IOSR Journal of Business and Management (IOSR-JBM)
e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 20, Issue 12. Ver. V (December. 2018), PP 63-79
www.iosrjournals.org

An Evolutionary Approach to Bombay Stock Exchange Prediction with Deep Learning Technique

Dr. Asif Perwej[1*], Prof. (Dr.) K. P. Yadav[2], Prof. (Dr.) Vishal Sood[3], Dr.Yusuf Perwej[4]
[1]Assistant Professor, School of Management Studies, Sangam University, Bhilwara, Rajasthan, India*
[2] Vice Chancellor, Sangam University, Bhilwara, Rajasthan,India
[3]Dy.Dean, School of Management Studies, Sangam University, Bhilwara, Rajasthan, India
[4]Assistant Professor, Department of Information Technology, Al Baha University, Al Baha, Kingdom of Saudi Arabia(KSA)
Corresponding Author:Dr. Asif Perwej

Abstract:*The prediction and analysis of stock market data have gotten avitalrole in today's economy. The financial prediction or in particular stock market prediction is one of the hottest fields of research lately due to its commercial applications owing to high stakes and the kinds of temptations gain that it has to offer. In this paper, we are using two types of models and compared those using historical data from the Bombay Stock Exchange. The models used were based on firstly, the Deep learning parameters updated through particle swarm optimization and secondly, the Deep learning parameters updated through the least mean square and data taken from Bombay Stock Exchange (BSE). The Deep learning algorithms vary extensively in the preference of network structure, activation function, and other model parameters, and their performance is known to depend heavily on the procedure of data representation. Its capability to extract features from a huge set of raw data without believingonearlier knowledge of predictors makes deep learning potentially tempting forBombay stock market prediction at high frequencies. The predicting the Bombay stock price of a company based on the historical value day to day open, high, low, close and volume of the related index. The outcome of the experiment has been depicted with the assistance of appropriate curves where a comparative analysis of the two types of models is done on the basis ofseveralparameters. This prediction model is used with some general indicators to maximize the return and reduce the risk of the stock market.From the outcome obtained, it is understandable that the deep neural network models are competent of recognizingthe patterns existing in the Bombay stock markets.*

Keywords:*Bombay Stock Exchange, Deep Learning, Particle Swarm Optimization, Prediction, Least Mean Square,Stock Market.*

Date of Submission: 17-12-2018 Date of acceptance: 01-01-2019

I. Introduction

Nowadays the stock market is a major part of a country's economy, understanding it is becoming more and more necessary. The prophecy the stock market is an essential field of identifying the future value of the price for preferable exchange of finance between the companies having the shares. In the scope of forecasting the financial status, stock market prediction has become a major one [1]. Once the financial status of a company has been as a matter of fact predicted, the investors will be confidently investing their money. A stock is really a part of a company and gives the owner of the stock the right to a part of the company's present and future net assets [2]. These stocks are bought and sold by investors on several stock markets worldwide. The stock market is a supply and demand market and has a lot of stockholders dedicating time and money to [3]develop models and techniques to assess the trend of the market for monetary gain. Accordingly, a lot of notice had been devoted to the analysis and prediction of future values and trends of the financial stock markets, and due to applications in several business transactions, stock market prediction has become anavidtopic of research [4]. There is no suspicion that the large number of the people related to stock markets are trying to achieve benefit. The profit is obtainedby investing in stocks that have a good future (short or long term future). The various methods exist for predicting the stock market and one of the most common is to use [5] software that uses graphical and statistical technique to prophecy the market [6]. The objective of prophecy research has been largely beyond the capability of conventional artificial intelligence research which has basically focused on developing intelligent systems that are supposed to emulate human intelligence. By its nature the stock market is mostly sophisticated (non-linear) and volatile. For the predictive model, we propose to use deep learning[7] to capture the impact of news events over a history that is longer than a day. Deep learning is essentially aselfreliant, self-teaching system in which you use existing data to train algorithms to discover patterns and then use that to make predictions about new data [8]. Deep learning is a subset of machine learning in artificial intelligence that has networks,competentof learning unsupervised from data that is [9] unlabeled or unstructured. Deep learning models can gain a state-of-the-art precision, sometimes exceeding human-level performance. This modelis trained by using a huge set of labeled data and neural network architectures that contain several layers [10]. Farther the market simulation shows that our model is more proficient of making profits compared to foregoing techniques. To our knowledge, we are the first to use a deep learning model for event-driven Bombay stock market prediction [11], which gives the best reported outcome in this paper.

II. Related Work

This section introduces the related work from the stock market prediction method For the past few decades, soft computing has been used for stock market prediction. Halbert white in [12] reported some results of an on-going project using neural network modeling and learning techniques to search for and decode nonlinear regularities in asset price movements. The main focus on case of IBM common stock daily returns.Kyoung-jae Kim and Won Boo Lee [13] developed a feature transformation method using genetic algorithms. This technique decreasesthe dimensionality of the feature space and removes irrelevant factors involved in stock price prediction. Tiffany Hui-Kuang and Kun-Huang Huarng in [14] used neural network because of their capabilities in handling nonlinear relationship and also implement a new fuzzy time series model to improve forecasting. The fuzzy connection is used to predict the Taiwan stock index. In the neural network, fuzzy time series model where asinsample observations are used for training and out-sample observations are used for forecasting. The Asif Perwej, Yusuf Perwej, Nikhat Akhtar and Firoj Parwej [15] introduced radial basis function (RBF) and function linked artificial neural network (FLANN) algorithms for prophecy of financial data. The based our models on data taken and differentiate those using historical data from the Bombay Stock Exchange (BSE). The RBF and FLANN parameters updated by Particle swarm optimization (PSO). Another research done on genetic algorithms (GAs) by Kyoung-jae Kim [16] to predict stock market by using GA not only to improve the learning algorithm, but also to reduce the complexity of the feature space. Thus, this technique decreasesdimensionality of the feature space and improve the generalizability of the classifier. Tong-SengQuah in [17] presented methodologies to select equities based on soft-computing models which focus on applying fundamental analysis for equities screening. This paper compares the performance of three soft-computing models, namely multilayer perceptrons (MLP), adaptive neuro-fuzzy inference systems (ANFIS) and general growing and pruning radial basis function (GGAP-RBF). Roman et.al performed an analysis on multiple stock market return using Back propagation and RNN [18]. The Chen et al. [19] proposed a double-layer neural network for high-frequency forecasting, with links specially designed to capture dependence structures among stock returns within different business sectors. M. thenmozhi et al in [20] presented studies on artificial neural networks that have the capacity to learn the underlying mechanics of stock markets. Actually, artificial neural networks have been widely used for forecast financial markets. In spite of, such applications to Indian stock markets are scarce. This paper enforcesneural network models to forecast the daily returns of the BSE Sensex. The multilayer perceptron network is used to construct the daily returns model and the network is trained using an error back propagation algorithm. In [21], they used ML algorithms like Least Square Support Vector Machine (LSSVM) and Particle Swarm Optimization (PSO) for stock market prediction. J .Štěpánek et al [22], a proposed data mining approach using genetic algorithms (GA) to solve the knowledge acquisition problems that are inherent in constructing and maintaining rule-based applications in the stock market.TheJacek in [23] presented neuro-evolutionary method for a short-term stock index prophecy is presented. The data is gathered from the German stock exchange (the target market) and two other markets (Tokyo Stock Exchange and New York Stock Exchange) together with EUR/USD and USD&JPY exchange rates. The artificial neural networks endorsement by genetic algorithm are used as the prediction engine. In [24], Kim et.al proposed a different approach for stock market prediction. Emad W. Saad et al. in [25] compared three networks for low false alarm stock trend predictions. The short-term trends, especially tempting for neural network analysis, can be used profitably in scenarios such as option trading, but only with considerable risk. Jia H. i et.al con- ducted a comparison study between Feed Forward MLP an Elman Recurrent Network for predicting the stock value of company [26]. Qui-yong Zhao in [27] predicted accuracy of price date by the traditional BP network by considering asingle closing price as the time series vector. But, in this paper author also add another factor vector to the BP network input vector so that low training accuracy caused by the a large number of factors can be recovered. Ching-Hseuecheng, Tai-Liang chen, Liang-Ying Wei in [28] this paper proposed a hybrid forecasting model using multi-technical indicators to predict stock price trends. There are four method mentionsuch as select the essential technical indicators, the favorite indicators based on a correlation matrix and use CDPA to minimize the entropy principle approach. In 2016 Chiang et al. proposed a dynamic stock prediction system using Predicted Square Error (PSE) and neural network [29]. Feng Li et al. in [30] analzsed complexity of the interior and a variety of exterior structure of the stock price system based on BP neural network to provide a prediction model for stock market by utilizing three-layered feed forward neural networks, presents topology of network, principles of determining the number of hidden layers selection and pre-treatment of sample data and determination of preliminary parameters. Chong et al. [31] applied a deep feature learning-based stock market prediction model, which extract information fromthe stock return time series without relying on prior knowledge of the predictors and tested it with high-frequency data from the Korean stock market.

III.Familiarization to the Stock Market

A stock market is a collection of buyers and sellers of stocks.In this section we are going to discuss some of the basics of stock market i.e. What is stock market, stock market index, stock exchange and many other concepts related to the stock market.

3.1 Stock Market

A stock market is essentially a genuine forum where buyers and sellers of an asset can meet to trade standardized share instruments. The stock markets are handlingby stock exchanges, which accept submissions from public business firm to be listed for the public sale of their securities [32]. In order to become listed, business firm must meet extensive qualifying thresholds, designed to safeguard traders and ensure a level of standardization in shares across the market. Stock markets assess the price of underlying assets, by matching the prices willing to be paid by both sellers and buyers to identify the market price of an asset at any particular time.A stock that is highly in demand will increase in price, whereas as a stock that is being heavily sold will deficiency in value.

3.2 Stock Market Index

The stock market index is the procedure of showing the comprehensive performance of all the companies that are listed on the stock market with a single number. This stock market index Number is used by all the investors and traders to understand the present comprehensive performance of the market and in the prophecy future of the stock market depending on the bygone values of the stock market index.Theprocedure of stock selection could be the type of industry, market capitalization or the size of the company [33]. The value of the stock market index is computed using the values of the basic stocks. Any transformation taking place in the underlying stock prices influencesthe overall value of the index.If the prices of most of the underlying securities enlargement, then the index will enlargement and vice-versa. In this procedure, a stock index reflects the overall market sentiment and direction of price transaction of products in the financial, commodities or any other markets.There are two major classes of indexes in usefirstly equally weighted price index in this index, is calculated by taking the average of the prices of a set of companies and secondlymarket capitalization weighted index, in this index, each of the N companies' prices is weighted by the market capitalization of the company.

3.3 Stock Broker

Prior to comprehend about what the stock brokers do, you need to know who is a stock broker. Whenever you want to buy or sell shares from the stock exchange you have to do so via registered stock brokers. Astock broker is a professional individual who carries out buy and sell orders for stocks and other securities via a stock market, or over the counter, for a fee or commission. Stockbrokers are commonly allied with a brokerage firm and manage transactions for retail and institutional customers. A stock broker acts an agent to investor and hasexclusivelyperformed a service for the investor. This means that the broker will buy for the buyer and sell for the seller, each time making certainthat the best price is acquired for the client [34]. Pursuant to an investor, a stockbroker is believed as the one who provides a valuable service and information to help in making the right investment decision.There are two types of stock brokers. Firstly the full-time service brokers are traditional brokers offering trading (commodities, stocks and currency), research and advisory, investment banking, sales and asset management under one roof. They also permit investing in Mutual Funds, IPOs, Forex, FDs, Bonds and Insurance. In some incident full-service brokers also have in-house banking and demat account services. Secondly the discount brokers offer low brokerage, high speed and the state-of-the-art execution platform for trading in stocks, commodities and currency derivatives. They offer a no-frills brokerage service for do-it-yourself traders who comprehend the market well. They are also known as online brokers who offer savings of 75 to 90% on brokerage.

3.4 What Role Does It Play

The role of stock markets is a vital one, as the portal to funding private business. Stock markets assent companies access to pools of private capital, which can be used to fund business rise and development, or the purchase of new assets. In the absence of access to private funding, businesses would be more constrained in the projects they can contemplate funding and would because be unable to truly capitalize on the equity in their company. By the way, stock markets make it practicable for business owners to ultimately cash out their positions for a profit, by selling their shares in the open market. The economic gain of stock exchanges and stock trading are untold, and without the facility in place for trading equity in this way, companies would quest expansion to a greater extent challenging.For individuals and traders, stock markets also have avitalrole to play in the sense that they assent traders to generate a return on their capital, such that it is appropriate for capital to be invested in the financial markets. This dual functionality makes it essential for both traders and companies as a means of facilitating the capital effluent.

3.5 How Do TheFinancial Markets Work

The stock markets work because sellers and buyers are willing to engage in trade in an asset. When stocks are purchasing, their value growsincrementally on a per share basis to reflect the growing demand for that asset. When shares are sold, the value is depressed by an equal amount, alike to the influence of the share sale on the total market price at any given time. This creates the mild price curve that markets tend to see, where the overarching trends cycle (unless the is underlying growth or depression in the value of the company), allowing traders to purchase low and sell high in where the chance exist to do so.

3.6 How Can It Be Traded For Profit

In order to trade shares fruitfully, theirrequirementfirstly is capital growth in the size of a position. This is attained by trading in markets that are expected to increase, or by shorting markets that you expect to slide, and any difference between the closing and opening value can represent a capital gain or loss on the particular position. Share dealers are fortunate in being able to gain in a secondary way, therewith to this level of capital growth, through the dividend yield. Dividends are shares in company profits paid out to shareholders at regular intervals ubiquitously the year, and those traders that hold dividend paying shares when the dividend is declared are entitled to receive the payment. This creates two distinct layers of fruitfully for traders engaging in stock market trading.

3.7 Stock Exchange

A stock exchange is a corporation or mutual organization which endues tradingconveniencefor stock brokers to trade in stocks and other securities. The securities traded on a stock exchange comprise shares issued by companies, unit trusts, derivatives, pooled investment products and bonds [1]. The primary function of a stock exchange is to facilitate the transactions associated with both buying and selling of securities. Afterwards, sellers and buyers of shares and stocks can track the price changes of securities from the stock markets in which they operate. Stock exchanges may also ensuefacilities for the issue and redemption of such securities and instruments and capital events comprise the payment of income and dividends. A number of major stock exchanges globally and each of these plays a part in determining the overall economic condition [32]. In the end, stock exchanges have several roles in the economy, this may contains raising capital for businesses, government capital-raising for development projects, facilitating company rise, profit sharing, corporate governance, creating investment opportunities for small investors, mobilizing savings for investment, barometer of the economy.

3.8Stock Market Prediction and Classification

Theefficient-market hypothesis(EMH) affirm that financial markets are informationalefficient, or that prices on traded assets (e.g., stocks, bonds, or property) already reflect all known information, and immediate modification to reflect new information. Hereupon, according to theory, it is not possible to consistently outperform the market byusing any information that the market already knows, except via luck. Knowledge or newsin the EMH is defined as anything that may affect prices that is unknowable in the present and thus become visible randomlyin the future. Stock market prediction brings withit the challenge of vindicate whether the financial market is predictable or not, since there has been no consensus on the validity of Efficient Market Hypothesis (EMH). The Efficient Market Hypothesis (EMH) is an application of 'Rational Expectations Theory' where people who enter the market [35], use all available & relevant information to make better conclusions. The only caveat is that information is costly and arduous to get.This Efficient Market Hypothesis implies that stock prices reflect all available and pertinent information, so you can't outguess the market or systemically beat the market. This means it's not possible for investors to either purchase undervalued stocks or sell stocks for inflated prices.Subsequently, the Random Walk Theory or the Random Walk Hypothesis is a mathematical model of the stock market.The random walk theory is the phenomena of an incident determined by a series of random movements, in other words, the incident that cannot be predicted.The random walk theory recommendsthatvariations in stock prices have the same distribution and are independent of each other, hereupon, the past movement or trend of a stock price or market cannot be used to predict its subsequent movement. In short, this is the opinion that stocks take a random and unpredictable path [36].Random walk theory demands that it is not possible to predict which way prices will go in the world of investments. In this scenario, shares and some other financial assets follow a random walk. In other words, it is infeasible to know whether the next price movement will be up or down, or how steeply that rise or fall might be.Pursuantto many economists, this volatility means that investors will never be able to outperform the market conformable.Feweconomists, although, disagree and claim that asset prices do follow a non-random walk, and they can be predictable. They emphasize that it is possible to outperform the markets consistently.According to the random walk theory, stock price variations have the same distribution and are completely independent of one another.Hereupon, it is not feasible to use the past trends to predict where a market will go.

IV. Deep Learning

The field of artificial intelligence is essentially when machines can do tasks that typically need human intelligence. It encompasses machine learning, where machines can learn by experience and gain skills without human participation. The deep learning is a subset of machine learning where artificial neural networks, the algorithms influence of the human brain, learn from huge amounts of data [37].Unlike conventional machine learning algorithms, many of which have a finite capacity to learn no matter how much data they obtain, deep learning systems can ameliorate their performance with access to more data [9]. Deep learning permits machines to solve sophisticated issues,even when using a data set that is very diverse, unstructured and interconnected.Any Deep neural network will be made up of three types of layers first input layer, second hidden layer and third the output layer shown in figure 1. The deep learning is a Neural Network be made up of a hierarchy of layers, whereby each layer transforms the input data into more abstract representations. These sequencecyof layers, between input and output, recognize the input features and create a series of new features based on the data, just as our brain. In deep learning the several layers a network has, the higher the level of features it will learn. The output layer amalgamatesall these features and makes a prediction.

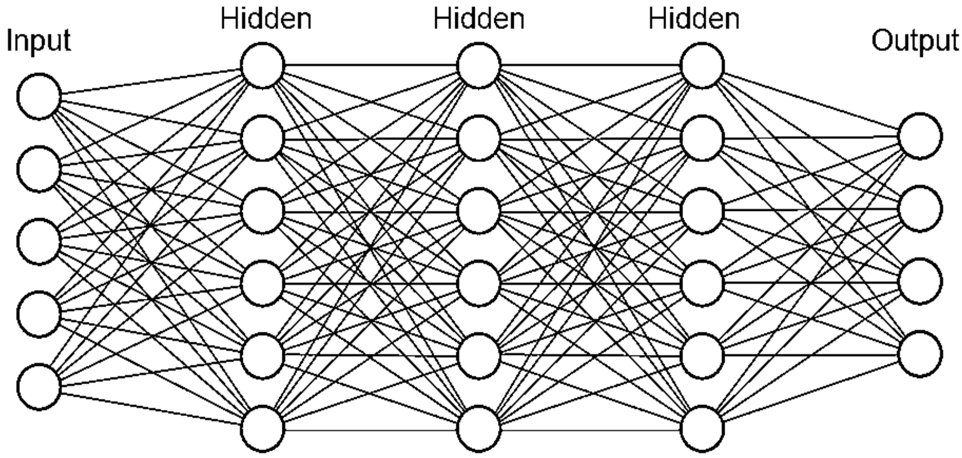

Figure 1. The Deep Learning

In the above diagram, the first layer is the input layer which receives all the inputs and the last layer is the output layer which endues the desired output. All the layers in between these layers are called hidden layers [38]. There can be n number of hidden layers. The number of hidden layers and the number of perceptrons in every layer will entirely depend on the use-case we are trying to solve.

V.How Deep Learning Works

Deep learning networks learn by discovering sophisticated structures in the data they experience. By building computational models that are composed of several processing layers, the networks can create several levels of abstraction to represent the data [39]. In an endeavor to re-engineer a human brain, deep learning studies the basic unit of a brain called a brain cell or a neuron shown in figure 2. Inspired from a neuron an artificial neuron or a perceptron was developed [40]. At the moment, let us comprehend the functionality of biological neurons and how we mimic this functionality in the perceptron or an artificial neuron.If we focus on the structure of a biological neuron, it has dendrites, which are used to obtain input. These inputs are summed in the cell body and using the Axon it is passed on to the next biological neuron as shown in the below image. Correspondingly, a perceptron receives several inputs, applies different transformations and functions and provides an output [40]. As we know that our brain consists of several connected neurons called neural network, we can also have a network of artificial neurons called perceptrons to form a deep neural network.

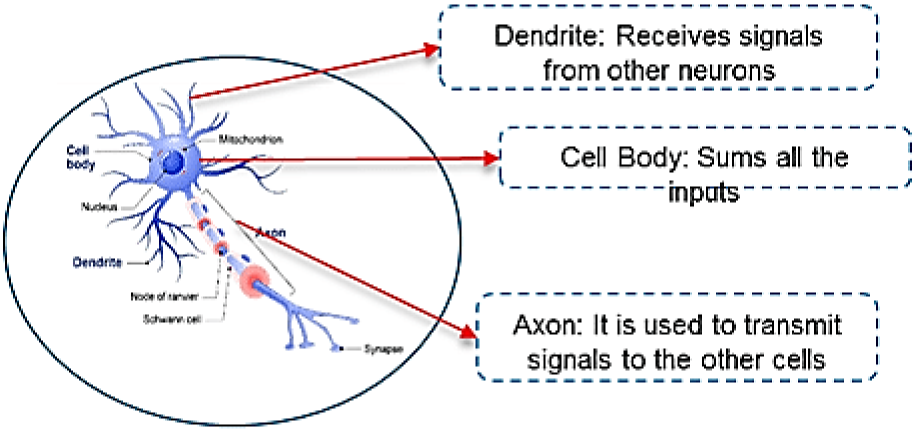

Figure 2. The Neuron

VI. Deep Learning Model for Bombay Stock Exchange Prediction

A prediction or forecast,is a statement about a future incident [41].Thestock market prediction is the act of an attempt to determine the future price of a company stock or other financial instrument traded on an exchange. The victorious prediction of a stock's future price could yield considerable profit. A deep neural network is a neural network with a certain level of insolubility, a neural network with more than two layers. Deep neural networks use intricate mathematical modeling of process data in complex ways.Deep learning eliminatesthe manual identification of features in the data and, alternatively, relies on whatever training process it has in order to explore the useful patterns in the input examples.

4.1. The Structure

For everyBombaystock, we seek a predictor function f in order to predict the Bombay stock return at time t + 1, r_{t+1},given the features fu_t extracted from the information acquiredat time t [42]. We assumethat r_{t+1} can be decomposed into two parts, firstly the predicted output $\hat{r}_{t+1} = f\ 0\ \varphi(R_t)$, and secondly the unpredictablepartγ, which we regard as Gaussian noiselike

$$r_{t+1} = \hat{r}_{t+1} + \gamma$$

$$\gamma \sim N(0, \beta)$$

Where $N(0,\beta)$ indicate a normal distribution with zero mean and varianceβ. The representation fu_tcanbe either a linear or a nonlinear transformation of the raw level information R_t. Denoting the variation function by ϕ, we have

$$ut = \phi\,(R_t)$$

$$\hat{r}_{t+1} = f\ 0\ \phi(R_t)$$

For our experiment, we define the raw level information as the past returns of the Bombay stocks in our sample. If there are S stocks in the sample and g lagged returns are selected, R_t will have the form

$$R_t = [r_1,t,\ldots\ldots\ldots r_{1,t-g+1},\ldots\ldots\ldots\ldots r_{M,t,}\ldots\ldots\ldots\ldots r_{,S,t-g+1}]^T \epsilon R^{Mg}$$

Wherer_i t indicate the return on Bombay stock i at time t.

4.2. Deep Neural Network

A Neural network specifies the nonlinear relationship between two variables $V_l$andV_{l+1}through a network function, which typically has the form

$$V_l+1 = \delta\,(WV_l + b)$$

Where δ is called an activation function, and the matrix *W* and vector *b* are model parameters. Thevariables$V_l$andV_{l+1} are said to form a layer, when there is only one layer between the variables,theirrelationship is called a single-layer neural network [15]. The Multilayer neural networks, augmented withadvanced learning methods are commonly referred to as deep neural networks [42]. A deep neural networks for thepredictor function, *y = f(u),* can be constructed by serially stacking the network functions as follows.

$$V_l+1 = \delta_1(W_1\, u + b_1)$$
$$V_l+1 = \delta_2(W_2 V_1 + b_2)$$
$$y = \delta_L(W_L V_{L-1} + b_L)$$

Where*L* is the number of layers.

Given a Bombay stock dataset $\{u^n, T^n\}^N_{n=1}$ of the inputs and targets, and an error function $\varepsilon(y^n, T^n)$ that measures the distinction between the output $y^n = f(u^n)$ and the target T^n, the model parameters for the entire network, $\theta = \{W_1, \ldots\ldots\ldots\ldots W_L, b_1, \ldots\ldots\ldots\ldots b_L\}$, can be selected so as to minimize the sum of the errors.

$$\min \theta \left[J \sum_{n-1}^{N} \varepsilon\left(y^n,\ T^n\right) \right]$$

Given an appropriate selected of ε(.) Its gradient can be acquired analytically through error backpropagation [43]. In this case, the minimization issue in the above can be solved by the usual gradient descent method. Anarchetypalchoice for the objective function that we take in this paper has the form like

$$J = \left[\frac{1}{N} \sum_{n-1}^{N} \left\| y^n,\ T^n \right\|^2 + \lambda . \sum_{l-1}^{L} \left\| W_l \right\|_2 \right]$$

Where $\|.\|$ and $\|.\|_2$ respectively denote the Euclidean norm and the matrix L_2 norm. The second term is a regularizedadded to keep away from overfitting, while λ is a user-defined coefficient.

4.3. Training Criteria

The training criterion should be easy to assess and be highly correlated to the final target of the Bombay stock task, so that the reformation in the training criterion would lead to the reformation in the final assess score. Preferably, the model parameters should be trained to minimize the expected loss

$$MP_{EL} = E(MP\ (W, b; o, y)) = \int_0 MP\ (W, b; o, y)\ p\ (o)\ d\ (o)$$

Where MP (W, b; o, y) is the loss function given the model parameters {W, b}, the observation o, and the corresponding output vector y, and p (o) is the probability density function of observation o. Unluckily, p (o) is typically uncertain and needs to be estimated from the training set, and MP(W, b; o, y) is not well-defined (the desired output vector is uncertain) for samples inconspicuous in the training set. For this reason, the deep neural networkmodel parameters are frequently trained to optimize the empirical standard.

4.4. Model Initialization

Thedeep neural networks are a highly nonlinear model and the training criterion with respects to the modelparameters is nonconvex, the initial model can greatly impress the resulting model.There are many heuristic tricks in initializing the deep neural networkmodel. Most of theseadroitnessis based on two considerations. In the first instance, the weights should be initialized sothat each neuron operates in the linear range of the sigmoid function at the start ofthe learning. If weights were all very huge, many neurons would saturate (close tozero or one) and the gradients would be very small.Whenthe neurons operate in the linear range, alternatively, the gradients are huge enough (close to the maximum value of 0.25) that learning can proceed efficaciously. Bear in mind that the excitation value depends on both the input values and the weights. In the second instance, it is vital to initialize the parameters arbitrarily. This is because neurons in the deep neural networks are symmetric and interchangeable. If all the model parameters have identical values, all the hidden layer neurons will have the same output value and detect the same feature patterns in the lower layers. Arbitrarily initialization serves the objective of symmetry breaking.TheLeCun at al. [44] indicated to draw values from a zero-mean Gaussian distribution with standard deviation $\sigma \mathbf{W}^{\ell+1} = \frac{1}{\sqrt{FN_\ell}}$ to initialize the weights in layer ℓ , where $FN_\ell$is the number of connections feeding into the node.

4.5. Dropout

The fundamental thought of dropout is to randomly omit a certain percentage of the neurons in each hidden layer for each presentation of the samplesduring training. This means during the training each random combination of the(1− ζ) remaining hidden neurons necessity to perform well even in the absence of theomittedneurons.Thisdemandeach neuron to depend less on other neurons to detectpatterns [45].On the other hand, dropout can be considered a technique that adds random noise tothe training data. This is because each higher-layer neuron gets input from a randomcollection of the lower-layer neurons. The stimulation received by each neuron is different, even if the same input is fed into the deep neural networks. In the company of dropout, deep neural networks need towaste some of the weights to eliminate the effect of the random noise introduced. Assuch, dropout essentially eliminatesthe capacity of the deep neural networks, and thus can make a better generalization of the resulting model.When a hidden neuron is dropped out, its activation is set to 0 and so no error signal will pass via it [46]. This means that other than the random dropout operation, no other modification to the training algorithm are needed to implement this feature. Atthe test time, however, instead of using a random combination of the neurons ineachhidden layer, we use the average of all the possible combinations. This can be effortlessly achievedby discounting all the weights involved in dropout training by (1− ζ), and then using the resulting model as a normal deep neural networks.During the dropout training, we need to again and again pattern a random subset of activationsat each layer. Thiswouldretardation the training notably. For this reason,the speed of the random number generation and sampling code is critical to decrease the training time.

4.6. Pattern Randomization

The pattern randomization is beside the point to the batch training since all patterns are used to estimate the gradient. If the entire training set can be loaded into thememory, pattern randomization canbecomfortably done by permuting an index array.Patterns can then be drawn one by oneaccording to the permuted index array. Since the index array is typically much smallerthan the features, this would cost less than permuting the feature vectors themselves,particularlyif each data pass needs a dissimilar randomization order. This adroitness alsoassuresthat each pattern will be presented to the training algorithm once for eachdata pass, and thus will not influence the data distribution. This property will assure that the model learned is coherent.

4.7. Momentumt

It is well-known that the convergence speed can be better if the model update isbased on all the foregoing gradients as an alternative only the current one. Nesterov's accelerated gradient algorithm [47], which is proved to beoptimal for the convex condition, is an example of this adroitness. In the deep neural networktraining, thisis typically attained with a convenient technique named momentum.Where ρ is the momentum factor, which typically takes value of 0.9 to 0.99 whenstochastic gradient descentorminibatch training is used.The momentum tranquil the parameter updateandbring down the variance of the gradient estimation.

4.8. Training Deep Neural Networks

In this section, we build the predictor functions, $\hat{r}_{i,t+1}$ = fi(fut); i = 1,……………..,S using deep neural networks, andequivalence,their predictive performance with a univariate autoregressive model with ten lagged variables**.** We make use of a three-layer network model of the form.WhereCoLU is the corrected linear unit activation function defined as CoLU(x) = max(x; 0), with max being an element-wise operator. CoLU is known to endow a much faster learning speed than standard sigmoid units, while maintaining or even make better the performance when applied to deep neural networks [48].

$$V_1 = \mathrm{CoLU}(W_{1fut} + b_1)$$
$$V_2 = \mathrm{CoLU}(W_2 v_1 + b_2)$$
$$\hat{r}_{i,t+1} = (W_3 V_2 + b_3)$$

VII. Learning Methods

7.1 Particle Swarm Optimization (PSO)

Particle swarm optimization works with a set of presumable solutions and constraints on an optimization issue. The optimization issue has to have a target condition then the algorithm works, to solve the issue and provide the optimal values. The particle swarm optimization was developed in 1995 by Russell Eberhard and James Kennedy [15]. These researchers started out to stare at computer simulations of bird flocking, and then worked to surpassing the algorithm based on this research.ThePSO algorithm has a number of desirable properties, containingenuityof implementation, scalability in dimension, andbetter empirical performance, it has been applied to solve many real-world issue, such as short termload forecasting [49], and the soft sensor [50]etc.In particle swarm optimization SP denotes the size of the swarm population. In PSO, each particle j (j =1, . . . , SP) has a position yj = (y_j,1, y_j,2, . . . , y_j,n) in the searchspace and a velocity Vj = (V_j,1, V_j,2, . . . , V_j,n) to gesture itspresentstate. A position y_i denotes a presumable solution. Thepositiony$i$ and the velocity Vj are updated with the best positionbp_j = (bp_j,1, bp_j,2, . . . , bp_j,n) encountered by the particle so far andtheoptimal position bpg = (bp_g,1, bp_g,2, . . . , bp_g,n) found by the entirepopulation of particles according to the following equation.

$$V_{j,}\,(t+1) = \omega V_{j,}\,(t) + lf_1 r_1\,(bp_j,\,(t) - y_j,\,(t)) + lf_2 r_2\,(bp_g,d\,(t) - y_j,d\,(t))$$

$$y_{j,d}(t+1) = y_{\,j,d}\,(t) + V_{j,d}(t+1)$$

Where lf_1 and lf_2 are two learning factors which control the impact of the social and cognitive components, rj (j = 1, 2) are random numbers in the range [0, 1] and ω is the inertia weight, which make sure the convergence of the PSO algorithm and is decreased linearly.When PSO is incorporated to update parameters of deep neural networkstructure, the dissimilar parameters of deep neural networks,including the mean, variance and weights of the output layer are to be updated. These parameters as a whole be regarded as a particle and each particle searches for the solution by investigation the fitness function. PSO is initialized with a group of random particles and then discovery for optima by updating generations. In every iteration, each particle is updated by following two optimal values. The first one is the best solution (fitness) it has achieved so far after that best value is also stored. This value is called pbest [51]. Another best value that is tracked by the particle swarm optimizer is the best value, acquired so far by any particle in the population. This optimal value is a global best and called gbest. When a particle takes part of the population as its topological vicinal, the best value is a local best and is called lbest.

7.2 Least Mean Square

The Least Mean Square (LMS) algorithm, presented by Widrow and Hoff in 1960 [52] is an adaptive algorithm, which uses a gradient-based method of steepest decent. The LMS algorithm uses the estimates of the gradient vector oftheobtainadata. LMS includesan iterative procedure that makes successive reformation to the weight vector in the direction of the negative of the gradient vector which eventually leads to the minimum mean square error [53].Differentiate to other algorithms LMS algorithm is relatively uncomplicated, it does not expect the correlation function calculation nor does it expect matrix inversions. The least-mean-square (LMS) algorithm is based on the use of coinstantaneous estimatesof the autocorrelation function r_x(j,k) and the cross-correlation function r_{dx}(k).

$$\hat{r}_x(j,k,n) = x_j(n)x_k(n)$$

$$\hat{r}_{dx}(k,n) = x_k(n)d(n)$$

The use of a hat in $\hat{r}_x$ and $\hat{r}_{dx}$ is intended to signify that these quantities are evaluated. Anonstationaryenvironment,in which case all the sensory signals and the desired reaction assume time-varying forms too. Thus, use the instead of $\hat{r}_x(j,k,n)$ and $\hat{r}_{dx}(k,n)$ n place of $r_x(j,k)$and r_{dx}(k)now we get

$$\hat{w}_k(n+1) = \hat{w}_k(n) + \eta\left[x_k(n)d(n) - \sum_{j=1}^{p}\hat{w}_j(n)x_j(n)x_k(n)\right]$$

$$= \hat{w}_k(n) + \eta\left[d(n) - \sum_{j=1}^{p}\hat{w}_j(n)x_j(n)\right]x_k(n)$$

$$= \hat{w}_k(n) + \eta\left[d(n) - y(n\right] x_k(n)$$

Where k = 1,2, 3,4, . . . , P and y(n)is the output of the spatial filter computed at iteration nin conformity with the LMS algorithm; that is,

$$y(n)= \sum_{j=1}^{p} \hat{w}_j(n) x_j(n)$$

We have used $\hat{w}_k(n)$ in place of $w_k(n)$ to emphasize the fact that include estimates of the weights of the spatial filter.In the technique of steepest descent applied to a familiar environment, the weight vector w(n),made up of the weights $w_1(n)$, $w_2(n)$,. . . , $w_p(n)$, starts at some initial value w(O),and then follows a precisely defined trajectory (along the error surface) that lastly terminates on the optimum solution w_o,endowthat the learning-rate parameter η is selected decently [53]. In contrast, in the LMS algorithm applied to an unfamiliar environment, the weight vector $\hat{w}(n)$, representing an estimate of w(n),follows a random trajectory. For this cause, the LMS algorithm is sometimes referred to as a stochastic gradient algorithm. As the number of iterations in the LMS algorithm approaches infinity, $\hat{w}(n)$ performs a random walk (Brownian motion) about the optimum solving a problem w_o.Thedeep neural network structure has the following parameters used forBombay stock. Firstly the n centres of deep neural networks where there are n nodes in the hidden layer, and secondly the n variance of deep neural networks where there are n nodes in the hidden layer and eventually the weights in the output layer. It is also vital to recognize that the LMSalgorithm can operate in a stationary or nonstationary environment.

VIII. Data Preprocessing and Input Selection

The data for the stock market prediction test has been collected for thestock indices, namely Bombay Stock Exchange(BSE). The time series data of all the stock indices were collected from 5^{th} October 2018 to 5^{th} November2018. Thus, therewere 2000 data sample ofthe Bombay Stock index. The data collected for the stock indicesconsisted of the ending price, opening price, and the lowest value in the day, highestvalueon the day and the total volume of stocks traded in each day [54]. The proposed prediction model is developed forpredictingtheending price of the index in each day of the prediction period of Bombay Stock Exchange.The various technical and fundamental indicators are used as inputs to the network.Technical indicators are any class of metrics whose value is derived from genericprice activity in a stock or asset. Technical indicators look to predict the futureprice levels [55], or simply the general price direction, of a security by looking at pastsampleof Bombay Stock Exchange.The details of the parametersand how they are calculated from the obtainable data is given below.

8.1 Simple Moving Average

A simple moving average (SMA)is an arithmetic moving average calculated by adding recent ending prices and then dividing that by the number of time periods in the average calculation. A simple, or arithmetic, moving average that is calculated by adding the ending price of the security for a number of time periods and then dividing this total by that same number of periods.The various SMAs used in the test are.

$$\frac{1}{ND}\sum_{i=1}^{ND} TP_i$$

ND is a number of days and TP_i is present price.

1 days (SMA1)
10 days (SMA10)
30 days (SMA30)
60 days (SMA60)

8.2Accumulation &Distribution Oscillator

Accumulation &distribution are a momentum indicator that efforts to assess supply and demand by determining whether investors are generally buying (accumulating) or selling (distributing) a few stocks. The Accumulation& distribution measure seeks to identify the difference between stock price and volume flow. The Accumulation& distribution of a security are calculatedfrominitialcalculating the cash flow multiplier and then multiplying the money flow multiplied by the period's volume.

ADO = ((Close – Low) – (High – Close))/ (High – Low) * Period's Volume

Each and every day's ADO has been taken during the experiment.

8.3On Balance Volume

The On Balance Volume (OBV) measures buying and selling pressure as a cumulative indicator that adds volume on up days and subtracts volume on down days. The On Balance Volume (OBV) line is straightforwardly a running total of positive and negative volume. A period's volume is positive when the close is above the foregoing close. A period's volume is negative when the close is below the foregoing close. To utilize OBV requires acomprehensionof how the indicator operates, how it can be used to aid trading verdict, and its restrictions.

If Present's close > Yesterday's Close
OBV = Yesterday's OBV + Present's Volume
If Present's close > Yesterday's Close
OBV= Yesterday's OBV – Present's volume

8.4Exponential Moving Average (EMA)

An exponential moving average (EMA) is a type of moving averagethat places a greater weight and importance of the most recent data points. Moving averages lag because they are based on bygone prices. In spite of this lag, moving averages help smooth price action and filter out the noise. The exponential moving average is also designated to as the exponentially weighted moving average. An exponentially weighted moving average responsemorenotably to recent price modify than a simple moving average (SMA), which enforce an equal weight to all observations in the period.

EMA = (ending price - previous day's EMA) * smoothing constant + previous day's EMA
Where the smoothing constant is2/(number of time periods + 1)

8.5Williams %R

The Williams %R is a momentum indicator that is the inverse of the Fast Stochastic Oscillator. Also indicated to as %R, Williams %R reflects the level of the close relative to the highest high for the look-back period. In contrast, the Stochastic Oscillator reflects the level of the ending relative to the lowest low. The %R rectifiesfor the inversion by multiplying the raw value by -100. As an outcome, the Fast Stochastic Oscillator and Williams %R produce the appropriate same lines, only the scaling is dissimilar. Williams %R oscillates from 0 to -100. Readings from 0 to -20 are considered overbought. Readings from -80 to -100 are considered oversold. Amazingly, signals derived from the Stochastic Oscillator are also applicable to Williams %R.

Calculation of Williams %R = (Highest high in n periods – Present's close)*100 (Highest high in n-periods – Lowest low in n-periods)
For this experiment n= 9 days

8.6Price Rate of Change (PROC)

The price rate of change (PROC) is a technical indicator of momentum that measures the percentage transformation in price between the present price and the price n periods in the past. The PROC calculation compares the present price with the price "n" periods ago. The plot forms an oscillator that fluctuates above and below the zero line as the price Rate-of-Change transmigration from positive to negative. As a momentum oscillator, PROC signals contain centerline crossovers, divergences and overbought-oversold readings.

(Present's close – Close x-periods ago) * 100 (Close x-periods ago)

IX. Model Setup Using Deep Neural Network

The data pattern is taken from stock exchange data and data pattern is collected from the historical values of Bombay Stock Exchange data.The deep neural network function has a hierarchy ofhidden layers having 12 centres the activation function is a Gaussian-Bernoulli RBM one which depends on the gradient descent which is gradient descent of inputpatternfrom the designated centres.Thus the first layer is a hierarchy dependancy.The resultant is multiplied by a weight corresponding to each centreand all of these are summed up to give a value which is called the plant output.Thedeep neural network function has thus had12centres (12 *12 = 144 weights, as 12 input parameters are being fed into the deep neural network), 12 variances and 12hierarchy weightscorresponding to the 12centres for each input pattern.These parameters are trainedusing either particle

swarm optimization and least mean squarealgorithm.The total data set of a particular Bombay Stock market index is fragmented up into two, one forthe training of the network and the rest for testing the performance of the network afterfreezing the weights. In this experiment we take approx 2000 to 3000 daily
statistical data on the Bombay stock index as training set. The remainder 1200 values are set aside for testing.

X. Training and Testing Process for Bombay Stock Exchange

10.1 TrainingProcess for Bombay Stock Exchange

In deep neural network model, 12 *12 = 144 weights, (associated with mean), 12 for a variance and 12 as hierarchy weight constitutes a solution to the model. All these weights are improved through particle swarm optimization and least mean square algorithm. The input data set is also normalizedbeforethe network training. The weightsabide static till all of the training data set is fed into the network, equivalence with the desired output and their respective error stored. The mean error for theentire epoch is numerated, and then the adaptive weight update takes place. TheLeast Mean Square (LMS) update algorithm and PSO is used in our experimenttoupgradethe weights by adding the product of the convergence constant, theparticularinput with the mean error for the epoch to the weights of the preceding epoch. The price function of the training process is the Mean Square Error (MSE). It is appropriate to end the training of the network when the minimum level of the costfunction is observed. Consequently, for each iteration (epoch), the mean square error iscalculated and plotted. Each of the iterations involves training the network with the5000-odd patterns, calculation of mean error, weight update and representing theMSE. The number of iterations is determinative upon by the gradient of the MSE curve. If itis observed that there is no vital decrease in the MSE then the trainingexperiment can be stopped. There happena trade-off between the time taken andstandardof training. The high number of iterations inclinesto give a superior level of training thenetwork at the cost of time taken to train.ThePSO is used to train parameters of the structure using particles. We have used 40particles whereby each particle symbolizesa solution to the issue. Theseparameters as a whole symbolize one particle and each particle searches optimal solution to the issue. Each particle has a fitness value associated with it and it isanerror in our test.Themain objective is to make each particle search for the optimal fitsolution and in turn minimize error. In this process each particle learns from itspast solutions and determines the pbest- the best position for each particle in aspecificiteration.

10.2 Testing Processfor Bombay Stock Exchange

At the conclusion of the training process of the network, the weights are frozen for testingthe network on inputs that were set apart from the training set. The testing setsampleis the input to the network and the output, the predicted index close priceisdifferentiatedwith desired output or genuine close price. The percentage of mistakesarerecorded for every data set. TheMean Absolute Percentage Error (MAPE) is used to assess the performance of thetrained prediction model for the test data. Mean Absolute Percent Error (MAPE) is the most general measure of forecast error. MAPE functions best when there are no extremes in the data (including zeros). It is quite dissimilar from normal MSE, asevident from the equation below. In our simulation, we have calculated both MSEand MAPE, but the analysis and differentiate is done on the basis of MAPE only.Theattempt is to minimize the MAPE for testing samplesin the quest for discovery apreferablemodel for forecasting stock index price movements. The MAPE is the average absolute percent error for each time period or prophecy minus genuine divided by genuineprice.

$$\frac{1}{ND}\sum_{i=1}^{ND}\frac{\left|F_i - A_i\right|}{A_i}$$

XI. The Outcome for ActualBombay Stock Exchange Closing Value

In this experiment we take approx 2000 to 3000 day to day statistical data on the Bombay stock index 1000 as a training set, it was found that only 2000 days data is sufficient enough to train the models for 1 day, 10 days, ahead prediction. For 30 and 60 days ahead prediction, upto 3000 days data is used to train the network. Thismodel is tested with fresh 1200 days data out of which only 100 are shown for legibility. We have used 40 particles whereby each particle depictsa solution to the difficulty. In the first part testing of deep neural network parameters tuned with least mean squaremodel and second part testing of deep neural network parameters tuned with particle swarm optimization model shown in below section.

11.1 Testing of Deep Neural Network Parameters Tuned With Least Mean Square Model

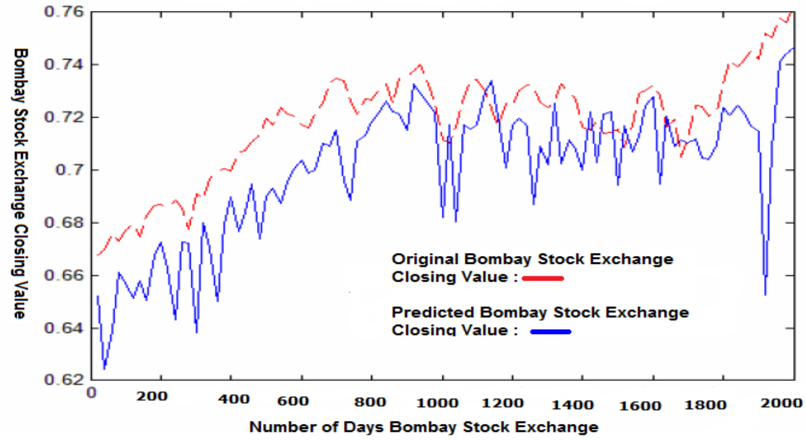

Figure 3. The First Days Predicted Bombay Stock Exchange Closing Value WithLeast Mean Square

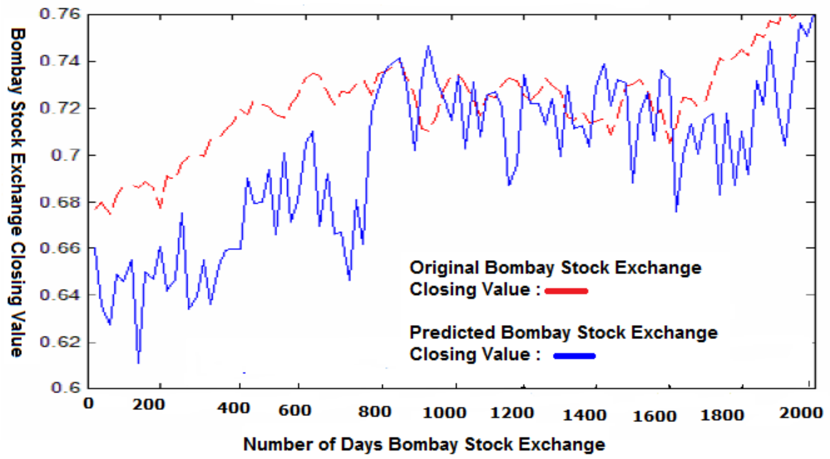

Figure 4. The Ten Days Predicted Bombay Stock Exchange Closing Value With Least Mean Square

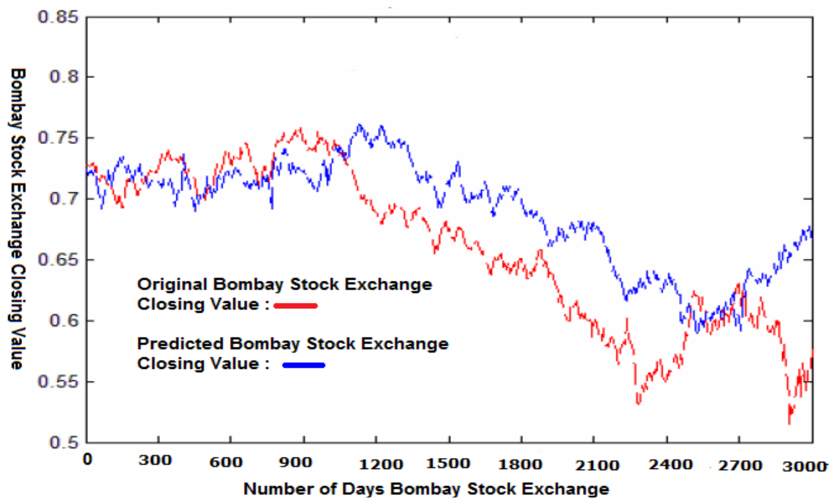

Figure 5. The Thirty Days Predicted Bombay Stock Exchange Closing Value With Least Mean Square

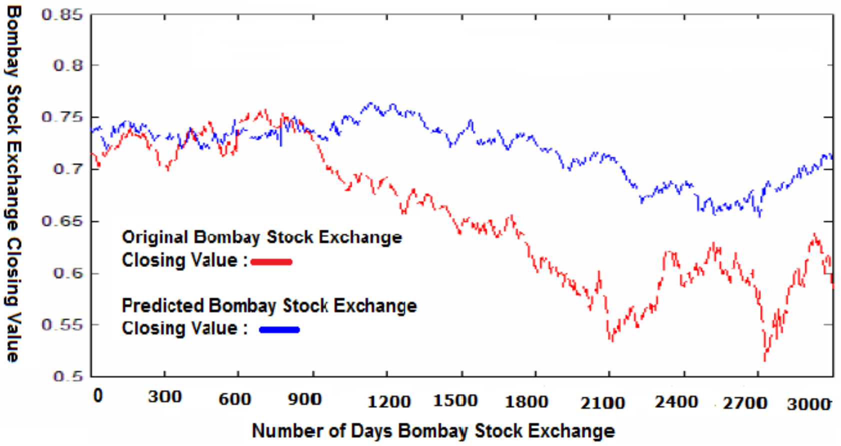

Figure 6. The Sixty Days Predicted Bombay Stock Exchange Closing Value With Least Mean Square

11.2 Testing of Deep Neural Network Parameters Tuned With Particle Swarm Optimization Model

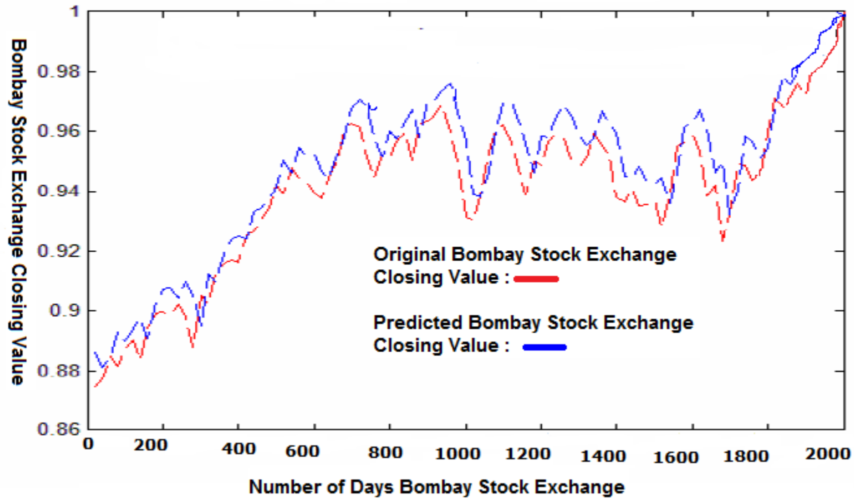

Figure 7. The First Days Predicted Bombay Stock Exchange Closing Value WithParticle Swarm Optimization

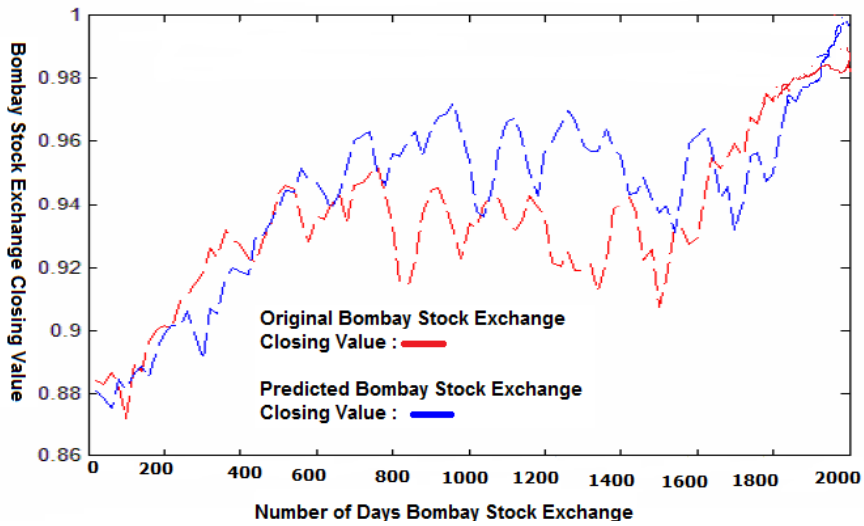

Figure 8. The Ten Days Predicted Bombay Stock Exchange Closing Value With Particle Swarm Optimization

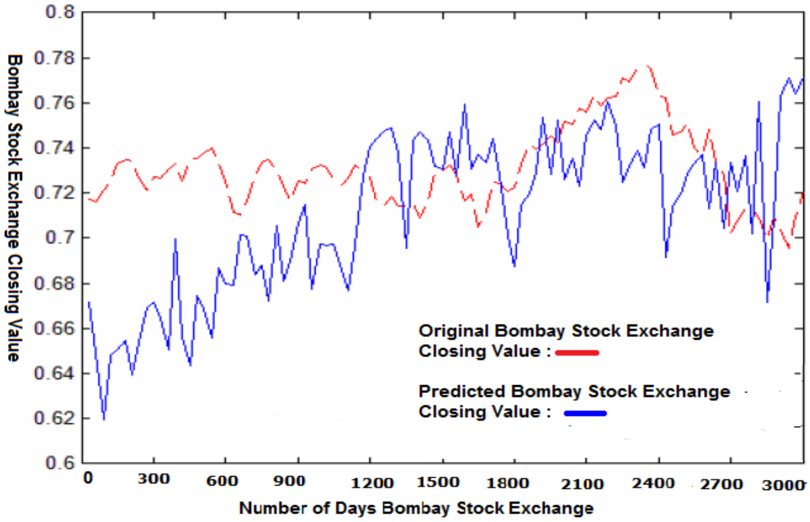

Figure 9. The Thirty Days Predicted Bombay Stock Exchange Closing Value With Particle Swarm Optimization

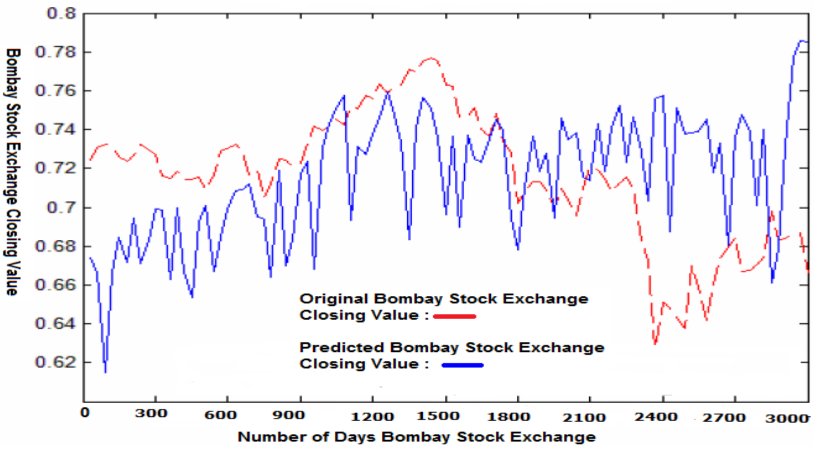

Figure 10. The Sixty Days Predicted Bombay Stock Exchange Closing Value With Particle Swarm Optimization

11.3 Model Comparison Through MAPE for Bombay Stock Exchange

In this sectionmodel comparison through MAPE for Bombay Stock Exchange between deep neural network parameters tuned with least mean squaremodel and deep neural network parameters tuned with the particle swarm optimization model shown in below table 1.

Table 1. The Model Comparison Through MAPE for Bombay Stock Exchange Between LMS and PSO

Number of Days Prediction for Bombay Stock Exchange	Number of Days Training Completed for Bombay Stock Exchange	Deep Neural Network Parameters tuned by Least Mean Square MAPE (%)	Deep Neural Network Parameters tuned by Particle Swarm Optimization Model MAPE (%)
First day Predicted Bombay Stock Exchange	2000 Days Training Completed for Bombay Stock Exchange	1.494%	1.187%
Ten Days Predicted Bombay Stock Exchange	2000 Days Training Completed for Bombay Stock Exchange	3.986%	3.471%
Thirty Days Predicted Bombay Stock Exchange	3000 Days Training Completed for Bombay Stock Exchange	5.897%	4.534%
Sixty Days Predicted Bombay Stock Exchange	3000 Days Training Completed for Bombay Stock Exchange	8.968%	5.341%

XII. Conclusions

At the present time, predict the stock market does not only have monetary gain, but aspresentlythe stock market is a vital part of a country's economy, understandingit is becoming more and more essential.In this paper, wehave proposed a deep learning algorithm for forecasting of Bombay Stock Exchange. The deep learning algorithms do this via several layers of artificial neural networks, which mimic the network of neurons in our brain. This permits the algorithm to perform several cycles to narrow down patterns and make better the predictions with each cycle. The financial data have been taken from a huge database and has been based on the stock prices onthe preeminent Bombay stock exchange. The deep neural network is the capability to extract features from a huge set of raw data without trust on previous information of predictors. This makes deep learning, especially convenient for stock market prediction, in which promiscuous factors influence stock prices in a complex, nonlinear trend.In this paper, we have formulated a comparison between variouslearning models like PSO and LMS. The differentiation of bothmethodsis done by taking deep neural network models and outcome are compiled. During the higher number of days ahead predication, deep neural network parameters updated with PSO algorithm works the optimal. In the end, this paperproposedpractical insights and potentially useful directions for further examine into how deep learning networks can be effectively used for stock market prediction.

References

[1]. Yusuf Perwej , Asif Perwej , "Prediction of the Bombay Stock Exchange (BSE) Market Returns Using Artificial Neural Networks and Genetic Algorithms", Journal of Intelligent Learning Systems and Applications (JILSA), Scientific Research Publishing, USA. Vol. 4, No. 2, May 2012, pages 108-119, ISSN Print: 2150-8402 , ISSN Online: 2150-8410, DOI: 10.4236/jilsa.2012.42010

[2]. U. Thissen, R. Van Brakel, A. P. De Weijer, W. J. Melssen, L. M. C. Buydens, "Using support vector machines for time series prediction", Chemometrics and intelligent laboratory systems 69, no. 1, pp. 35-49, 2003.

[3]. Zhang Wenping, Li Chunping, Ye Yunming, Li Wenjie, Eric W.T. Ngai, "Dynamic Business Network Analysis for Correlated Stock Price Movement Prediction", Published by the IEEE Computer Society, 2015,

[4]. T.Helstrom, and K.Holmstrom, "Predicting the stock market". Published as Opuscula ISRN HEV-BIB-OP-26-SE. 1998.

[5]. M.H. FazelZarandi, B. Rezaee, I.B. Turksen, E. Neshat, "A type-2 fuzzy rule-based expert system model forstock price analysis", Expert Systems with Applications, vol. 36, no. 1, pp. 139-154, January 2009.

[6]. Andrew Beattie. The Basics Of Business Forecasting. Accessed: 2015-05-11, 2011.

[7]. P. Rozycki, J. Kolbusz, B.M. Wilamowski, "Dedicated Deep Neural Network Architectures and Methods for Their Training" in INES'15-Intelligent Engineering Systems, Bratislava, pp. 3-5, Sept 2015.

[8]. P. Kontschieder, M. Fiterau, A. Criminisi, S.R. Bulo, "Deep Neural Decision Forests", IEEE InternationalConference on Computer Vision Santiago, pp. 1467-1475, 2015.

[9]. Yusuf Perwej, "An Evaluation of Deep Learning Miniature Concerning in Soft Computing", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), ISSN (Online) : 2278-1021, ISSN (Print) : 2319-5940, Vol.4, Issue 2, Pages 10 - 16, February 2015,DOI : 10.17148/IJARCCE.2015.4203

[10]. Y. Bengio, "Deep Learning of Representations for Unsupervised and Transfer Learning, " in Proceedings of the Unsupervised and Transfer Learning challenge and workshop, 2011.

[11]. Ko I., Kazutaka S.," Stock Market Prediction from News on the Web and a New Evaluation Approach in Trading", 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), IEEE, Kumamoto, Japan, July 2016DOI: 10.1109/IIAI-AAI.2016.157

[12]. Halbert White," Economic prediction using neural networks: the case of IBM daily stock returns" Department of EconomicsUniversity of California, San Diego.

[13]. K. Kyoung-jae, L. Won Boo."Stock market prediction using artificial neural networks with optimal feature transformation". Neural Computing and Applications (2004),Volume: 13, Issue: 3, Publisher: Citeseer, Pages: 255-260

[14]. Tiffany Hui-Kuangyu and Kun-Huang Huarng, "A Neural network-based fuzzy time series model to improve forecasting", Elsevier, 2010, pp: 3366-3372.

[15]. Asif Perwej, Yusuf Perwej, Nikhat Akhtar, and Firoj Parwej, "A FLANN and RBF with PSO Viewpoint to Identify a Model for Competent Forecasting Bombay Stock Exchange", COMPUSOFT, International journal of advanced computer technology, ISSN: 2320-0790, Volume-IV, Issue-I, pp 1454-1461, January-2015DOI : 10.6084/ijact.v4i1.60

[16]. K. Kim, I. Han. "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index". Expert Systems with Applications, Volume 19, Issue 2, August 2000, Pages 125–132.

[17]. Tong-SengQuah, "Using Neural Network for DJIA Stock Selection ", Engineering Letters, 15:1, EL_15_1_19, 2007.

[18]. Roman J., and Jameel A. (1996). "Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns." In Twenty-Ninth Hawaii International Conference on system sciences 2: 454-460.

[19]. H. Chen, K. Xiao, J. Sun, and S. Wu, "A double-layer neural network framework for high-frequency forecasting," ACM Transactions on Management Information Systems (TMIS), vol.7, no. 4, article no. 11, 2017.

[20]. M. Thenmozhi, "Prediction stock index returns using artificial neural networks", Delhi Business Review X Vol. 7, No. 2, December 2006.

[21]. Hegazy O., Soliman O.S. and Salam M.A. (2014)." A Machine Learning Model for Stock Market Prediction." arXiv preprint arXiv :1402.7351

[22]. JacekMandziuk and MarcinJaruszewicz , "Neuro-evolutionary approach to stock market prediction" , Proceedings of InternationalJoint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007,IEEE.

[23]. J .Štěpánek, ,J. Šťovíček, R. Cimler " Application of Genetic Algorithms in Stock Market Simulation" Cyprus International Conference on Educational Research (CY-ICER-2012)North Cyprus, US08-10 February, 2012.

[24]. Kim K.J. and Han I. (2000). "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index." Expert systems with Applications 19 (2) :125-132

[25]. Emad W. Saad, Danil V. Prokhorov,and Donald C. Wunsch, "Comparative Study of Stock Trend Prediction Using Time Delay, Recurrent and Probabilistic Neural Networks", 1998,IEEE.

[26]. Jia H. (2016). "Investigation into the effectiveness of long short term memory networks for stock price prediction."arXiv preprint

arXiv :1603.07893
[27]. Qui-yong Zhao, Xiaoyu Zhao, Fu Duan , “Prediction Model of Stock Prices Based on Correlative Analysis and Neural Networks” 2009 Second International Conference on Information and Computing Science , 2009, pp: 189-192 , IEEE.
[28]. Ching-Hsuecheng, Tai-Liang Chen, Liang-Ying Wei, “ A hybrid model based on rough set theory and genetic algorithms for stock price forecasting”, 2010, pp. 1610-1629.
[29]. Chiang, W.-C.; Enke, D.; Wu, T.; Wang, R. “An adaptive stock index trading decision support system” , Expert Syst. Appl. 2016, 59, 195–207.
[30]. Feng Li, Cheng Liu , “Application Study of BP Neural Network on Stock Market Prediction”, 2009 Ninth International Conference on Hybrid Intelligent Systems, 2009, pp: 174-178 , IEEE.
[31]. E. Chong, C. Han, and F. C. Park, “Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies,” Expert Systems with Applications, vol. 83, pp. 187–205, 2017.
[32]. Lawrence Shepp, "A Model for Stock Price Fluctuations Based on Information", IEEE Transactions On Information Theory, vol. 48, no. 6, June 2002.
[33]. Laszlo Gerencser, BalazsTorma, ZsanettOrlovits, "Fundamental Modelling of Financial Markets", *ERCIM News*, vol. 78, pp. 52, Jul. 2009.
[34]. Burton G. Malkiel, "The Efficient Market Hypothesis and Its Critics", The Journal of Economic Perspectives, vol. 17, no. 1, pp. 59-82, 2003.
[35]. Fama.E.F: Efficient Capital Market: A Review.of Theory and empirical Work. Journal of Finance, 1970 (25), pp.38-417.
[36]. SaeedB.,Alexander J.,” Random walk sampling for big data over networks”, International Conference on Sampling Theory and Applications (SampTA), IEEE , Tallin, Estonia, July 2017., DOI:10.1109/SAMPTA.2017.8024453
[37]. PariwatOngsulee,”Artificial intelligence, machine learning and deep learning”,15th International Conference on ICT and Knowledge Engineering (ICT&KE), IEEE , Bangkok, Thailand, Nov. 2017DOI:10.1109/ICTKE.2017.8259629
[38]. Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, May 2015
[39]. K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", Proc. CVPR, pp. 770-778, 2016.
[40]. Yusuf Perwej,Firoj Parwej, “A Neuroplasticity (Brain Plasticity) Approach to Use in Artificial Neural Network” International Journal of Scientific & Engineering Research (IJSER), France , Vol.3, Issue 6, June 2012, Pages 1- 9, ISSN 2229 – 5518. DOI: 10.13140/2.1.1693.2808
[41]. Yusuf Perwej , Asif Perwej , “Forecasting of Indian Rupee (INR) / US Dollar (USD) Currency Exchange Rate Using Artificial Neural Network” International Journal of Computer Science, Engineering and Applications (IJCSEA), Academy & Industry Research Collaboration Center (AIRCC) ,USA ,Vol.2, No.2, April 2012, Pages 41- 52. DOI : 10.5121/ijcsea.2012.2204
[42]. Y. LeCun, Y. Bengio, G. Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, 2015.
[43]. Hecht N.,”Theory of the backpropagation neural network”,International 1989 Joint Conference on Neural Networks, IEEE, Washington, DC, USA, USA, 1989. DOI:10.1109/IJCNN.1989.118638
[44]. LeCun, Y., Bottou, L., Orr, G.B., Müller, K.R.: Efficient backprop. Neural Networks: Tricks of The Trade, pp. 9–50. Springer, Berlin (1998)
[45]. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580 (2012)
[46]. Steven Rennie, VaibhavaGoel, Samuel Thomas, "Annealed dropout training of deep networks", Spoken Language Technology (SLT) IEEE Workshop on.IEEE, *2014*.
[47]. Nesterov, Y.: A method of solving a convex programming problem with convergence rate O (1/k2). Sov.Math.Dokl. 27, 372–376 (1983)
[48]. Nair, V., & Hinton, G. E.,”Rectified linear units improve restricted boltzmann machines”, In Proceedings of the 27th international conference on machine learning (icml-10) (pp. 807–814), 2010
[49]. B. Wang, N.-L.Tai, H.-Q.Zhai, J. Ye, J.-D.Zhu, and L.-B. Qi “A new ARMAX model based on evolutionary algorithm and particle swarm optimization for short-term load forecasting,” Electric Power Systems Research, vol. 78, no. 10, pp. 1679–1685, 2008.
[50]. H. Wang and F. Qian, “An improved particle swarm optimizer with behavior-distance models and its application in softsensor,” in Proceedings of the 7th World Congress on Intelligent Control and Automation
[51]. Sheng S.,”A Particle Swarm Optimization (PSO) Algorithm Based on Multi-agent System”,International Conference on Intelligent Computation Technology and Automation (ICICTA), IEEE, Hunan, China, Oct. 2008DOI:10.1109/ICICTA.2008.367
[52]. Widrow, B., & Hoff, M. E. ,” Adaptive switching circuits”, In IRE WESCON ConventionRecord, 1960
[53]. Bernard Widrow, "Thinking About Thinking: The Discovery of the LMS Algorithm", IEEE Signal Processing Magazine, pp. 100-106, 2005.
[54]. Crone Sven, F. NikolaosKourentzes, "Feature selection for time series predictionA combined filter and wrapper approach for neural networks", Neurocomputing, vol. 73.10, pp. 1923-1936, 2010.
[55]. Inteak Kim, Song-Rock Lee, "A Fuzzy Time Series Prediction Method based on Consecutive Values", IEEE International Fuzzy Systems, 1999, vol. 2, pp. 703-707

Dr. Asif Perwej. “An Evolutionary Approach to Bombay Stock Exchange Prediction with Deep Learning Technique” IOSR Journal of Business and Management (IOSR-JBM), Vol. 20, No. 12, 2018, pp. -.63-79.