Forward Forecast of Stock Price Using Sliding-window Metaheuristic-optimized Machine Learning

Radha Angad Sawant, Prof. Dr. B. M. Patil

Students and Prof. of Department of Computer Engineering,
MBES College of Engineering, Ambajogai.
BATU University, Lonere
Maharashtra, India

Abstract: — Stock market trading is an activity in which investors need fast and accurate information to make effective decisions. Since many stocks are traded on a stock exchange, numerous factors influence the decision-making process. Moreover, the behaviour of stock prices is uncertain and hard to predict. For these reasons, stock price prediction is an important process and a challenging one. This leads to the research of finding the most effective prediction model that generates the most accurate prediction with the lowest error percentage. This paper reviews studies on machine learning techniques and algorithm employed to improve the accuracy of stock price prediction. The neural network, one of the intelligent data mining technique that has been used by researchers in various areas for the past 10 years. Prediction and analysis of stock market data have got an important role in today’s economy. The various algorithms used for forecasting can be categorized into linear (AR, MA, ARIMA, ARMA) and non-linear models (ARCH, GARCH, Neural Network). In this paper, we are using four types of deep learning architectures i.e. Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) for predicting the stock price of a company based on the historical prices available. Here we are using day-wise closing price of two different stock markets, National Stock Exchange (NSE) of India and New York Stock Exchange (NYSE). The network was trained with the stock price of a single company from NSE and predicted for five different companies from both NSE and NYSE. It has been observed that CNN is outperforming the other models. The network was able to predict for NYSE even though it was trained with NSE data. This was possible because both the stock markets share some common inner dynamics. The results obtained were compared with ARIMA model and it has been observed that the neural networks are outperforming the existing linear model (ARIMA).

Keywords – Machine Learning, Deep Learning, Finance, Stock Price Prediction, Time Series Analysis, Recurrent Neural Networks (RNN), Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), Artificial Neural Network, Mean Absolute Percentage Error, National Stock Exchange, New York Stock Exchange, etc.
I. INTRODUCTION

The stock market provides opportunities for brokers and companies to make investments on neutral ground [2]. Stock prices are predicted to determine the future value of companies’ stock or other financial instruments that are marketed on financial exchanges [3]. However, the stock market is characterized by nonlinearities, discontinuities, and high-frequency multi-polynomial components because it interacts with many factors such as political events, general economic conditions, and traders’ expectations [4]. Therefore, making precise predictions of stock values are challenging.

Machine learning (ML) is coming into its own that can play a key in a wide range of critical applications. In machine learning, support vector machines (SVMs) have many advanced features that are reflected in their good generalization capacity and fast computation [6]. They are also not very sensitive to assumptions about error terms and they can tolerate noise and chaotic components. Notably, SVMs are increasingly used in materials science [8], the design of engineering systems [9] and financial risk prediction [10].

Funding or investment is the technology and art of growing money with the aid of placing money to work [3]. Stock marketplace has been a center of attraction for the investors as well as funders for an extended period of time. It historically provided the very best returns of any financial asset which was near 10% over the long time. In stock market, its miles possible to make multiple returns as well as to lose the precept and pass bankrupt. The major key to fulfillment is to buy and sale the stock at the proper time for the right cost.

Forecasting of stock marketplace has been a vital subject matter in one-of-a-kind fields of computational sciences because of its viable monetary profit. Stock marketplace is a place wherein excessive capital is invested and organizations do buying and selling of their shares [1]. Inventory marketplace forecasting poses the task of disproving the green market hypothesis, which states that the marketplace is efficient and can’t be expected. Researchers have labored hard to prove the truth that monetary markets are predictable. With the advancement and availability of era, inventory markets at the moment are greater reachable to investors. Various fashions were proposed, each in industry and academia, for stock market prediction ranging from machine learning knowledge of, to data mining, to statistical models.

In this study, time collection forecasting models are constructed to forecast the market analysis index by using the usage of a back propagation classification algorithm (a wildly used technique) [1]. Models are built in several workflows a good way to look at a suitable span of time for the inventory market information.

II. LITERATURE SURVEY

- Sahaj Singh Maini and Govinda.K(ICISS 2017) proposed an approach towards prediction of stock market trends using machine learning models like Random Forest model and Support Vector Machine. The context driven from news articles and the dataset from the timeline 2000 to 2016 was used for prediction of Dow Jones Industrial Average Index. The study predicts that Random Forest model using a 1-gram model [1] for text analysis produced an accuracy of 84.3 percentage and on using a 2-gram model produced an accuracy of 86.2 percentage. The linear Support Vector Machine using 1-gram model and 2-gram model for text analysis produced predictions with an accuracy of 82.2% and 84.6%, while the nonlinear Support Vector Machine produced predictions with an accuracy of 85.1 percentage for both 1-gram and 2-gram models. [1] It was observed that the Random Forest Model outperforms the Support Vector Machine while using the given dataset.
Pankaj Kumar and Dr. Anju Bala proposed a model for stock market using machine learning algorithms such as Decision tree, Linear model, Random forest and further their results have been compared using the classification evaluation parameters such as H, AUC, ROC, TPR, FPR, etc. The study [2] predicts that Random forest is the most effective model to predict the binary classification as it yields the highest accuracy of 54.12 percent whereas decision tree and linear model gives the accuracy of 51.87% and 52.83% respectively. It also predicts that study of problem solving of binary classification data [2] based on the machine learning models gives the best 2out of models which is experimentally used.

Ching-Hsue Cheng, You-Shyang Chen (ICMLC 2007) proposed a model to predict RGR of firms which employs Decision Tree C4.5, [3] Bayes net, Multilayer perception and Rough sets techniques. They used the RGR dataset in Taiwan stock market. The process was based on using revenues, assets, profits, income, and other data as condition attributes to determine the potential future growth of its revenue. The imperative results indicate that the rough sets outperform the listing methods because of its accuracy and understandable tools.

Lamartine Almeida Teixeira and Adriano Lorena Incio de Oliveira (IEEE 2009) proposed a technique that consists of a combination of a nearest neighbor classifier and some well-known tools of technical analysis, namely, stop loss, stop gain and RSI filter. The results obtained [4] are compared to the results that would be obtained by adopting a buy-and-hold strategy. The key performance measure in this comparison was profitability. The proposed method was shown to generate considerable higher profits than buy- and-hold for most of the stock data.

Radu Iacomin (ICSTCC 2016) proposed a new algorithm on predicting the stock markets. PCASVM was implemented to both eliminate the false predictions and to determine what features are important. Comparing to the simple methods [5] from SVM and evolving to GASVM and PCASVM, the solution to the main problem and sub-issues was more efficient and showed promising results for a real prediction using recent data sets.

Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., Deng, X. proposed a technique for stock market prediction on the basis of sentiments of Twitter feeds [6] which was experimented on SP 100 index. A continuous Dirichlet Process Mixture model was used to learn the daily topic set. Stock index and Twitter sentiment time-series were then regressed to make a prediction.

Sneha Soni (2010) proposed combination of three supervised machine learning algorithms, classification and regression tree (CART), linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) for classification of stock market data. [7] In section IV and V of IJCSE Vol. 02, No. 09, 2010, experimental results and performance comparison section show that classification and regression tree misclassification rate is only 56.11 percent whereas LDA and QDA show 74.26% and 76.57% respectively.

The work proposed by Shailendra Shrivastava and Sneha Soni (2010) in this paper is unique in comparison to other works in literature as they used a combination of supervised machine learning algorithms for classification of Indian stock market data while other works consist of unsupervised machine learning algorithms. [7] From this paper it was concluded that classification and regression tree, supervised machine learning algorithm is best as compared to linear and quadratic discriminant analysis.
Hiral R. Patel and Satyen M. Parikh proposed the development of the prediction model on crude commodity based on its price movement due to news released by various sources. The decision strategy would be driven by analysing stock price fluctuations. [8] The paper decides the prediction method to be used in model by performance comparison of following prediction techniques: Regression Modelling techniques, Classification Techniques, Statistical Techniques.

Ching-Te Wang and Yung-Yu Lin in 2015 12th International Conference on FSKD proposed a Web robot to capture data from the stock market. The system explores and analyzes the information to predict stock prices in the see-saw process. Using a group of cement, medical industries as the examples, [9] this paper discusses the topics of Web robot, Genetic Algorithm and Support Vector Machine, which can provide a framework for data analysis and predict the stock market. Support Vector Machine (SVM) is effective classification in the supervision learning. SVM can map data from high dimension into low-dimension space. When the data are classified in the procedure, SVM can collect non-overlapping data and then distinguish each other to a classification situation. The main characteristic of this system is to collect data automatically by using web robots and establish regular expressions, Xpath to analyzing web pages. [9] From the experimental results, the system has shown a better performance. Consequently, the method can crawl the valuable data, analyze huge information efficiently and provide the function to predict the prices in the stock market.

Girija V Attigeri, Manohara Pai M M, Radhika M Pai, Aparna Nayak in 2015 in their paper considered both technical and fundamental. [10] Technical analysis is done using historical data of stock prices by applying machine learning and fundamental analysis is done using social media data by applying sentiment analysis. Exploiting social media data in addition to numeric data increases the quality of the input and gives improved predictions. The aide of big data technology allows predictions at real-time. [10] The algorithm used for sentiment analysis uses summative assessment of the sentiments in a particular news article or tweet, this could be improved for better.

III. PROPOSED MODEL

In this project, we have demonstrated a machine learning approach (deep learning) to predict stock market trend using different neural networks. Results show how history data has been used to predict stock movement with reasonable accuracy. Also, with T test result analysis we can conclude that LSTM performs better in comparison to Back propagation and Linear Regression. For this implementation, we can conclude that if we incorporate all the factors that affect performance of the stock and feed them to neural network with proper data preprocessing and filtering, after training the network we will be able to have a model which can predict stock momentum more accurately and precisely for the better idea of stock value so that firms may have increased profit ratio as compared to what is might be going currently at that time. This will also lead to more transparency regarding stock as it will be easier for firms to analyze loses and achieve great success.
IV. OBJECTIVES

- To design approach for a generate pattern from large set of data of stock market for prediction for users.
- To predict an approximate value of share price.
- To assess whether prices over short period display any systematic patterns or they are indistinguishable.
- To know and understand using this project the current market situation for further investments.
- Accurately prediction and recommendations of market index prices.

V. LIMITATIONS

- Required proper project plan.
- System required i.e. PC, Laptop.
- Internet Connection required.
- Database or dataset required.

VI. CONCLUSION

This system presents computational issues of supervised machine learning algo i.e. classification and regression tree, linear and quadratic discriminant analysis set of rules on Indian stock market data for its class with committed intention for maximizing earnings of market analyst and investors to make selection for selling, buying or holding inventory (stock) of a specific employer on the idea of type rule. most of the algorithms, classification and regression tree set of rules is right because result show that class and regression tree algorithm type effects are less difficult to interpret and understand compared to linear and quadratic discriminant analysis algorithm as it gives consequences inside the form of tree structure. a good way to examine the classification performance of various gadget getting to know set of rules, classifiers are applied
on identical data and outcomes are as compared on the idea of misclassification and accurate type charge and according to experimental outcomes.

ACKNOWLEDGEMENT
I would prefer to give thanks the researchers likewise publishers for creating their resources available. I'm conjointly grateful to guide, reviewer for their valuable suggestions and also thank the college authorities for providing the required infrastructure and support.

REFERENCES
[2] Pankaj kumar and Dr. Anju Bala, “Intelligent Stock Data Prediction using Predictive Data Mining Techniques“ at CSED Department, Thapar University, Punjab, India