59%	67.40%
	Dubai Investments	58.06%	57.46%	57.67%
	ARAMEX	34.55%	38.46%	44.19%
	Union Properties	67.10%	70.16%	68.66%
5% (-0.05 to 0.05)	Emaar	97.13%	97.02%	96.85%
	DIB	97.01%	97.04%	96.22%
	Dubai Investments	96.85%	97.25%	95.34%
	ARAMEX	89.66%	92.13%	92.44%
	Union Properties	95.32%	96.19%	95.91%

Table 6.37 comparison of error intervals of neural network prediction model - Mode 1 on all the five stocks.

In addition, ARAMEX and Union Properties stocks were also used in mode 2, to verify the results obtained by mode 1 on these securities. Here, only one training method was chosen, which is half of data used for training and the other half used for validation. The results on the average absolute error and error intervals are listed in the below tables.

Validation Criteria	Company	Mode 2: Predicting the next three days given the previous six days, using Neural Networks (training on ½ of data only)		
		Day 1	Day 2	Day 3
Average Absolute Error	Emaar	1.61%	3.17%	4.59%
	DIB	1.04%	1.75%	2.33%
	Dubai Investments	1.31%	2.25%	3.04%
	ARAMEX	1.97%	3.08%	3.77%
	Union Properties	1.14%	1.78%	2.28%

Table 6.38 results of mode 2 – Comparison of neural network model on the five stocks

Error Interval	Company	½ of Data for training and ½ for validation		
		Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	Emaar	33.97%	13.25%	8.44%
	DIB	72.01%	54.17%	39.53%
	Dubai Investments	55.45%	22.44%	14.64%
	ARAMEX	37.98%	20.16%	21.71%
	Union Properties	68.91%	61.54%	54.38%
5% (-0.05 to 0.05)	Emaar	97.01%	85.15%	62.82%
	DIB	97.01%	92.31%	88.35%
	Dubai Investments	97.22%	92.74%	88.46%
	ARAMEX	93.02%	81.78%	70.93%
	Union Properties	95.94%	89.53%	86.75%

Table 6.39 comparison of error intervals of neural network prediction model - Mode 2 on all the five stocks.

CHAPTER 7

7. CONCLUSION, LIMITATIONS AND FUTURE WORK

Conclusion

Predicting the stock future prices was the aim of this study. The study was conducted on Dubai Financial Market as an emerging market, and the focus was on the market leading stocks, like Emaar Properties stock. There were two prediction models developed in this study. The first model was developed with the famous back propagation feed forward neural network. The second model was developed with polynomial classifiers, as a first time application for PCs to be used in stock prices prediction. The inputs to both models were identical, and both models were trained and tested on the same data.

In general, both models achieved outstanding results in terms of average error percentage and prediction accuracy. Both models did score around 1.5% average error of the next predicted day, 2.5% average error on the second predicted day, and around 4% average error in the third predicted day. The prediction accuracy of the two models was certainly remarkable , where around 60% of the predicted prices of the first day, 50% of the predicted prices of the second day, and 35% of the predicted prices of the third day, were all within -1% to 1% of the actual prices of the three days.

When comparing the neural network and polynomial classifiers prediction models, it was found that first order polynomial classifier performed slightly better or as good as the neural network. Whereas the second order polynomial classifier could barely achieve similar results on the stocks that were used in this study. Further work can be done using other stocks in similar emerging markets and mature markets, to verify the same conclusion.

Limitations

It is very important to address the limitations of this study, in order to better understand the scope and the unique conditions of this study.

- This study was conducted on Dubai Financial Market. The DFM is a very emerging market and has been established only seven years ago. Just like the case with any emerging market; the market index in the first few years was pretty much stable in terms of prices volatility. This is due to the limited number of listed securities, listed brokers, and investors in the market at that stage. Apparently, and at a later stage, when the market has created more awareness among different types of investors, and when more securities were listed in the market; DFM has experienced more active prices movement (inclining and declining) and some stocks prices have scored ten times higher than the original listing price during the market incline phase.

- This study was applied to three companies only (Emaar, DIB, and Dubai Investments). Each one of these securities is considered a leading stock in its sector (real estate, banks, and investments). The amount of daily trades made on each of these securities is relatively high, compared to other stocks at the same sector. This implies that taking any other security, and especially if it is an inactive one, may not lead to similar results obtained by the three selected securities. Therefore, the results of this study can be generalized to other companies within the same market or other markets. The scope of this study could be broadened to include companies in this DFM and/or other security markets in other GCC region and to make comparisons among companies across firms.

- Each of the three selected companies was listed at the first year of establishing the DFM. That could guarantee 6 years of historical prices, which is considered the maximum possible historical data available to train the intelligent system on. Most of the other listed securities have between 2 to 4 years of historical prices, and only few other securities have a bit longer history. Obviously, training the system on a shorter period will result in different outcomes, compared to the ones obtained by the three selected stocks. The results of this study are limited to the time period selected. The models tested may behave differently in other time periods.

Future Work

Throughout the progress of this study and throughout the analysis conducted on the results obtained by the two prediction models; there are several enhancements to be added to this study:

1. This study was entirely based on historical prices of the selected stocks, which can be classified as technical analysis. Initially, it was planned to include other inputs in this work, specifically factors from fundamental analysis that focus on the selected companies and related ratios. Although there are several studies showing that fundamental analysis can be useful in stocks prediction when combined with technical analysis, however, the results obtained by just using the historical prices were good enough to be recorded and addressed.

2. There are other factors in technical analysis that can be also included for future enhancement. For instance, high, low, open and close prices of each trading day can be used as inputs to both models. Also, the volume of each trading day can be further utilized.

3. In recent studies on stocks prediction, artificial systems are being hybridized. Usually two artificial systems are combined to eliminate certain limitation in each system, and to improve the overall prediction power. For example, the

neural network developed here didn't show any improvement when the input was doubled from six days to twelve days, whereas some other artificial systems could utilize the expansion of input vector to better predict the future prices. Some recent artificial systems that have been combined lately with neural networks are Genetic Algorithms (GAs) and Space Vector Machines (SVMs).

4. The output of both prediction models is the future price. This output can be commercialized in order to have practical usability. For example, the output can be designed to provide a buy or sell signal, based on the price tendency. There are several neural network models that were developed to provide such signal where trading strategies can use that to generate certain profits.

REFERENCES

[1] Hyun-jung Kim, Kyung-shik Shin, A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets, Applied Soft Computing, 7, 569-576, 2007.

[2] Tae Hyup Roh, Forecasting the volatility of stock price index, Expert Systems with Applications (2006), doi: 10.1016/j.eswa.2006.08.001.

[3] Guoqing Zhang, B. Eddy Patuwo, Michael Y. Hu, Forecasting with artificial neural networks: the state of the art, International Journal of Forecasting, 14, 35-62, 1998.

[4] Tsang, P.M., et al., Design and implementation of NN5 for Hong Kong stock price forecasting, Engineering Applications of Artificial Intelligence (2007), doi: 10.1016/j.engappai.2006.10.002.

[5] www.dfm.ae

[6] G. Armano, M. Marchesi, A. Murru, A hybrid genetic-neural architecture for stock indexes forecasting, Information Sciences, 170, 3-33, 2005.

[7] An-Sing Chen, Mark T. Leung, Hazem Daouk, Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index, Computers & Operations Research, 30, 901-923, 2003.

[8] Cheng-Lin Liu, Hiroshi Sako, Class-specific feature polynomial classifier for pattern classification and its application to handwritten numeral recognition, Pattern recognition, 39, 669-681, 2006.

[9] Qing Caoa, Karyl B. Leggioa, Marc J. Schniederjansb, A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market, Computer & Operations Research, 32, 2499-2512, 2005.

[10] Yousef Al-Assaf, Hany El Kadi, Fatigue life prediction of composite material using polynomial classifiers and recurrent neural networks, Composite Structure, 77, 561-569, 2007.

[11] Jar-Long Wang, Shu-Hui Chan, Stock market trading rule discovery using pattern recognition and technical analysis, Expert Systems with Applications, 33, 304-315, 2007.

[12] S. P. Kothari, Capital market research in accounting, Journal of Accounting and Economics, 31, 105-231, 2001.

[13] Ping-Feng Pai, Chih-Sheng Lin, A hybrid ARIMA and support vector machines model in stock price forecasting, Omega, 33, 497-505, 2005.

[14] www.emaar.ae

[15] Tong-Seng Quah, Bobby Srinivasan, Improving returns on stock investment through neural network selection, Expert Systems with Applications, 17, 295-301, 1999.

[16] Khaled Assaleh, M. Al-Rousan, Recognition of Arabic Sign Language Alphabet Using Polynomial Classifiers, EURASIP Jornal of Applied Signal Processing, 13, 2136-2145, 2005.

[17] K. T. Assaleh, W. M. Campbell, Speaker Identification using a Polynomial-Based Classifier, Fifth International Symposium on Signal Processing and its Applications ISSPA, Brisbane, Australia, 22-25 August, 1999.

[18] William M. Campbell, Khaled T. Assaleh, Charles C. Broun, Speaker Recognition With Polynomial Classifiers, IEEE transactions on speech and audio processing, Vol. 10, No. 4, May 2002.

[19] www.dib.ae

[20] www.dubaiinvestments.com

[21] Kyoung-Jae Kim, Artificial neural networks with evolutionary instance selection for financial forecasting, Expert Systems with Applications, 30, 519-526, 2006.

[22] William Leigh, Ross Hightower, Naval Modani, Forecasting the New York stock exchange composite index with past price and interest rate on condition of volume spike, Expert Systems with Applications, 28, 1-8, 2005.

[23] Li-Chiu Chi, Tseng-Chung Tang, Artificial neural networks in reorganization outcome and investment of distressed firms: The Taiwanese case, Expert Systems with Applications, 29, 641-652, 2005.

[24] Serdar Yumlu, Fikret S. Gurgen, Nesrin Okay, A comparison of global, recurrent and smoothed-piecewise neural models for Istanbul stock exchange (ISE) prediction, Pattern Recognition Letters, 26, 2093-2031, 2005.

[25] Niall O'Connor, Michael G. Madden, A neural network approach to predicting stoch exchange movements using external factors, Knowledge-Based Systems, 19, 371-378, 2006.

APPENDIX A

RESULTS OF NEURAL NETWORK PREDICTION MODEL – MODE 3 ON DUBAI ISLAMIC BANK AND DUBAI INVESTMENTS STOCKS

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.131 AED (1.28%)	0.219 AED (2.24%)	0.291 AED (3.04%)	0.115 AED (1.05%)	0.196 AED (1.78%)	0.252 AED (2.32%)	0.117 AED (1.19%)	0.189 AED (1.92%)	0.245 AED (2.49%)
Error Standard Deviation	0.3349	0.5315	0.6969	0.3353	0.5352	0.6977	0.2819	0.4384	0.5568

Table A-1: results of mode 3 - neural network prediction model on DIB stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	56.12%	31.73%	21.07%	71.79%	52.98%	39.88%	63.32%	45.34%	34.54%
5% (-0.05 to 0.05)	96.23%	90.65%	84.38%	96.48%	92.2%	88.56%	95.8%	90.23%	86.65%
10% (-0.10 to 0.10)	99.11%	98.02%	94.56%	99.01%	97.72%	96.95%	97.92%	96.12%	95.76%

Table A-2: error intervals of prediction error of neural network in all training methods for the next three days – mode3 on DIB stock

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.048 AED (1.33%)	0.071 AED (2.24%)	0.085 AED (2.93%)	0.038 AED (1.31%)	0.061 AED (2.22%)	0.079 AED (2.94%)	0.041 AED (1.3%)	0.0793 AED (2.16%)	0.103AED (2.89%)
Error Standard Deviation	0.0994	0.1525	0.1903	0.0989	0.1381	0.1723	0.1115	0.1690	0.2109

Table A-3 results of mode 3 - neural network prediction model on Dubai Investments stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	57.38%	25.55%	16.67%	55.85%	22.72%	15.74%	58.52%	29.15%	17.95%
5% (-0.05 to 0.05)	96.52%	92.9%	87.99%	97.84%	93.67%	88.96%	94.97%	90.09%	86.70%
10% (-0.10 to 0.10)	99.96%	98.93%	98.73%	99.93%	98.93%	97.85%	100.00%	98.87%	98.60%

Table A-4: error intervals of prediction error of neural network in all training methods for the next three days – mode3 on Dubai Investment stock.

APPENDIX B

RESULTS OF POLYNOMIAL CLASSIFIERS PREDICTION MODEL – MODE 3 ON EMAAR AND DUBAI ISLAMIC BANK STOCKS

1^{st} order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			1/2 of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.116 AED (1.38%)	0.205 AED (2.52%)	0.243 AED (3.32%)	0.109 AED (1.12%)	0.182 AED (2.01%)	0.238 AED (2.79%)	0.166 AED (1.32%)	0.269 AED (2.26%)	0.331 AED (2.91%)
Error Standard Deviation	0.2563	0.4208	0.5073	0.271	0.4392	0.5308	0.3237	0.5251	0.6341

Table B-1: results of mode 3 – first order polynomial classifier prediction model on Emaar stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	52.23%	21.62%	13.95%	62.56%	33.45%	20.41%	59.85%	35.87%	21.36%
5% (-0.05 to 0.05)	96.11%	88.12%	81.82%	97.73%	92.82%	88.40%	96.98%	91.74%	85.79%
10% (-0.10 to 0.10)	99.41%	97.38%	95.29%	99.22%	98.73%	98.1%	99.56%	98.11%	96.43%

Table B-2: error intervals of prediction error of first order polynomial classifier in all training methods for the next three days – mode3 on Emaar stock

2nd order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			½ of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.152 AED (1.71%)	0.292 AED (3.45%)	0.397 AED (4.73%)	0.124 AED (1.3%)	0.202 AED (2.42%)	0.277 AED (3.23%)	0.176 AED (1.37%)	0.271 AED (2.17%)	0.334 AED (2.72%)
Error Standard Deviation	0.3012	0.5445	0.7091	0.2856	0.4687	0.5567	0.3368	0.5504	0.6565

Table B-3 results of mode 3 – second order polynomial classifier prediction model on Emaar stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	43.34%	24.29%	17.1%	55.62%	32.85%	18.22%	57.44%	45.66%	37.77%
5% (-0.05 to 0.05)	94.56%	76.31%	61.35%	96.13%	89.79%	81.49%	96.36%	89.88%	84.6%
10% (-0.10 to 0.10)	99.24%	95.11%	88.9%	99.29%	98.28%	96.3%	99.12%	98.41%	96.87%

Table B-4: error intervals of prediction error of second order polynomial in all training methods for the next three days – mode3 on Emaar stock

1st order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			½ of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.126 AED (1.29%)	0.219 AED (2.42%)	0.295 AED (3.11%)	0.121 AED (1.21%)	0.198 AED (1.82%)	0.262 AED (2.39%)	0.127 AED (1.26%)	0.195 AED (1.94%)	0.247 AED (2.49%)
Error Standard Deviation	0.3307	0.5374	0.7034	0.3414	0.5382	0.7048	0.2862	0.4392	0.5582

Table B-5: results of mode 3 – first order polynomial classifier prediction model on DIB stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	57.22%	32.15%	20/26%	70.15%	53.35%	39.28%	64.31%	46.25%	35.71%
5% (-0.05 to 0.05)	96.31%	90.32%	84.73%	97.07%	91.57%	88.65%	95.75%	91.05%	86.74%
10% (-0.10 to 0.10)	99.1%	97.56%	95.41%	99.05%	97.62%	95.52%	98.26%	95.62%	95.33%

Table B-6: error intervals of prediction error of first order polynomial classifier in all training methods for the next three days – mode3 on DIB stock

2nd order polynomial classifier

Validation Criteria	1/3 of data for training & 2/3 for validation			½ of data for training & 1/2 for validation			2/3 of data for training & 2/3 for validation		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
Average Absolute Error	0.122 AED (1.38%)	0.325 AED (2.51%)	0.397 AED (3.18%)	0.175 AED (1.38%)	0.341 AED (2.44%)	0.385 AED (2.98%)	0.141 AED (1.34%)	0.221 AED (2.15%)	0.283 AED (2.77%)
Error Standard Deviation	0.7085	1.1872	1.3093	0.6924	1.1693	1.3233	0.3593	0.591	0.7125

Table B-7 results of mode 3 – second order polynomial classifier prediction model on DIB stock

Error Interval	1/3 of Data (training)			½ of Data (training)			2/3 of Data (training)		
	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3	Day 1	Day 2	Day 3
1% (-0.01 to 0.01)	62.1%	40.17%	28.72%	67.44%	42.82%	29.52%	63.73%	42.16%	34.63%
5% (-0.05 to 0.05)	95.22%	87.37%	84.75%	95.97%	88.24%	85.87%	96.55%	90.26%	86.62%
10% (-0.10 to 0.10)	98.36%	95.42%	95.12%	98.57%	97.02%	94.79%	98.24%	97.27%	96.72%

Table B-8: error intervals of prediction error of second order polynomial classifier in all training methods for the next three days – mode3 on DIB stock

VITA

Saeed Mohammed Al-Salkhadi was born in Abu Dhabi, 1980. He was educated in local public schools and graduated from the university of Sharjah in 2003.

Mr. Al-Salkhadi joined Dubai World Central as the project coordinator at the Dubai Logistics city.

He began his higher studies in 2004 when he enrolled in the American University of Sharjah as a student in the Engineering Systems Management program and graduated in June 2007