**Market Efficiency of U.S. Stock Market: A Test for Semi-Strong Form Efficiency**

**Abstract**

The purpose of this analysis is to test efficient market hypothesis in a semi-strong efficient market using the closing prices of U.S. Stock market data. These tests are related based on different sectors of an industry to draw conclusions on the consistency of the predicted stock prices based on each sector. Stock return prediction is done using a hybrid model of neural networks and time series modelling.

**Contribution & State of Art**

The role of capital market is to allocate ownership of capital stocks in an economy. The ideal way to define is a capital market is a market where companies can make investment decision and investors can make these decisions based under the assumption that prices fully reflect all the information that is publically available. A capital market where prices “always reflect” the information available is known to be an efficient capital market. Efficient capital market can also be described as a market in which people having access to the same information on stock prices try to achieve similar investment goals. The stock market has many profit motivated professionals and private investors who continuously search for mis-valued securities. Both these profit motivated professionals and private investors have similar investment objectives where each is motivated to gain a high profits on returns, certainty in the investment, low risk, and so forth. These goals must adhere to the securities law which states that both parties in a transaction or an investment must have access to the same publically available information.

Efficient market hypothesis (EMH) states that it is not possible to consistently outperform the market—which is usually based on the composite decisions of millions of investors with similar goals and equal access to the same publically available information. In a more general sense EMH explained by Fama [1] states that the stock prices follow a random walk theory and cannot be predicted based on their past behaviour. Here “efficiency” in hypothesis means that the market can quickly digest any new information available on the company, industry, economy, or value of the company and can impound this information accurately into securities prices. Hence in efficient capital markets we can expect people to earn fair returns for the risks taken, no more and no less.

From the hypothesis it does not mean that the investors or participants are denied the profits of investment, it means that for a given risk the benefits from investing in largely competitive markets will be fair (on the average). For instance in an efficient market a news of increase in the earnings of a company will be immediately and accurately assessed by the combined judgement of millions of participants and quickly reflected in the stock prices of the particular company. The result of this efficiency can be seen when you buy the company’s stock before, after, or during release of the news on the earnings. Taking the risk of purchasing or owning the security will yield only in a fair market rate of return in case of efficient market hypothesis.

Efficient market hypothesis is categorized into three forms of market efficiency. The weak form of the efficient market hypothesis holds when values from historical price and volume data cannot be employed to outperform the market. In case of semi strong form of market efficiency, all the publically available information and historical price data is considered to reflect the prices in the future and strong form of efficient market holds when both publically available information and private information is used to predict the future stock prices. While it proves to be a useful benchmark in all its forms, semi strong form of EMH is most widely suggested form to predict the future stock prices. Hadi [8] described and tested various forms of efficient markets and observed that generally semi strong form is considered to be efficient in accounting based research.

Sample studies which provide aggregate analysis on the efficient market hypothesis for Australian data are Hogan, Sharpe and Volker [2] and Groenewold and Kuay [3]. Their results are consistent with those found and summarized by Fama [1] for US data which shows the expected returns when measured over short durations like a daily or weekly. Another study by Groenewold, N [4] tests both weak and semi strong forms of EMH and reports consistent results for daily observations from Australia and New Zealand stock exchange based on the log of prices.

Investments which can be used to gain abnormal profits or returns (in general known as anomalies) violate the efficient market hypothesis in both its semi strong and strong form. In finance the most common anomalies include abnormal returns or unfair profits which are relative to unexpected increase in earnings announcements, size of the firm, month of January, day of the week, undue recommendations from the analysts’, impact of the federal budget deficit announcement, and many others. These anomalies are catalogued and extensive literature survey is provided by Raghubir and Das [[5]](http://www.sciencedirect.com/science/article/pii/S016792360100121X#BIB42), and Hong, H., & Stein, J. C [6].

This paper studies and tests the semi strong market efficiency of pricing data on daily and quarterly observations of NYSE stock exchange. The study also tries to relate the same tests based on the sectors of industry to draw conclusions based on the consistency of the predicted stock prices based on each sector. Various studies have been made based on the tests for semi strong form of efficiency in relation to predicting stock prices using various models and theories. Leuthold, R. M., & Hartmann, P. A [7] uses econometric forecasting model to test the semi strong form of efficient markets and reported that Live-Hog future markets are inefficient. The efficiency of Indian Capital Market was test and concluded as efficient in its semi strong form of efficient market hypothesis by Khan, A. Q., & Ikram, S [9] using Correlation Coefficient and Linear Regression analysis. London Metal Exchange market was tested to be efficient in the sense of semi strong form of efficient market hypothesis by Barry A. Goss [10] and results comprised that the future prices for the metals like copper, zinc and lead fully reflect publically available information. The paper by Ardiansyah, M., & Qoyum, A [11] discusses the concepts of Islamic capital market focussing on Jakarta Islamic Index (JII) to test its market efficiency using market adjusted model and mean adjusted model. Stock split announcements in capital markets were tested for semi strong form of efficiency by Raja, Sudhahar, and Selvam [12] on Indian Stock Market with respect to IT (Information Technology) companies.

Studies also show that some of the capital markets do not hold the semi strong form of efficient markets. This kind of behaviour is shown when the predicted prices do reflect all the information available or we cannot retrieve accurate prices with the data we have. The predictability of expected returns on Athens Stock Exchange was examined using the financial statement information and results indicated that the Greek market does not fully incorporate the information into stock prices, Alexakis, C., Patra, T., & Poshakwale, S [13]. Pele, D. T., & Voineagu, V [14] proposed a model for stock’s return decomposition for testing efficient markets on Romanian Capital Markets and results concluded that the hypothesis of market efficiency cannot be rejected in the weak sense.

All these studies basically review many tests and methods on semi strong form of efficient market hypothesis using the data from various capital markets around the globe. Before we discuss the method used in this paper to predict the expected returns we should know the two approaches which are popular in investment analysis – fundamental analysis and technical analysis. Investment analysis where prediction of the stock prices is based on fundamental factors which are internal or related to the company or its industry, for example product earnings’, competition, management, consumer spending and others is termed as fundamental analysis. A fundamental analyst recommends a purchase for a company which shows a consistent increase in its earnings and believes that the profit will be fair. Therefore, a semi strong form of efficiency opposes the concept of fundamental analysis. Analysis where stock prices can be predicted from historical stock market data which includes changes in prices of stocks and trading volume is termed as technical analysis. A technical analyst believes that all the internal information available on a stock (i.e. fundamental factors) is reflected in the stock price behaviour and suggests a purchase based on the recent prices and volume information. The weak form of efficient market hypothesis opposes the principle of technical analysis.

Our paper uses neural networks model to test the semi strong form of efficient markets by predicting the returns using the pricing information on stocks. The data is categorized mainly based on the quarters and daily stock information. Stock prices and information related to capital markets is time dependent data and forecasting the future prices is majorly dependent on historical prices and all publically available information on the stocks. Pricing information can be considered as a time series which is defined as a sequence of values measured over a period of time (discrete or continuous). Financial information of a company is released publically every year in the form of quarterly and annual statements for a fiscal year. These statements include costs, revenues, expenses, price, earnings and many other factors. Predicting stock prices based on many related factors gives a better result. So we consider a multivariate time series which consists of variables whose values change over time. Hence a multivariate time series analysis is used to study the behaviour of time dependent pricing data and forecast values based on the history of variation of the data.

In general techniques used in time series analysis assume that the relationship among the variables is linear, but data with temporal variables do not exhibit normal regularities and hence are difficult to predict accurately. It seems appropriate to use non-linear modelling and factor analysis like neural networks for temporal data. Neural network approach is a data driven approach and analysis is determined based on the data available. This approach uses the concept of constructing a machine based on the available data. We use this approach to predict the expected returns of the stocks from the multivariate time series based on all the information available publically and historic pricing information. Many studies have used various statistical methods to forecast stock price. Ince, H., & Trafalis, T. B [15] used Principal Component Analysis (PCA) and Factor Analysis in order to identify the most influential inputs for forecasting model with Neural Networks (NN) and Support Vector Regression as the inputs. The results indicate that neural network approach outperformed support vector regression in forecasting the future value of stock price based on technical analysis.

Another study based on trading strategies guided by forecasts of the direction of movement of price on Taiwan Stock Exchange [16] was done using probabilistic neural networks (PNN). The forecasts from statistical performance of the PNN model are measured and compared with the generalized methods of moments. It is observed that the empirical results from using neural networks in investment strategies had higher return values than other strategies in the study. In yet another analysis of forecasting the stock market activity, a hybrid analysis is performed using technique from pattern recognition and artificial intelligence method from neural networks on NYSE data [17]. This hybrid methodology predicted returns that are superior to the rand walk theory. Combining methods (like principal component analysis, genetic algorithms and decision trees) or hybrid techniques [18] have provided better financial factors which are important for stock prediction and future investment decisions rather than using a single statistical technique. Ludvigson, S. C., & Ng, S [19] used dynamic factor analysis approach to model the conditional mean and volatility of stock market returns. This approach allows undermining the limitations from empirical analyses.

Neural Networks was employed in many studies to predict stock price behaviour which yielded in excellent profits [20]. Financial applications like option pricing, stock index trading to currency exchange have used radial basis function networks and neural networks over the decades. Neural Networks are considered to be universal functional approximators which can be used to map a non linear function without actually having any assumption about the data.

We are trying to implement neural network methodology on U.S. stock market data using daily closing prices to compute the returns of each company. The expected returns are computed based on daily and quarterly stock price data. Many studies have used U.S. stock market data to forecast returns implementing many different methods like Gaussian models [26], Bayesian approach, support vector regression [21] and artificial immune algorithms. Giles, C. L., Lawrence, S., & Tsoi, A. C [24] used recurrent neural networks on ‘Chicago Mercantile Exchange’ for financial prediction and White, H. [25] used neural networks to predict the returns of IBM daily stock prices. Stock return predictability measures from regression analysis were used to validate the adaptive market hypothesis [22] which states that returns are subject to market conditions. Observing the non-linearity of the stock market data we have tried to employ neural network factor model to forecast the returns of U.S. stock data.

The study below presents the following sections where data is described, methods used are explained and results are discussed. Section 2 describes the data considered for the analysis and description of every financial indicator used. Section 3 presents the neural network architecture we used in our analysis and the moving average time series model of statistical prediction. The performance and results from the network model are noted in section 4 which is followed by discussion and conclusion of our analysis.

**Data**

Data accumulation and parsing is a recursive process and when we have to analyze financial data then collection of data is an on-going process. The dataset considered is a combination of both profit and loss statement and balance sheet. P&l statement (mostly known as income statement) has indicators which help us in identifying the performance and profitability of a company whereas Balance Sheet has indicators which help us in identifying the financial status of each company. Thus using these indicators from both income statement and balance sheets we try to forecast the stocks with better expected returns.

Every public company issues financial statements quarterly and annually for a fiscal year which includes costs, revenues and expenses. To analyze the profitability and performance of a company it is essential for stock investors to understand the various aspects of these financial statements as they play an important role in whether or not we should invest in a company or stock. The dataset has about 88 indicators for quarterly data on each stock. Each stock or ticker has data on prices, revenue, gross profit, net income, earnings per basic share and other indicators. The expected return value for each stock is calculated for the purpose of this analysis using the quarterly pricing information. We are trying to identify stocks with better expected returns using the indicators in the table below. The final dataset has 15 indicators with quarterly data on 5500 stocks approximately. Each indicator is chosen according to its relation in identifying the stocks with better expected returns. To this data set we will add the daily closing prices, trading volume and the daily US Treasury yield curve rates for each stock. This data is available on the Data and Charts Center of [treasury.gov](file:///C:\madhumita\Desktop\DS-Sem3-Spring2017\DS-670-BigDataBusinessAnalytics\Assignments\Assignment13-FinalManuscript\DS-670-Assignment13-Final-Manuscript.docx) which coordinates with the U.S. Department of the Treasury.

**Method**

Neural Networks are valued as one of the foremost techniques in performing nonlinear forecasting tasks and complicated pattern recognition across many applications and domains. Some such interesting applications of neural networks are in decoding protein sequences and deterministic chaos. With these successful applications in varied domains, it is interesting to know whether applying this technique to economic time series will yield in extracting the non-linear regularities of the economic market. Decoding the regularities in asset price movement like day to day or month to month fluctuations of stock prices (which were undetected previously) will be the key to earn benefits from investing.

Efficient market hypothesis states that asset or stock prices follow random walk theory. Apart from the expected risk free returns, the price movement of an asset is totally unpredictable based on the publically available data on the asset such as the daily price and the volume history of the stock itself or that of any other stock. It should be observed that insider information (publically unavailable) cannot be considered for forecasting returns to gain unusual profits. The absence of predictability often opens profit opportunities in stock markets only for a small period of time. For example if we expect a strong increase in the price of an asset, an investor goes a long and invests in this asset which drives the price to the expected level, thereby quickly exploiting the profit opportunity which was available a few moments ago. Understanding the simple form of efficient market hypothesis it is feasible to believe that undetected regularities (high and low price movements) exist in historical price data and may persist to exist. Many researchers (mostly academic) and practitioners made tremendous efforts using various techniques to predict the price movements in stock markets and develop financial strategies to convert these predictions into profits. Among the various econometric forecasting techniques used, neural networks methodology seemed to provide better results in trading and financial forecasting. Hence, using neural networks learning methods to find these non-linear regularities in the pricing information may prove useful to the market participants.

The prediction of expected returns using the stock price information of U.S. market is made up of many neural networks that learn the relationship between technical and economical factors. The aim is to forecast the daily and quarterly returns of the all the stocks. Figure 1 depicts the basic architecture of the prediction system. It basically converts all the factors considered (technical and economic) into a space pattern which form the input to the neural network. The input factors are discussed in detail later.

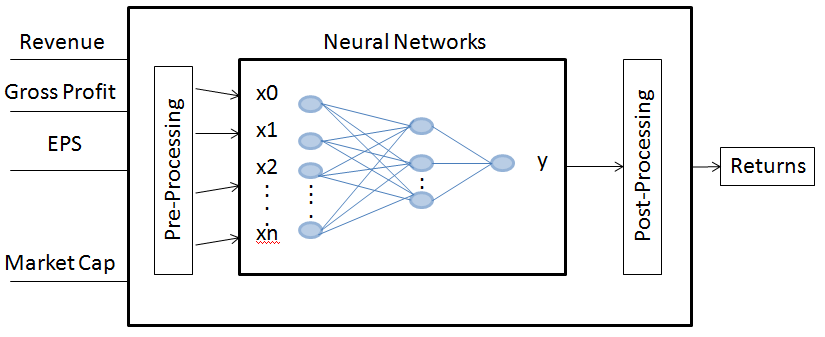


Figure 1: Basic architecture of predicted system

**Neural Networks**

Neural network model is considered as an intelligent data mining technique which is used in pattern recognition problems like stock price predictions. It is a non-linear, non-parametric, data driven model which is trained to map the past values (here past prices and volume information) of a time series. The idea of using a neural network was initially inspired from the nervous system of human beings. The nervous system consists of vast number of processing units known as neurons (Figure 2). Every neuron receives signals from other neurons or from outside and they process these signals using activation function which in turn produces an output. The output from one neuron is sent to other neurons. The impact of each input is different from other inputs. For instance impact of ith neuron on jth neuron is given by a value ‘w’, which is the weight of the connection between both the neurons i and j (Figure 3). That is, each neuron or unit in the network receives input from the lower level neurons and they perform weighted summation to derive the output. We can either use a linear function or sigmoid function as the output function.

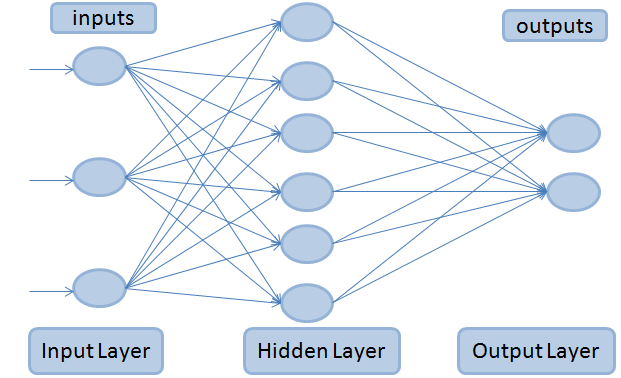


Figure 2: Architecture of Feed Forward MLP network.

The connection is considered to be strong if the weight ‘w’ is high and vice versa. A feed forward multi layer perceptron neural network is made of layers of neurons (Figure 2). The input data is connected to the first layer of the network. After this we can have one or more layers in between which are known as hidden layers. The units or neurons in these hidden layers are called hidden units. These hidden layers are connected to the final output layer which returns the results. Unlike the recurrent neural networks, feedback networks have all the connections towards the output layer. Figure 2 depicts the basic feed forward neural network architecture. It is a three layered network: the input layer, the hidden layer, and the output layer. These three layers are connected completely to form a hierarchical structure. wij

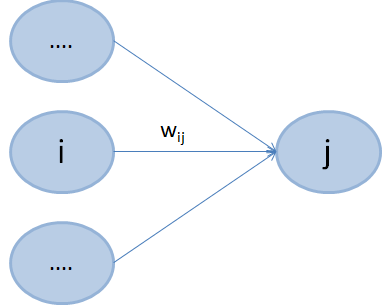


Figure 3: Perceptron neuron’s connections

The overall network function is designed in an equation which shows the inputs (xi), hidden units with the corresponding weights for each layer (wji for first hidden layer and wkj for the second hidden layer) and the output unit (yk(x,w)). The activation function for the output unit can be considered a sigmoidal or linear activation function. The activation function h() for the hidden units is assumed to be sigmoidal and for the output units is linear.

yk(**x**,**w**) = σ (**h(**xi+ wj0(1)**)** + wk0(2)) *[ if yk is sigmoidal ]*

= **h(**xi+ wj0(1)**)** + wk0(2) *[ if yk is linear ]*

The common choice for the activation function is yk = σ(ak) where,

σ(a) = 

Note that in our case there is one network output per observation and that output is the forecasted log-returns for the observation. Thus, in our case k=1 in the neural network equation. The learning algorithm used in feed forward multi layer perceptron (MLP) networks is the error back propagation. Error back propagation method is considered as a representative learning rule for hierarchical MLP networks. This learning method allows the network to learn the patterns in the available data and compute the weight of the connection in the inverse direction with respect to the error function. The error function is the regularized sum of squared error. This back propagation methodology selects training vector from the training dataset and processes it from the input layer of the network to the output layer. The error is calculated in the output layer and propagated back to previous layers so that the weights on the connections can be corrected. This process continues until the error reaches a pre defined value. It can be noted that any continuous function can be approximated using a three layered feed forward MLP network with any kind of precision. With any increase in the number of layers and neurons of the neural network, we can see that the learning speed will decrease gradually.

**ARMA models**

The ARMA model or autoregressive moving average model is generally applied to time series data. Most commonly it is used to compute and estimate the Value at Risk (VAR), an estimation model used in simulation forecasting of asset prices. The ARMA model can be used to forecast the future values for a given time series data. Forecasting future stock prices is based on the past prices and volume information which can be converted to a time series data. Many commercial banks employ ARMA methods to predict VAR values.

ARMA has the ability incorporate both the autoregressive models and moving average models. This is usually referred as ARMA ( p, q ) model. The order of autoregressive terms is represented by p and the order of moving average model is represented by q. The autoregressive model of order p is indicated by AR ( p ) notation. The autoregressive model AR ( p ) is written as:

**Xt = c +** **Xt-i + εt**

where c is a constant, ϕ is the parameter of the model and ε is the error term. The moving average model of order q is indicated by MA ( q ) notation. The moving average model MA ( q ) is written as:

**Xt = εt +** **εt-i**

where θ is the parameter of the model and ε is the error term. . The autoregressive moving average model of order p, q (p – autoregressive terms and q – moving average terms) is indicated by ARMA ( p, q ) notation. This has two models - AR ( p ) and MA ( q ).The autoregressive moving average model ARMA ( p, q ) is written as:

**Xt = εt +** **Xt-i +** **εt-i**

The flexibility and efficiency of the neural networks gives it the ability to deal with robust and uncertain data. This means that to forecast the prices or returns of an asset, neural networks in combination with time series analysis can be efficiently used in stock market analysis. Many previous studies have shown that neural networks outperform linear models, statistical models and classical forecasting. Combining both features of neural networks and ARMA time series model, building a network for prediction of returns is much more complicated. In order to get a robust model all the crucial factors must be considered while designing a prediction model. One of the major factors in networking is the structure of the model which includes number of layers, connections and neurons. Many other factors influence the performance of the model like training algorithm, each neuron’s activation function, dividing the train and test data sets, normalization of data and measurement evaluation (for weights). We have two kinds of neural networks, a feed forward multi layer perceptron and recurrent neural networks, which are used to forecast a company’s asset based on the asset’s share value and historical information. The back propagation learning algorithm is used to train the data in the networks.

The inputs to our neural network are various indicators from the financial statements which determine the company’s performance for a fiscal year. There are three forms of financial statements a company releases quarterly and annually and they are income statement, cash flow statement and balance sheet. The model which is proposed in this paper can be observed as a time series prediction model. The inputs considered are 20 indicators (technical and economical) from the financial statements and the output is the expected returns of every stock. Table 1 shows the different types of financial statements.

|  |  |
| --- | --- |
| Balance Sheet | Comprises of Assets, Liability and Equity |
| Income Statement | Reports Company’s Performance |
| Cash Flow Statement | Presents Operating, Investing and Financing Activities |

Table 1: Types of Financial Statements

The proposed model uses a four layered neural network with back propagation learning algorithm to train the network. The four layers are the input layer, two hidden layers and an output layer. A neural network always takes normalized data as input. The data from various indicators is normalized and is given as input to the first layer. Hence we have 20 neurons in the input layer. The choice of the number of hidden units depends on the number of inputs and number of outputs. The ideal choice is to consider two thirds of the size of input layer, plus size of the output layer and it should be less than the number of inputs. Our model has two hidden layers with eight and seven hidden units in each hidden layer. The output layer has only one neuron which gives the predicted value of returns of stocks (computation is shown in Table 2). Figure 4 shows the proposed neural network model with two hidden layers and Table 2 gives the description of new explanatory variables.

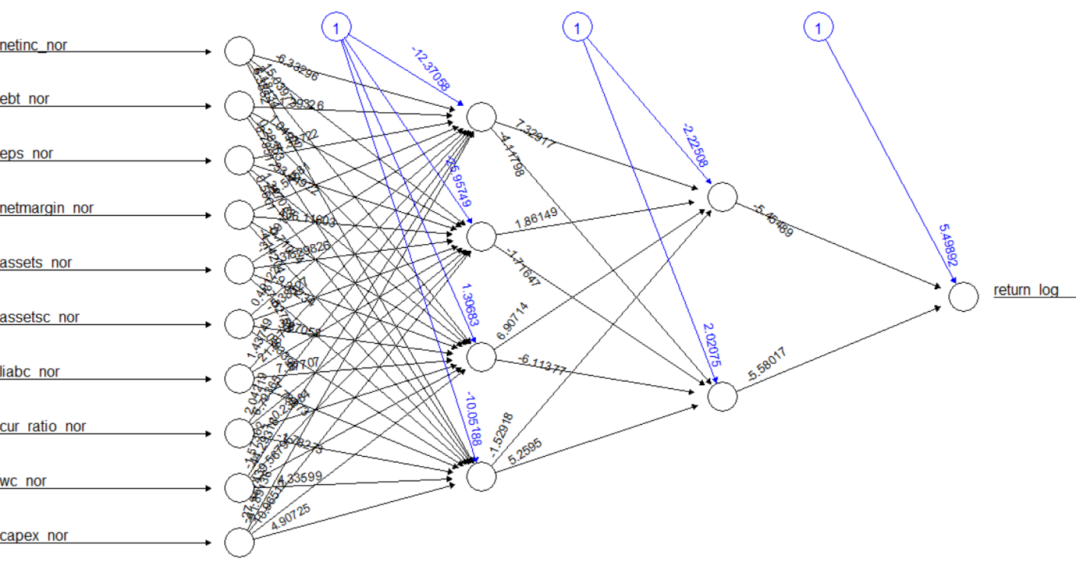


Figure 4: Neural network sample.

|  |  |
| --- | --- |
| Notation | Description |
| Pt | Stock Price |
| pt | log Pt |
| Rt | Return = (Pt /Pt-1) – 1 |
| rt | Return from ‘risk free’ investment |
| Rt – rt | Excess return |

Table 2: Explanatory variables

The hybrid approach of using neural networks with time series analysis addresses the fundamental problems by predicting a noisy time series. Predicting stock prices in an economic system is subject to change vigorously. For such forecasting the prediction rules continuously change and hence learning and prediction must adhere to these changing rules. A prediction system known as moving simulation was introduced for this kind of changes. In moving simulation method the prediction is carried out by simulation while moving the objective learning and prediction time intervals. Observing figure 5, the learning of data is done for M months in the past and then prediction is done for the next L months. The system continues this process until we can forecast the returns of the stocks.

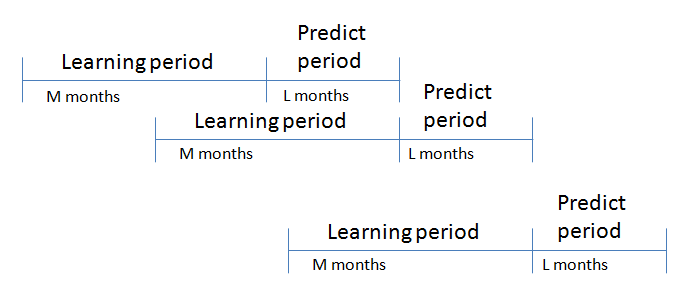


Figure 5: Moving Simulation method

The proposed neural network model in this paper follows this simulation method. For instance we train the neural network for stock market data from September 2014 quarter using the high speed learning algorithm (error back propagation algorithm). We use this model to predict the expected returns of December 2014 quarter. This is learning the data for M past months – September quarter and predicting the values for next L months – December quarter.

Neural Networks are considered as highly dynamic, non-linear, large scale, continuous and time based systems. The knowledge from a neural network is stored on the numerous nodes in the form of a weight matrix where each weight depicts the relationship between the nodes. Neural network is a data driven approach where learning and training of the model is based on the available data. This feature makes it very robust to deal with the uncertainty of data. Neural networks are highly capable in pattern recognition and dealing with machine learning challenges. It is believed that data driven methods are best ways to handle the real time prediction problems. The next section explains in detail the results from our proposed neural networks model with time series data analysis to forecast the returns of stock market data.

**Results**

We analysed the data after the neural network model for each quarter was established. The expected returns were calculated based on these models for the test data. The neural network model for every quarter had 4 layers – input layer consisting of the technical, fundamental and macro-economic factors, two hidden layers with six and four hidden units respectively, and an output layer which results in the expected returns.

Initially, the selected indicator data is normalized this allows the algorithm to reach a global minimum faster. Before fitting the neural network model, we have prepared the formula (i.e. y ~ x1 + x2+...) as the function neural net in R does not accept strings. Later, the neural network architecture was employed with ‘formula’ and 2 hidden layers for the two thirds of the available dates (which form the training data). Using these models the future quarter’s returns have been predicted.

After the stock returns for test data has been computed, the stocks for every quarter are ranked based on these returns. These ranked stocks are divided into five quantiles – highest to lowest. The average of each quantile defines the average returns of stocks in that bucket. In the similar way, returns, ranking and quantiles have been computed for all the future five quarters. Based on every quarter’s average returns of each quantile we will be able to identify the trends in each quantile over a period. This will help us in identifying whether there was a positive trend or negative trend followed in a particular quarter. The difference between top quantile and the bottom quantile helps us in identifying the dispersion in returns in a quarter. Also, we can identify the expected returns of top quantiles or bottom quantile in each quarter which helps us in making investment decisions (buy or sell). To demonstrate the above theory in a particular sector we have used the proposed model in Finance sector for quarters – Q4 2014, Q1 2015, Q2 2015, Q3 2015 and Q4 2015.

The results from the quarterly stock analysis (Figure 6) have been explained in detail. Based on the overall forecasted values we can say that the expected return of financial stocks have shown a positive trend in all the quarters with Group 1 stocks out performing Group 5 stocks. This can be observed in both predicted values and actual values.

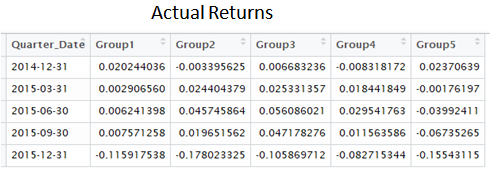
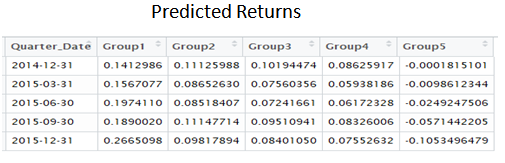


Figure 6: Quarterly Average Predicted and Actual Returns for each quantile.

The spread between the performance of group 1 and group 5 (Figure 7) has increased for every quarter. On the whole the averages depict a huge volatility in stock price movements. The return predictability is dependent on market conditions and reflects all the publically available information which is evidence that the U.S. stock markets are semi-strong form efficient.

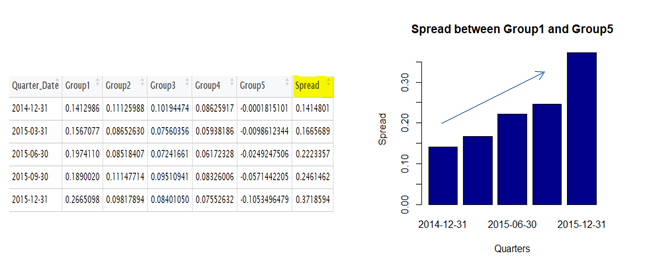


Figure 7: Spread between Group1 and Group5 for each quarter. Group1 – well performed stocks and Group 5 – poorly performed stocks.

**Discussion:**

Traditional statistical methods for prediction have their limitations in the applications with data sets which show nonlinearities. Many forecasting techniques are capable only of selecting general trends like positive and negative trends, and show complexity in modelling cycles which by no means are repetitive in period, amplitude, or shape. Even though such inadequacies are present, modelling techniques such as multivariate linear regression is used routinely in the Capital Markets and also have proved to be a very useful tool.

In this study we have showed that using a simple neural network employing publically available data we are able to successfully forecast a group of stocks which will out-perform the market as well as a group of stocks which will under-perform the market, directly contradicting the semi-strong form of the EMH. With their smooth interpolation properties neural networks allow models to fit better to the data and generalize significantly better.

We have used a hybrid approach of neural networks and time series analysis to calculate the return predictability. Based on the overall forecasted values we can say that the expected return of stocks have shown a positive trend in all the quarters. The spread between the performance of group 1 and group 5 has increased for every quarter. On the whole the averages depict a huge volatility in stock price movements.

We have assumed that the price for each stock is highly influenced by the fundamental, technical and macro-economic factors while the regression analysis [22] was done under the assumption that the market follows a weak-form of efficient market hypothesis (EMH). This assumption was made to test for evidence of AMH (adaptive market hypothesis) in return predictability. The hybrid approach followed in this study helped to test for semi-strong form of efficient market hypothesis (EMH) based on return predictability. The returns were computed from the pricing information for each stock and predicted for the future time using the neural networks and time series analysis.

Regression analysis in [22] uses Dow Jones Industrial Average (DJIA) as a technical indicator. The index is a price-weighted average of 30 blue-chip stocks, accounting for 25–30% of the total value of U.S. stocks. Index created for 30 blue chip stocks (large cap stocks) for century-long U.S. stock market data from 1900 to 2009. We have used daily closing prices of both small and large cap stocks captured from exchange traded fund (ETF) for the duration of 2011 to 2016 for U.S. stock market. In our model Relative Strength Index (RSI) is used as a technical indicator. Relative Strength Index compares the magnitude of recent gains and losses over a period of time based on the change of price movements.

Neural Networks and Factor Models were implemented to forecast the returns of a stock in the market. We use a multivariate time series to study the behaviour of stocks. To weigh the use of neural networks we compare the results and metrics to Regression Analysis, which was conducted on the measures of return predictability of stocks. Many different metrics were used to compare and compute the results. In Neural Networks sequential testing of market efficiency plays a major role in computing t the returns accurately based on all the historical and present data available while sequential testing of market efficiency was not taken into consideration in linear regression.

Our analysis addresses that the factors chosen drive the performance of the model but also the Investability of our model with real-time data. This entails simulation of a portfolio controlling for risk and liquidity of the assets traded. We showed that predicted returns reflect the publically available information and are dependent on market conditions which prove that markets are semi-strong form efficient.

**Conclusion:**

The hybrid approach combining Neural Networks and Time-Series used in this study was able to categorize between the stocks that performed well and stocks which performed poorly in the market. This approach also outperformed the multiple linear regression analysis approach overall. Investors who use this modelling technology should be able to improve their analysis choosing the best investment alternatives based on better predictions. However, further research is required in using this hybrid approach in stock prediction analysis. To further validate the proposed technique in investment analysis, we need to perform qualitative and quantitative tests on our data, considering different time periods and different industries or sectors. With the many advantages, investors and researchers should seriously consider the application of neural network and time-series techniques to investment analysis.

**References**

[1] Fama, E. F. (1991). Efficient capital markets: II. The journal of finance, 46(5), 1575-1617 and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The journal of Finance, 25(2), 383-417.

[2] Hogan, W. P., Sharpe, I. G., & Volker, P. A. (1982). Capital market efficiency and the relationship between equity returns, interest rates, and monetary aggregates in Australia. Journal of Economics and Business, 34(4), 377-385.

[3] Groenewold, N., & Kang, K. C. (1993). The semi‐strong efficiency of the Australian share market. Economic Record, 69(4), 405-410.

[4] Groenewold, N. (1997). Share market efficiency: tests using daily data for Australia and New Zealand. Applied Financial Economics, 7(6), 645-657.

[5] Raghubir, P., & Das, S. R. (1999). A case for theory-driven experimental enquiry. Financial Analysts Journal, 56-79.

[6] Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of finance, 54(6), 2143-2184.

[7] Leuthold, R. M., & Hartmann, P. A. (1979). A semi-strong form evaluation of the efficiency of the hog futures market. American Journal of Agricultural Economics, 61(3), 482-489.

[8] Hadi. “Review of Capital Market Efficiency: Some Evidence from Jordanian Market”. International Research Journal of Finance and Economics, Year: 2006, Issue: 3, pp 13-26.

[9] Khan, A. Q., & Ikram, S. (2010). Testing semi-strong form of efficient market hypothesis in relation to the impact of foreign institutional investors'(FII's) investments on Indian capital market. International Journal of Trade, Economics and Finance, 1(4), 373.

[10] Barry A. Goss (1983) The semi-strong form efficiency of the London Metal Exchange , Applied Economics, 15:5, 681-698, DOI: 10.1080/00036848300000044

[11] Ardiansyah, M., & Qoyum, A. (2012). Testing the Semi-strong Form Efficiency of Islamic Capital Market With Response to Information Content of Dividend Announcement: A Study in Jakarta Islamic Index. Journal of Modern Accounting and Auditing, 8(7), 1025.

[12] Raja, M., Sudhahar, J. C., & Selvam, M. (2009). Testing the semi-strong form efficiency of Indian stock market with respect to information content of stock split announcement: a study in IT industry. International Research Journal of Finance and Economics, 25(2009), 7-20.

[13] Alexakis, C., Patra, T., & Poshakwale, S. (2010). Predictability of stock returns using financial statement information: evidence on semi-strong efficiency of emerging Greek stock market. Applied Financial Economics, 20(16), 1321-1326.

[14] Pele, D. T., & Voineagu, V. (2008). Testing market efficiency via decomposition of stock return. Application to Romanian Capital Market. Romanian Journal of Economic Forecasting, 3(2008), 63-79.

[15] Ince, H., & Trafalis, T. B. (2007). Kernel principal component analysis and support vector machines for stock price prediction. IIE Transactions, 39(6), 629-637.

[16] Chen, A. S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. Computers & Operations Research, 30(6), 901-923.

[17] Leigh, W., Paz, M., & Purvis, R. (2002). An analysis of a hybrid neural network and pattern recognition technique for predicting short-term increases in the NYSE composite index. Omega, 30(2), 69-76.

[18] Tsai, C. F., & Hsiao, Y. C. (2010). Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. Decision Support Systems, 50(1), 258-269.

[19] Ludvigson, S. C., & Ng, S. (2007). The empirical risk–return relation: A factor analysis approach. Journal of Financial Economics, 83(1), 171-222.

[20] Kimoto, T., Asakawa, K., Yoda, M., & Takeoka, M. (1990, June). Stock market prediction system with modular neural networks. In Neural Networks, 1990., 1990 IJCNN International Joint Conference on (pp. 1-6). IEEE.

[21] Chiu, D. Y., Shiu, C. Y., & Lin, Y. S. (2011, June). USA S&P 500 stock market dynamism exploration with moving window and artificial intelligence approach. In Networked Computing and Advanced Information Management (NCM), 2011 7th International Conference on (pp. 341-345). IEEE.

[22] Kim, J. H., Shamsuddin, A., & Lim, K. P. (2011). Stock return predictability and the adaptive markets hypothesis: Evidence from century-long US data. Journal of Empirical Finance, 18(5), 868-879.

[23] Leigh, W., Frohlich, C. J., Hornik, S., Purvis, R. L., & Roberts, T. L. (2008). Trading with a stock chart heuristic. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 38(1), 93-104.

[24] Giles, C. L., Lawrence, S., & Tsoi, A. C. (1997, March). Rule inference for financial prediction using recurrent neural networks. In Computational Intelligence for Financial Engineering (CIFEr), 1997., Proceedings of the IEEE/IAFE 1997 (pp. 253-259). IEEE.

[25] White, H. (1988). Economic prediction using neural networks: The case of IBM daily stock returns.

[26] Petelin, D., Šindelář, J., Přikryl, J., & Kocijan, J. (2011, September). Financial modeling using Gaussian process models. In Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2011 IEEE 6th International Conference on (Vol. 2, pp. 672-677). IEEE.