FUTURE SHARE PRICE PREDICTION AND ITS COMPARISON WITH DECISION TREE AND FEED FORWARD NEURAL NETWORK

[1]Nijiya Nazar, [2]Reji Thankachan

[1]PG Scholar, [2]Assistant Professor

[1] Department of Electronics and Communication Engineering,
College of Engineering Karunagappally, Kollam, Kerala, India,

Abstract : A lot of endeavor has gone in to the study of financial market and how the price fluctuate along time, adopting technical analysis method of aiming to predict future rates. Throughout this paper we studied the statistical properties of stock-market data and prediction done with the help of statistical properties. Quarterly worked stock market data are standardized by re-scaling each index and then normalizing the whole time- series. Plot of the kurtosis and skewness will help to study the characterization of stock market data. Standardized normalized data further given to different predictors then plot the output and compare..

***IndexTerms* - financial data; standardization ;normalization ;decision tree predictor ;neural network predictors.**

I. INTRODUCTION

Financial and economic development for past centuries was done without the reference of advanced statistical and mathematical modeling (and signal processing). The new ideas reveled from Fischer Black and Myron Scholes modified all the myths in financial area. The Black-Scholes model,[1]also known as the Black-Scholes-Merton model, The main idea behind model explains variation of price over time of financial instruments such as stocks. This model also determines the price distribution of a European call option. The mathematization financial models[2] places policy making in the hands of the experts and increases investors' trust. Judgments are straightly based on discernible, realistic and reproducible quantitative analysis rather than the ”good hunch” of a ”market seasoned veteran”.

To know about the relevance of signal processing in finance system modeling [3] it may first be rewarding to explore the concept of a signal itself. A signal is represent as a sequence of numerical data that varies with respect to an underlying independent variable, mostly time.

In contrast, the heart of technical analysis aims to predict the future market value of a business using historical financial data. This is precisely the type of task for which signal processing is suited because the quantity of historical data is often immense and the sheer objectivity demanded in calculations is scarcely different from that seen in electrical engineering applications.

Financial continuous time varying data often consists of long-term trends in the presence of high-frequency and wildly fluctuating noise. To removing noises, techniques like moving-average filters, and low-pass filters are usually employed, before moving to the actual processing of waveforms. Advanced methods based on adaptive Kalman filters[4] used for the removal of noise and extracting meaningful information. In this paper standardizing stock-market data, and relate it to the behavior of the market, observable as changes in stock-market indices. we introduce our normalization of the distribution of the standardized data, and analyze the evolution of the normalized distribution.

A coordinate system: its x-axis is time measured in quarters, with the index i; and normalized distributions Zk[i] are plotted on its y-axis, with origin mean and units variants. We can analyze this normalized distribution in terms of skewness and kurtosis. Skewness measures the asymmetry of a distribution, and kurtosis the sharpness of its peaks and the thickness of its tail. We can take skewness and kurtosis to be indicators of the presence of extreme values in a distribution. Since a global financial crisis initially affects a small number of markets, then we can expect to observe several peaks in these higher moments just before the crisis, followed by a decline as all markets are more equally affected. Especially in the middle of the crisis, we can spontaneously observe a significant dip in skewness and a sharp peak in kurtosis.Using neural network and decision tree predictors for future value comparison.

II. RELATED WORKS

Financial data[5] is behaves as random signals or stochastic signal, such data consist of high frequency and wildly fluctuating noise. So many different techniques are used for removing noise. Moving Average method[6] is one of the technique to get an overall idea of the trends in a data set and it's useful for forecasting long-term trends. Using this, we can calculate five year moving average a four year moving average, and so on. Stock market analysis use a fifty to two hundred days moving average to help to study trends in the stock market. Relative Strength Index is another technique has been used. Same above technique applies momentum indicator invented by a well-known Technical Analyst Mr. Wells Wider to note the similarities and dissimilarities in the magnitude of current loses and profits in between a time lapse to measure the speed and changes of price movement of a security. It is primary attempt is to identify overbought or oversold conditions in a trading of an asset. Moving Average Convergence Divergence[7] is a another momentum indicator shows the connection between two changing average of cost. It is calculated by subtracting the twenty six days exponential moving average from twelve days EMA. Fibonacci retracements[8] is created by taking two extreme points (major peak and through) on a stock chart and dividing the vertical distance by the key Fibonacci ratios of 23.6%,38.2%,50%,61.8% and 100

III. RESEARCH METHODOLOGY

A series of data-points catalogued in time order is called a time-series[9], it is a progression which is organized in a consecutive equally distributed points on time. Time-series is a discrete time-data. This kind of dates are mostly used in statistics, signal processing, pattern recognition, econometrics, mathematical, finance, weather forecasting, earthquake prediction ,electro encephalography, control engineering, astronomy communication engineering.

Analysis of time-series data[9] is done by time series analysis method which helps to extract meaningful statistics and other characteristics of data. Time series forecasting is the model to predict future values based on previously observed values. Regression analysis not called a “time-series analysis” because current values of one or more independent time series affect the current value of another time series. Time-series analysis mainly divided into frequency-domain methods and time-domain methods. For getting better prediction we introducing quarterly worked financial data is converted as a standardized normalized data.

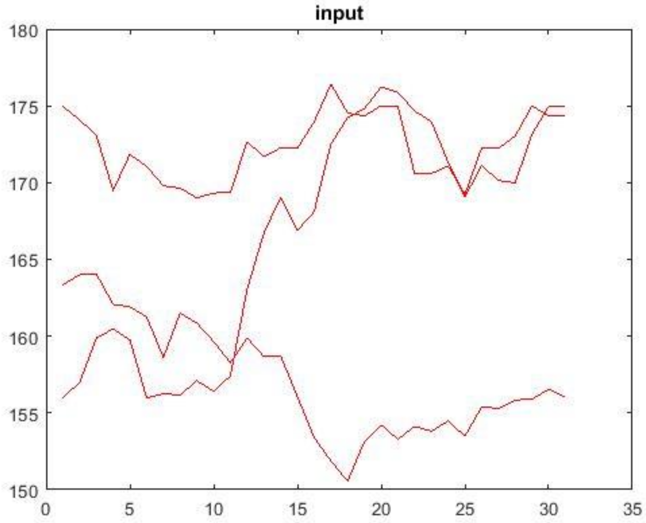

Fig.3.1 Financial Data

3.1 Standardization of data

Standardize the quarterly worked stock-market data by re-scaling the market indices[10]. . If $A_k = \{a_k[i]\}_{i\in N}$ is the time-series of re-scaled indices $a_k[i]$ of a stock market k, out of n markets in total (i.e. k = 1, 2, 3, ⋯ , n), then we define a new set of data $A_k^\dagger = \{a_k^\dagger[i]\}_{i\in N}$, where

$$a_k^\dagger[i] = \frac{a_k[i+1] - a_k[i]}{a_k[i]} \qquad (3.1)$$

Which is the proportional change in value over one step of the time-series .Then comparison is done with the terms of a_k. which provides a standardized measure of the movement of each market.

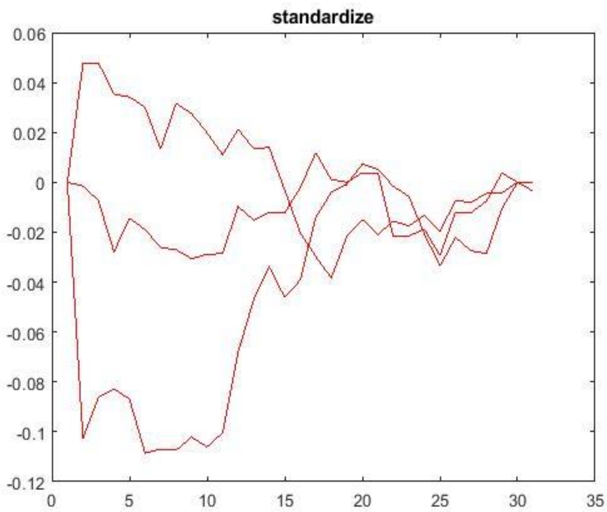

Fig.3.2 Standardized financial data out

3.2. Normalization of the standardized data

We now have n standardized data-sets $A_K^\dagger = \{a_k^\dagger[i]_{i\in m}\}$for 1 ≤ k ≤ n , and we now organize them into the following sequence:

$$N_i = \{Z_k[i] = \frac{a_k^\dagger[i] - m_i}{\sigma_i}\}_{k\in K} i \in N \qquad (3.2)$$

where K = { 1 , 2 , 3 , •••, n } and m i and σi respectively are the mean and standard deviation of { a ȷ [i] } k ∈ K at each time i .Thus{ N i } i ∈ N is a sequence of distributions. Introduce a coordinate system[6]: its x -axis is time measured in quarters, with the index i ; and normalized distributions Z k [i] are plotted on its y -axis, with origin m i and units σ i . We call this the standardized normalization , and it is plotted. This figure provides more precise and abundant information about the behavior of the stock market data.

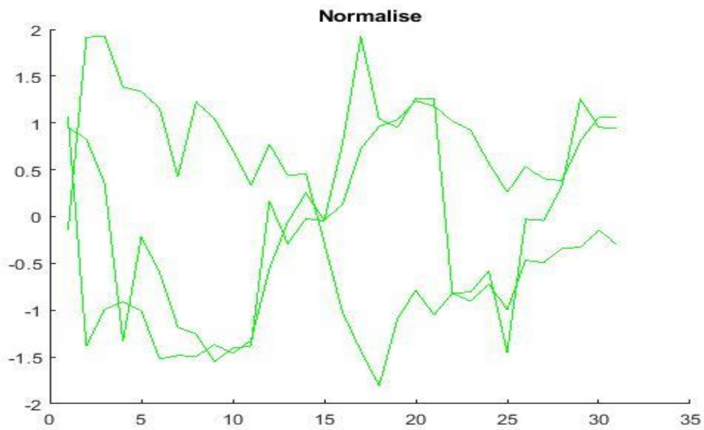

Fig.3.3 Normalization of the standardized data

We can analyze this normalized distribution in terms of skewness Skew $(Z_k[i]) = \frac{E((Z_k[i]-E(Z_k[i]))^3)}{V(Z_k[i])^{\frac{3}{2}}}$ (3.3) and kurtosis

Kurt $(Z_k[i]) = \frac{E((Z_k[i]-E(Z_k[i]))^4)}{V(Z_k[i])^2} - 3$ (3.4). Skewness measures the asymmetry of a distribution, and kurtosis the sharpness of its peaks and the thickness of its tail. We can take skewness and kurtosis to be indicators of the presence of extreme values in a distribution. After that the normalization of standardized data will given to the input of a predictors .Here we are using two type of predictors. Stock market data prediction technique is used for determine the company's financial instrument traded or stock future values in an exchange.

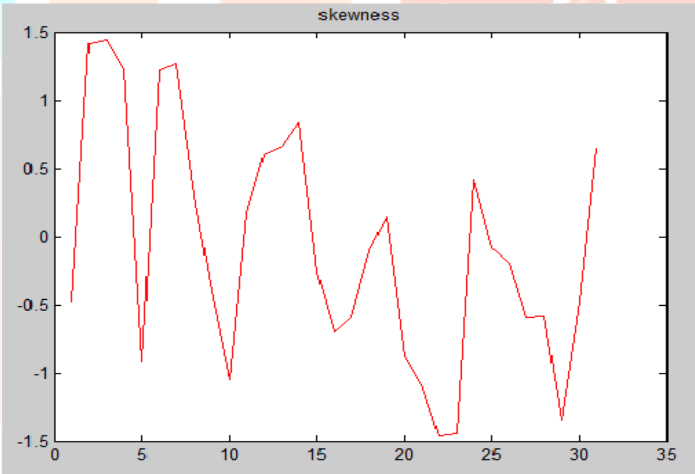

Fig.3.4 Skewness of normalized data.

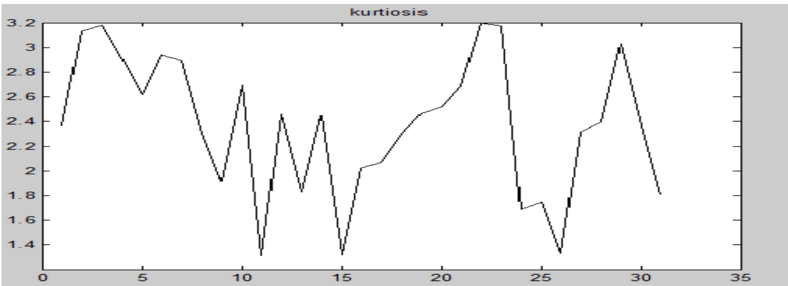

Fig.3.5.kurtiosis of normalized data.

A significant profit can be yield by successful prediction of stock's future price. Prediction methods mainly are in two ways, fundamental analysis and technical analysis. Data mining technology is got advanced in the advent of the digital computers. Stock market prediction has since progressed into the technological field. The most prominent technologies are artificial neural network and decision tree predictors. Normalized data is given in to the two predictors and compared with original predicted value and normalized predicted value after that finding the accuracy of a predictor with the help of out puts.

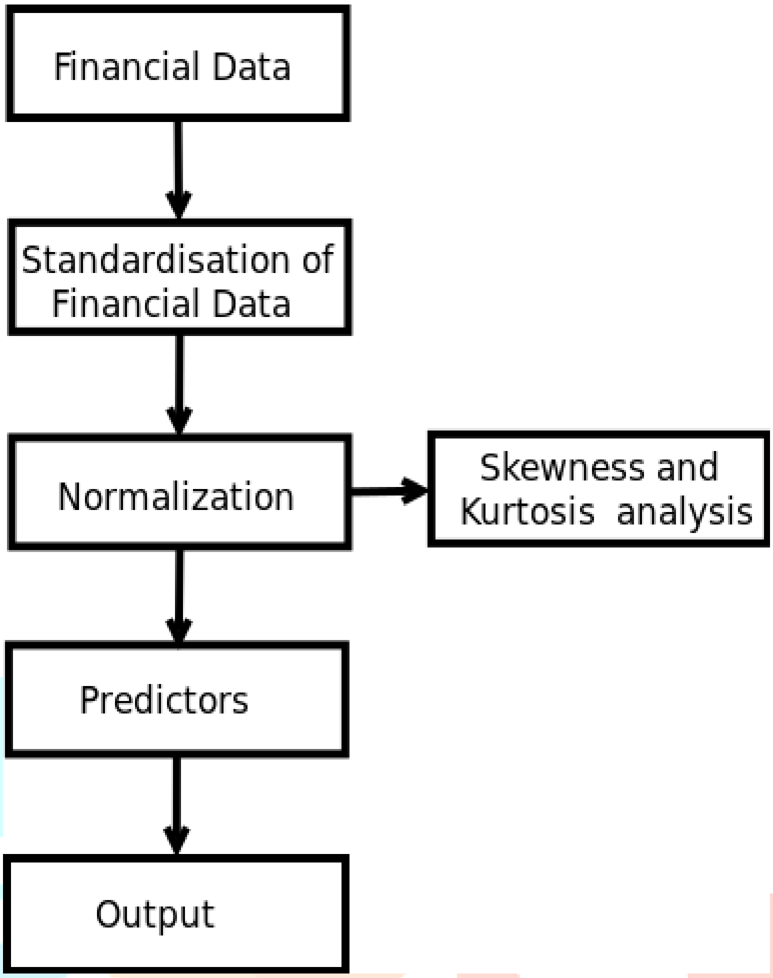

Fig.3.6 Flow chart of the Proposed Method

Neural network are a class of generalized non-linear non-parametric models encouraged by the studies of the human brain. The prominent advantage is that they can approximate any non-linear function to an arbitrary degree of accuracy with a suitable number of hidden units. Neural network can be used for prediction with various level of success. This was used to train the prior data expecting that it will discover concealed reliance and that it able to apply them for predicting the future. A time series is a sequence of vectors X_t = ($x_{t-n},\ldots..,x_{t-I},\ldots.,x_{t-2},x_{t-1}$) where x_{t-i} represents past value varies with time, such as stock price. The main aim is to predict future trends with the help of a past price patterns X_t. $\tilde{x}_t = F(x_{t-n},\ldots..,x_{t-I},\ldots.,x_{t-2},x_{t-1})$, here $F:R^n \rightarrow R$.

Where $\tilde{x}_t$ is a predictor value. $\tilde{x}_{t+1}$ can be predicted after $\tilde{x}_t$ is predicted based on the following function $\tilde{x}_{t+1} = F(x_{t-(n-1)},\ldots.., x_{t-i},\ldots..x_{t-2}, x_{t-1}, \tilde{x}_t)$.

Here we are using feed forward neural network. It is one of the simplest forms of ANN. Were the data or the input travel in the one direction, data passes through the input notes and exit on the output notes. This neural network may or may not have the hidden layers. This kind of neural network is responsive to the noisy data and easy to maintain.

Decision tree trend to be the method of choice for predictive modeling, because this method is very easy to understand and also very effective. The main process of decision tree is to split a data into smaller segments. In first stage is, to train the model and this is where the tree build, optimized and tested with the help of an existing compilation of data, in the second stage unknown outcomes is predicted using the model.

IV. RESULTS AND DISCUSSION

Experiments were conducted on APPLE INC Company's one year financial data sheet and are done on MATLAB platform. From the data sheet, choose close value prices and its plotted as three signals. Standardization of these signals will change the time index of signals and further undergo normalization process. With the help of normalized standardized data, future value is predicted with neural network and decision tree.

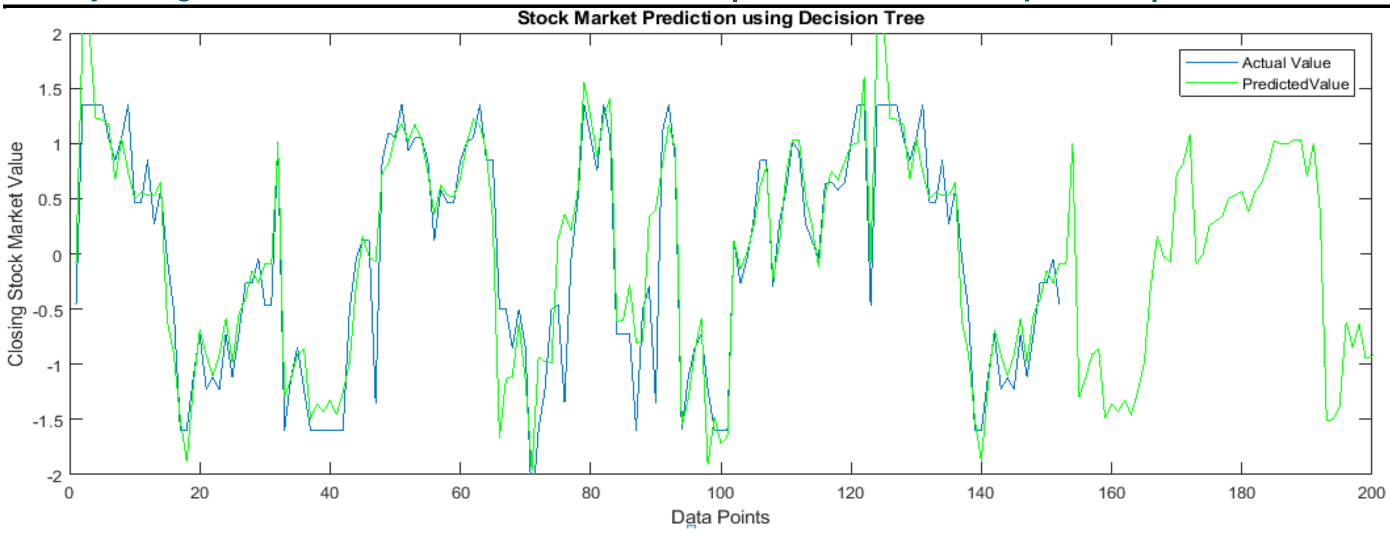

Fig.4.17Prediction using Decision Tree

From the Fig.4.1 shows prediction of future values of financial data. Blue line will shows the actual value predicted as earlier. Green line shows that the prediction with normalized data. Compare the values we can clearly say that prediction is not exactly correct with the actual predicted data. Decision tree predictions have less accuracy than neural network.

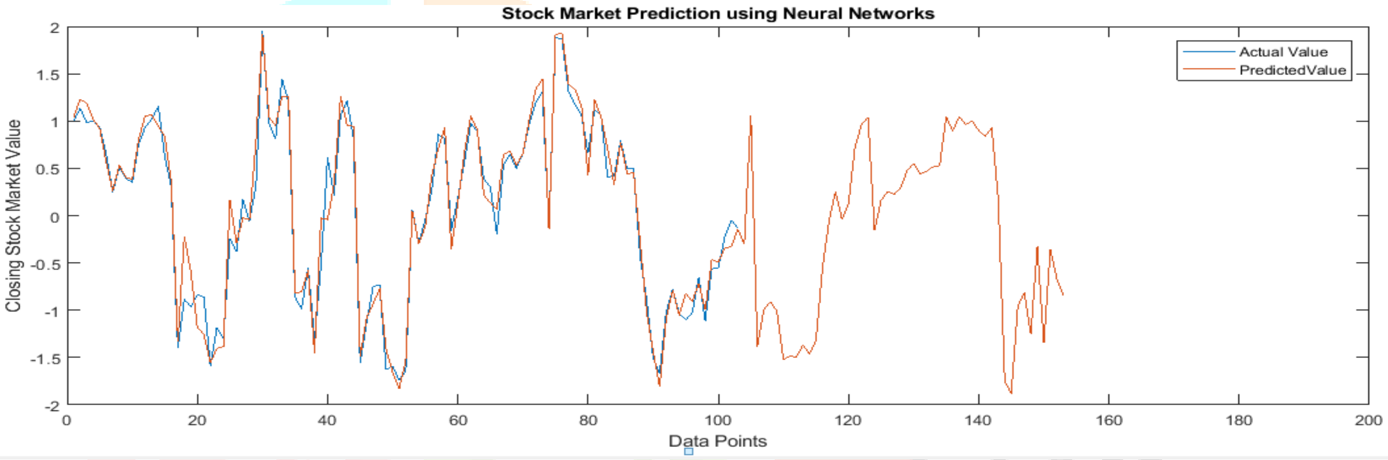

Fig.4.2Prediction using Neural Network.

Fig.4.2 shows future value prediction of a financial data with neural network. Blue line shows the predicted value earlier and the red line shows the future value prediction with normalized data. From the graph, prediction is exactly similar to the actually predicted value. Neural Network shows better prediction with high accuracy.

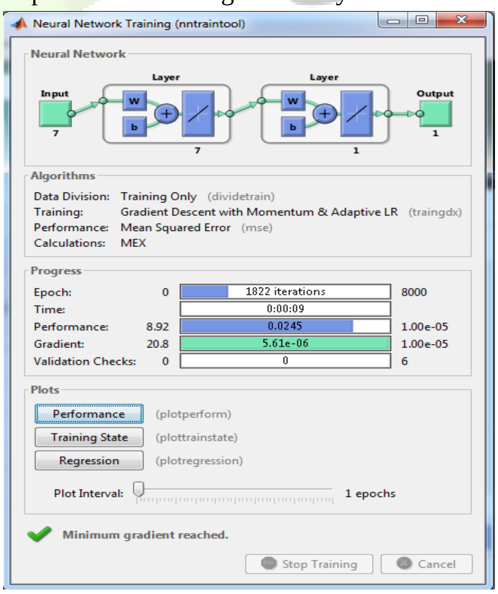

Fig.4.3Neural network training.

Fig.4.2 shows future value prediction of a financial data with neural network. Blue line shows the predicted value earlier and the red line shows the future value prediction with normalized data. From the graph, prediction is exactly similar to the actually predicted value. Neural Network shows better prediction with high accuracy.

Table 4.1: Accuracy Comparison

Predictors	Accuracy
Neural Network	98.4%
Decision Tree	93.7%
SVR	89.67%

V. CONCLUSION

Signal is a any sequence of numerical data that varies with respect to an underlying independent variable mostly time. In financial signal processing of technical analysis is lies the aim of using historical financial data to predict the future market value of business. Financial data are often results of erratic patterns in daily trades. Filtering is important because of highly volatility of such financial instruments caused by rapid fluctuation in their value, Filtering is crucial here, because of tremendous volatility of corresponding instruments generated by rapid inconstancy in their value, since noise consistently fluctuates very swiftly, its frequency is very high and difficult to understand.

Here we using standardization and normalization technique of financial data and find the skewness and kurtosis for understanding the current behaviour of data set, this is call it as a data processing. After that normalization of standardized data given to the feed forward neural network and decision tree. Comparing these two predictors the accuracy of neural network is 98.4% and decision tree make 93.2% .The exact values of future share price can be predicted with the guidance of normalization of standardized previous financial data.

REFERENCES

[1] Q. M. Khaliq, D. A. Voss, S. K. Kazmi Department of MathematicsWestern Illinois University Macomb, IL, USA 61455 Numerical Simulation of Black-Scholes Model for American OptionsA

[2] N. El Karoui, M. Jeanblanc-F'icquk, and S. E. Shreve, Robustness of the Black and Scholes Formula, Mathematical Finance, 1998), 93-126

[3] IEEE Signal Process. Mag. (Special Issue on Signal Processing for Financial Applications), vol. 28, no. 5, Sept. 2011.

[4] R. E. Kalman, “A new approach to linear filtering and prediction problems,” Journal of Fluids Engineering, vol. 82, no. 1, pp.. 35 - 45,1960

[5] Konstantinos Drakakis, “ Analysis of Financial Data through Signal Processing Techniques” 1UCD CASL2 University College Dublin 2008, no. 25, 1203 – 1214.

[6] Seng Hansun, “A New Approach of Moving Average Method in Time Series Analysis, Computer Science”, Universitas Multimedia Nusantara, Tangerang, Indonesia

[7] Gautier Durantin1,2 , Sebastien Scannella1 ,Thibault Gateau1 , Arnaud Delorme2 , Frederic Dehais1 “Moving Average Convergence Divergence Filter Preprocessing for Real-Time Event-Related Peak Activity Onset Detection : Application to fNIRS Signals” Chicago, IL, USA

[8] Ng Ee Hwa, “ Market Strategist for Chart Nexus Fibonacci Retracement – Waiting for the Right Moment”

[9] Konstantinos Drakakis, “ Analysis of Financial Data through Signal Processing Techniques” 1UCD CASL2 University College Dublin 2008, no. 25, 1203 – 1214.

[10] D. Jun et al., Analysis of the global financial crisis using statistical moments, Finance Research Letters (2016),

[11] Berkmen, S.P. , Gelos, G. , Rennhack, R. , Walsh, J.P. , 2012. The global financial crisis: explaining cross-country differences in the output impact. J. Int. Money Finance 31, 42–59 .