

# STOCK MARKET PREDICTION AND ANALYSIS USING MACHINE LEARNING

<sup>[1]</sup>Sumedh kapse, <sup>[2]</sup>Rajan kelaskar, <sup>[3]</sup>Rahul kamble, <sup>[4]</sup>Manoj sahu, <sup>[5]</sup>Prof. Atul Shintre

<sup>[1]</sup> <sup>[2]</sup> <sup>[3]</sup> <sup>[4]</sup> Student, PVPPCOE, <sup>[5]</sup> Faculty, PVPPCOE

<sup>[1]</sup> sumedhantipro.32@gmail.com, <sup>[2]</sup> rajankelaskar@gmail.com, <sup>[3]</sup> rk5306903@gmail.com,

<sup>[4]</sup> manoj\_sahu3968@gmail.com, <sup>[5]</sup> atulshintre@gmail.com

**Abstract—** Stock market prediction has always been an interesting research topic among researchers mainly due to its capital gain by trading stocks and or to understand the information hidden in stock market data. Many machine learning algorithms and statistical models have been proposed by researchers for stock price prediction and stock price movement prediction. This paper surveys machine learning methods and techniques for stock market prediction. Here we present recent growth in stock market prediction methods and models, perform a comparison among these models to find out the accuracy of the prediction of the stock market values and also figuring out the advantages and disadvantages of these individual models.

**Index Terms—** Stock Price Forecasting, Machine Learning, Support Vector Machine, Artificial Neural Networks.

## I. INTRODUCTION

Stock market prediction has been one of the most challenging goals of the Artificial Intelligence (AI) research community. This research is meant to go far beyond the capabilities of traditional AI research, which primarily focuses on developing systems that are supposed to emulate human intelligence, because stock market is generally Nonlinear and volatile. The fluctuation rate may depend on many factors: equity, interest rate, security options, warrants merger and ownership of large financial corporations or companies. Still, no one can consistently predict the stock market movement. That is why this of AI prediction requires an iterative process of knowledge discovery and system improvement which can be achieved with various models such as ANN and SVM.

Stock market is a promising financial investment that can generate great wealth. However, the volatile nature of the stock market makes it a very high risk investment. Stock market prediction have always been an important issue in the field of technology, financial and mathematics because of its very likely financial gain. As major amount of capital is traded and exchanged from the stock market, it is considered as highly important investment outlet. Also, the prediction of the stock brings the task of proving if the stock market can be predicted or not, with the huge number of fast computers and tremendous information all over the Internet, stock markets have become more accessible to both the strategic investors and the normal public.

Stock prediction has always been interesting topic for researchers, due to its monetary rewards, complexity and unpredictable nature. With the proliferation of computer technology more calculating and data processes abilities are accessible to individual / institutional researchers. The major motivation for our work is to understand any hidden pattern in the stock market using machine learning algorithms. Secondly to design a prediction model that can continually

predict future stock value with greater accuracy. Thus, make better trading decisions and earn money.

Researchers all around the world have proposed various prediction algorithms which uses historical data for stock forecasting. Our prediction model tries to predict the movement of stock price in the next few days by using historically acquired data of the stock. With this information, an investor can make excellent decisions for better investments so as to achieve good returns.

Our motivation for the project ideas have been because of the importance of stock market to the economy is peerless and the concept of stock exchange and stock trading has always been a fascinating topic to look into, the stock market involves different companies to have their stock listed on stock exchange which makes the stock more liquid and attractive to the investors. The size of the stock market exchanges internationally can be around hundreds of trillions of dollars in the present date. The prediction aspect of the stock market could lead to significant amount of profit for the organization and also considering its importance from industrial point of view we found this topic much suitable and appropriate to work.

## II. LITERATURE REVIEW

Kim and Han [1] also combined NNs with GA and predicted Korea Composite Stock Price Index 200. He achieved 82% of accuracy in predicting both weekly rising and declining stock market tendencies.

Kim and Han (2000) used a genetic algorithm to transform continuous input values into discrete ones. The genetic algorithm was used to reduce the complexity of the feature space.

Brofos [2] proposed an ensemble of random forest classifier, support vector machine and relevance vector machine classifiers, including an ensemble of k-nearest

neighbor classifiers. Using relief-f algorithm for identifying the informative features of the dataset. And implemented with customized Classification Standard finance sectors. This model was made to predict stock price for next quarter, it seems truly justifiable how information gathered in one quarter is not useful for prediction in next quarter.

Qian, Rasheed [3] proposed an ensemble stock market prediction using Artificial Neural Network (ANN), NN, Decision Tree. For ANN they used Levenberg-Margquardt learning algorithm with sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. They used simplest weighted voting and stacking method to combine the classifiers, But they couldn't achieve significant accuracy.

Pulido, Castillo and Melin [4] proposed Genetic Optimization of ensemble neural networks for complex time series prediction of the Mexican exchange. They implemented a Genetic Algorithm to optimize the ensemble 5 neural network architectures for each of the modules, and thus to find a neural network architecture that yields optimum results in each of the Time Series that dataset.

Using support vector machine Yuan [5] proposed an integrated model of Principal component analysis and Support vector Machine for stock prediction for Korean stock exchange and Hang Seng index for next following day. The PCA-SVM approach seems to be working for different stock with accuracy greater than 65%.

Mahdi Pakdaman Naeini, Hamidreza Taremiyan [6] In this paper, two kinds of neural networks, a feed forward multilayer Perceptron (MLP) and an Elman recurrent network, are used to predict a company's stock value based on its stock share value history. Both the networks are trained using backpropagation algorithm. The obtained results shows that amount of error using MLP neural network is less comparatively and MLP predicts close to the real one in comparison to the other methods.

### III. PREDICTION METHODS

Various different models have been considered for the prediction:

#### 1) Traditional time series prediction :-

Traditional statistical models are widely used in industries for time series prediction. These models are good at modeling linear relationships between the factors that affect the market and the value of the market. In economics, there are two types of time series forecasting: univariate and multivariate.

Multivariate models are nothing but univariate models expanded to find factors that could affect the behavior of the data . As the name suggests, the equation of the model has multiple variables. Many studies have compared the multivariate model with NNs.

#### 2) Neural Networks :-

Neural Networks have the ability to predict market directions more accurately than any other prediction techniques. The feature of NNs to discover non-linear relationship between the training i/o pairs makes them ideal for modeling nonlinear dynamic systems such as stock markets.

One of the benefits is the ability to learn and adapt the relationship through the data itself rather than assuming the functional form of the relationship. As Neural networks are known as a universal approximator, any relationship can be modelled to any degree of accuracy when enough data for the modeling are given. In addition, it withstand a level of tolerance to noise and incomplete data representation. Another benefit is that Neural networks have non-linear, non-parametric and adaptive learning properties and they also have the most practical effect in the modeling and forecasting. The nonlinear nature of Neural networks shows huge potential to solve many difficult problems. Due to the above mentioned characteristics of stock markets, Neural networks can be used for stock market prediction. Stock data is difficult to model because of its complexity, thus, non-linear models are more useful. Secondly, huge set of interacting input sequence is sometimes needed to explain certain stock.

#### 3) Support Vector Machine :-

Support Vector Machine has become an attractive topic of research because of its successful utilization in classification and regression, specifically in time series prediction and financial applications .

SVM is a specific type of learning algorithm characterize by the control capacity of the decision function, the use of the kernel function and the sparsity of the solution. Based on the special theory of the structural risk minimization principle to calculate a function by minimizing an upper bound of the error, SVM is very resistant to the overtraining problem, achieves the high generalization performance. Another important property of SVM is that training SVM is similar to solving a linearly restricted quadratic programming problem. Therefore the solution of SVM is unique and globally optimal, unlike Neural network training which needs nonlinear optimization with a possibility of getting stuck at local minima.

### IV. PROPOSED MODEL

We proposed a model to predict the price movement of a stock. Initially we examine different machine learning techniques. Then we will be examining various generative and non-generative models to get best prediction model. The proposed model is shown below:

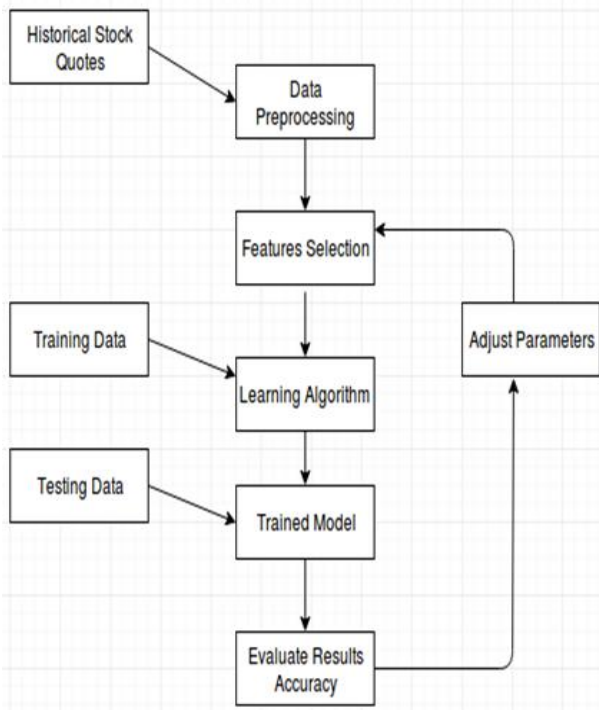


Fig.1 Proposed Model

In the process of predicting the future of stock market there needs to be a detailed procedure to follow as described in this figure, the initial step will be to collect some historical stock data of any firm and preprocess that data to make it acceptable to be used in the prediction algorithm, later some features are required to be selected and the historical training data once we have built the learning algorithm must be feed to the system, Now that the system is trained with the particular training model we will test the prediction aspect of this model with the help of testing data and carry out the prediction, once we acquire the prediction, we need to evaluate the result to find out its accuracy compared to the actual value.

Once the discrepancies are known we just need to adjust the parameters and repeat the steps to find better accuracy for further iterations with the chosen prediction model.

## ARTIFICIAL NEURAL NETWORK

An Artificial Neural Network is nothing but interconnected links with some weight value attached to it. The biological neural network have been the basic idea of the Artificial neural networks. Any Artificial neural network can be considered as large amount of interconnected units which are divided into three different layers they are: input layer, hidden layers and the output layer.

Inputs are fed to the input layer, and the weights associated with them are passed to the hidden layer. Having additional levels of hidden neurons gives more flexibility and accurate processing.

Artificial neural network can adapt to data patterns and relationship between input and the output, which results in better accuracy compared to traditional method.

An Artificial neural network architecture with two hidden layers is as follows:

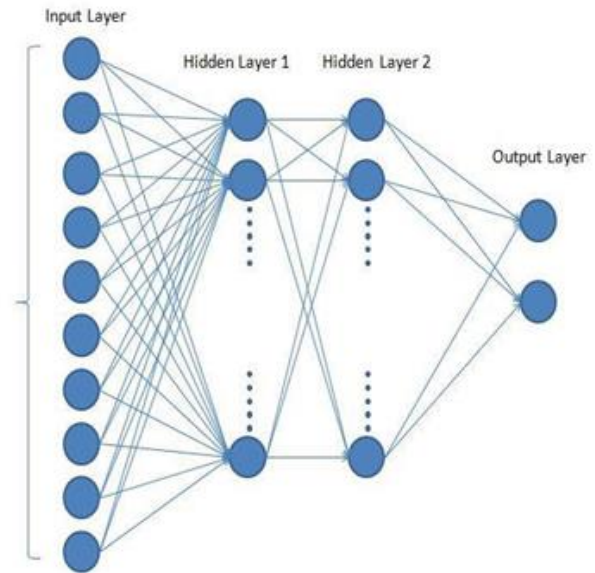


Fig.2 Architecture of ANN with two hidden layers

## V. CONCLUSION AND FUTURE SCOPE

In this paper, we analyzed recent progress in stock market prediction models. By comparing different prediction models, we found out that Neural Networks outperforms all other prediction techniques in terms of prediction accuracy. The capability of Neural Networks to map non-linear relationships from the training input/output pairs empower them to model non-linear dynamic systems such as stock markets more meticulously. It has been reviewed that NNs based on both perceptible and subjective factors are far superior to the ones based only on the quantitative factors.

The network will be trained to accept past predefined data sets and statistics upon which it will learn and adapt to the pattern and predict the future stock market value, which will be approximate and close to the original value and a comparative study will be presented to relate different models and their prediction for analysis.

## VI. REFERENCES

- [1] Kim, K. and Han, I. 2000, 'Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index', Expert System Appliaance, vol. 19.
- [2] J. Brofos, "Ensemble Committees for Stock Return," 4 April 2014. [Online]. Available: <http://arxiv.org/abs/1404.1492>.
- [3] P. Salamon and Hansen L. K , "Neural network ensembles," IEEE Trans.Pattern analysis and machine intelligence, vol. 12, no. 10, pp. 993-1001, 1990.
- [4] P. Melin, O. Castillo and M. Pulido, "GENETIC OPTIMIZATION OF ENSEMBLE NEURAL NETWORKS," International Journal of Innovative, vol. 9, no. 10, pp. 4151-4166, 2013.
- [5] In-Chan Choi and Yanshan Wang,, "Market Index and Stock Price Direction Prediction using Machine Learning Techniques," Elsevier Ltd, Korea, 2013.