Published in AJSE, Vol:21, Issue: 3
Received on 10th August 2022
Revised on 13th November 2022
Accepted on 30th December 2022

AIUB JOURNAL OF SCIENCE AND ENGINEERING
ISSN: 1608 – 3679 (print) 2520 – 4890 (Online)

Stock Price Prediction: An Incremental Learning Approach Model of Multiple Linear Regression

Md. Tanvir Mahtab, A. G. M. Zaman, Montasir Rahman Mahin, Mia Mohammad Nazim, and Md. Tanjirul Islam

***Abstract*—The endeavour of predicting stock prices using different mathematical and technological methods and tools is not new. But the recent advancements and curiosity regarding big data and machine learning have added a new dimension to it. In this research study, we investigated the feasibility and performance of the multiple linear regression method in the prediction of stock prices. Here, multiple regression was used on the basis of the incremental machine learning setting. The study conducted an experiment to predict the closing price of stocks of six different organizations enlisted in the Dhaka Stock Exchange (DSE). Three years of historical stock market data (2017-2019) of these organizations have been used. Here, the Multiple Regression, Squared Loss Function, and Stochastic Gradient Descent (SGD) algorithms are used as a predictor, loss function, and optimizer respectively. The model incrementally learned from the data of several stock-related attributes and predicted the closing price of the next day. The performance of prediction was then analysed and assessed on the basis of the rolling Mean Absolute Error (MAE) metric. The rolling MAE scores found in the experiment are quite promising.**

***Index Terms*—Dhaka Stock Exchange (DSE), Incremental learning, Multiple linear regression, Prediction, Stochastic Gradient Descent (SGD), Stock market**

I. INTRODUCTION

Stock market is one of the most stochastic financial sectors of today's world. The economic condition of a country is largely impacted by the movement of its stock market. Although the Efficient Market Hypothesis (EMH) indicates the impossibility of consistently predicting the movements of the market, the practice of leveraging the power of machine learning techniques in stock trading is rising day by day [1].

Md. Tanvir Mahtab is with DataWorks Ltd., Dhaka-1213, Bangladesh (email: mahtab.tanvir.cs@gmail.com)
A. G. M. Zaman is with the department of Computer Science, American International University-Bangladesh (email: agmzaman@aiub.edu)
Montasir Rahman Mahin is with Brainstation 23 Limited, Dhaka-1212, Bangladesh (email: mrahmanmahin@gmail.com)
Mia Mohammad Nazim is with Star Soft Ltd., Dhaka-1229, Bangladesh (email: miamdnazim@gmail.com)
Md. Tanjirul Islam is with the department of Computer Science, American International University-Bangladesh (email: tanjirulislamprince@gmail.com)

Along with the fundamental and technical analysis, which are being used traditionally, several machine learning algorithms seem to be effective to some extent in some cases [2]. The rapid improvement of the capability of computing devices has simplified the implementation of these highly data-intensive algorithms. Besides, real-time predictive analytics is getting much more attention in recent times, especially in the highly dynamic economic and financial sectors. Since the stock market data can be considered as streaming data, stock trading can also be benefited by using real-time predictive analytics [3]. In terms of machine learning for streaming data, the approach of incremental learning is often more helpful than its batch learning counterpart [4, 5]. There are several machine learning algorithms that are suitable to use based on the setting of incremental learning. One such algorithm is Linear Regression which is very effective in the prediction of time-series data like those of stock markets [6].

While the batch learning approach of machine learning has its own set of advantages, the necessity of incremental learning is increasing day by day. In today's stock market of high-frequency trading, this necessity is significantly relevant. An automated stock prediction system, based on an incremental learning mechanism, could be of great use in terms of getting market insights and making decisions instantaneously.

This research study is conducted using a Multiple Linear Regression (MLR) model. The model is designed and implemented based on an incremental learning setting. One of the aspects of incremental learning for streaming data is the ability of learning and predicting instantaneously which demands a model that should be computationally less expensive. A simple but effective machine learning algorithm like linear regression is a suitable candidate in this regard [7]. The study used stock data of six organizations (Dhaka Bank Ltd., BRAC Bank Ltd., BEXIMCO Pharmaceuticals Ltd., The ACME Laboratories Ltd., Aramit Cement Ltd., and Confidence Cement Ltd.) enlisted in the Dhaka Stock Exchange (DSE) [8]. Among these there are two private banks, two pharmaceutical companies and two cement companies. Historical stock data of three consecutive years (2017, 2018, and 2019) of these organizations have been used.

The objective of this research was to come up with an incremental learning-based model that could be used to predict the closing price of stocks with an acceptable degree of accuracy. The model is expected to have the ability to be used in the production environment.

The nature of the stock market is highly non-stationary [9]. Therefore, there is less or no chance of a system to be developed that can impose certainty on this randomness. Notwithstanding, an Artificial Intelligence (AI)-powered automated system, capable of predicting the closing price of stocks within an acceptable range of deviations, can be considered valuable in terms of getting market insights and decision-making [10]. Such a system should have significant demands among stock traders and investors [11].

The result of this experiment shows a significantly satisfactory predictive performance of the model. Though, at some points in time, there are few abnormal predictions, mainly due to the abnormal changes in the actual prices. But the rapid adaptability with the latest trend and the overall consistent decrease of the MAE scores indicate the strength and potential reliability of the model. The outcome of this research has a potential for the development of an artificially intelligent automated stock price prediction system that could provide services to its possible clients and customers.

This research study is organized into multiple sections. Section I introduces the overall concepts of the research, section II contains a glimpse and insights of the resources which have been studied for this research, section III focuses on the methodology used to conduct the research, section IV discusses the implementation of the incremental learning model, section V contains the analysis of the experimental results, and finally, section VI concludes the overall research findings.

II. LITERATURE REVIEW

Curiosity and the attempt of predicting something is an inherent human nature. When it associates with risk and money, like predicting the movement of the stock market, it becomes particularly of special interest and fascination [12]. With the advancements of computing devices and algorithmic techniques, research on predicting the stock market has increased dramatically. The impact of various parameters on the movement of stock prices is being studied quantitatively in a more structured and rigorous way [13]. Different machine learning algorithms are being trained on these parameters, including some parameters regarding public sentiment, and their performances of predictions are being evaluated from different points of view.

With the unprecedented increase in the amount of data, the technologies of managing and utilizing those are being improved rapidly [14]. Business organizations are realizing the necessity of using the insights of these tremendous amounts of data to make more effective data-driven decisions [15]. Many of these data come in the form of streaming data from different sources. Often it seems helpful to get insights and predict trends based on these data instantaneously. The necessity of incremental learning is realized in this context. Incremental learning has several benefits over batch learning. Unlike batch learning, incremental learning doesn't forget previous knowledge which makes this approach to learning even more natural [16]. Another advantage is- in the case of incremental learning, data can be loaded into the memory of the computer successively instead of all at once [17]. Therefore, if the dataset is too large to fit into the memory, incremental learning can be an option of choice [18]. Stock market data are dynamic and have a streaming nature. Very often decisions are required to be made on the fly in a high-frequency trading environment. This is a suitable scenario where the techniques of incremental learning can be implemented.

Predicting stock prices is a regression problem that falls under the supervised learning category. Many algorithms have been devised and tested for solving regression problems. Algorithms like Linear Regression and Support Vector Machine (SVM) are linear in nature [19]. On the other hand, ensemble methods like Random Forest, Bagging, Boosting, etc. are composed of multiple models [20]. The implementation of neural networks has gained much popularity in recent years [21-23]. The power of neural networks comes with the cost of powerful computing devices and time. These computational expenses sometimes make neural networks less suitable for many applications.

S. Banerjee, N. Dabeeru and R. Lavanya tried to get a suitable regression model between simple and multiple linear regression to predict the stock prices in their paper [24]. Where the multiple linear regression model seems to do better. Besides, the usability of Principal Component Analysis (PCA) and Support Vector Machine (SVM) in cooperation with multiple linear regression was also demonstrated.

The research of Md. F. Hossain et al. is aimed to investigate the applicability of SVM in the prediction of stock prices [25]. The performance of SVM was compared to some other traditional methods like Single Exponential Smoothing (SES), Double Exponential Smoothing (DES), and Autoregressive Integrated Moving Average (ARIMA). Among these, the SVM-based model seems to do better.

In their study, N. Hasan and R. I. Rasel proposed to use an Artificial Neural Network (ANN), which is basically a three-layer perceptron model (a feed-forward neural network), in association with a windowing operator [26]. Here, the sigmoid function was used as the activation function and the Sliding Window Validation was used as a validation technique. The evaluation metrics used were Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE).

The objective of the paper of Mustain B. Rubel et al. is to explore an efficient method for predicting the closing price of different stocks of the Dhaka Stock Exchange (DSE) [27]. They applied Artificial Neural

Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) on the historical stock data of several companies. The research found ANFIS as a significantly effective technique for predicting closing prices.

The paper of David M. Q. Nelson et al. investigates the applicability of the Long Short-Term Memory (LSTM) network, which is a variant of the Recurrent Neural Network (RNN), to anticipate future trends in the prices of stocks using historical data [28]. They also considered some indicators of technical analysis. The model was assessed by several metrics. The performance was compared with that of some other machine learning algorithms to test its capability. The results obtained were quite satisfactory. It has 55.9% accuracy on average, in terms of predicting the movement of the prices of stocks in the imminent future.

G. Montana and F. Parrella proposed an algorithm, based on Support Vector Regression (SVR), to mitigate the drawbacks of batch learning and incorporate incremental learning ability for streaming data [29]. They tried to solve the problem from the perspective of algorithmic trading.

There are several techniques for solving regression problems, and new techniques are being developed and improved regularly. While traditional techniques like simple and multiple linear regression, Autoregressive Integrated Moving Average (ARIMA), Support Vector Machine (SVM), etc. are common in use, more sophisticated techniques like Random Forest, Bagging, Boosting, Neural Networks, etc. are gaining attention as well [30, 31]. Computational expenses can be a reason for some sophisticated techniques to be less suitable for some applications. In those cases, traditional techniques might be considered more suitable. Another aspect is, some algorithms are inherently capable of functioning based on an incremental learning setting, whereas some need to be modified or redesigned to accomplish the goal.

III. METHODOLOGY

The incremental learning process model is shown below in Fig. 1. Three years of historical stock price data from the year 2017 to 2019 of six enlisted companies of the Dhaka Stock Exchange (DSE) were used to predict the closing price. These organizations' data were used to train the model and do the prediction at the same time.

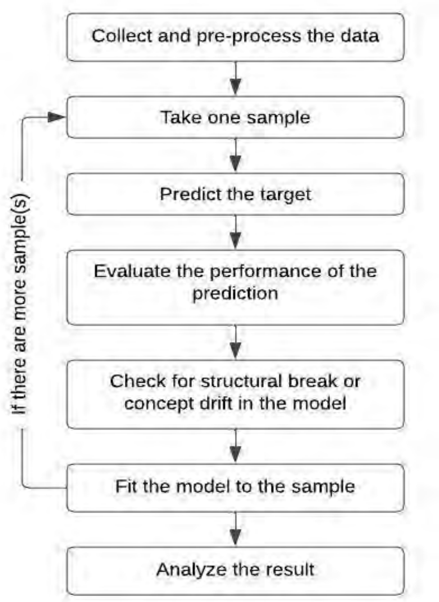

Fig. 1. Incremental learning process model

In the case of time series data, there is no way to split it randomly since the sequence of time will break. Our algorithm first takes the first sample of the entire dataset and does the prediction, performance evaluation, structural break checking, and model fitting as shown in Fig 1. Then it proceeds to the second sample and performs the same process. The whole process continues sequentially till the last sample of the dataset. Here, prediction can be considered as testing, and model fitting can be considered as training; both of these are done for each sample of the dataset in one iteration dedicated to each one. Therefore, there is no need to split the data into train and test sets. It is also noticeable that the prediction or testing is done prior to the fitting or training phase as opposed to the typical machine learning approach. This is done due to the reason that we want to know the performance of the model first by testing it for a sample and then train or adjust the structural break of the model for the same sample. This happens for each successive sample of the dataset and thus the model learns incrementally. The performance of prediction was evaluated using the Mean Absolute Error (MAE) metric.

IV. INCREMENTAL LEARNING MODEL IMPLEMENTATION

The implementation of the incremental learning model can be broken down into different stages such as data acquisition, data processing, feature selection, and implementation of the incremental learning model itself.

In the data acquisition step, three years of historical stock price data, from 2017 to 2019, of the selected stocks of the Dhaka Stock Exchange (DSE) were collected from Kaggle [32].

In the data processing step, the data were transformed into the intended format required in the following steps. The samples were sorted by ascending order of date. Then the date was converted to the ordinal date format to make it suitable for the incremental learning model.

Features were selected based on the Pearson Correlation Coefficient (PCC). Nine features were selected for training the model to predict the target variable, i.e., the closing price of the next day. These features were date, last traded price, highest price, lowest price, opening price, yesterday's closing price, total trade, value (in million), and volume. The Pearson correlation coefficient for Dhaka Bank Ltd. of DSE is shown below in Table I.

TABLE I
PEARSON CORRELATION COEFFICIENT FOR FEATURES AND TARGET

Feature Variable	PCC	Target Variable
Last traded price	0.947	Closing Price
Highest price	0.956	
Lowest price	0.954	
Opening price	0.953	
Yesterday's closing price	0.995	
Total trade	0.423	
Value (in million)	0.476	
Volume	0.411	

In multiple regression, the target variable is predicted based on multiple feature variables [33]. To accomplish this goal a hypothesis should be constructed as a function of the feature variables. This is a linear function that can be expressed generally as shown in (1).

$$\begin{aligned} y &= h_\theta(x_1, x_2, x_3, \dots, x_n) \\ &= \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \cdots + \theta_n x_n \end{aligned} \tag{1}$$

Here, $x_1, x_2, x_3, \dots, x_n$ are feature variables and y is the target variable. $\theta_0, \theta_1, \theta_2, \dots, \theta_n$ are parameters.

The hypothesis function in (1) aims to map the feature variables $(x_1, x_2, x_3, \dots, x_n)$ with the target variable (y) based on different values of the parameters $(\theta_0, \theta_1, \theta_2, \dots, \theta_n)$. This mapping provides an estimated value of y. Each combination of the values of parameters provides an estimated value of y for each value of x. Here the difference between the actual and estimated values of y will indicate whether the set of parameter values is optimal or not. The optimal set of the values of parameters would be that one for which this difference is minimum. To check the optimality, based on difference, a cost function is used [34], as shown in (2).

$$\begin{aligned} &J(\theta_0, \theta_1, \theta_2, \dots, \theta_n) \\ &= \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \dots, x_n^{(i)}) - y^{(i)})^2 \end{aligned} \tag{2}$$

Here, m is the number of total training samples.

The cost function in (2) can be plotted in a multi-dimensional space where the independent variables are $\theta_0, \theta_1, \theta_2, \dots, \theta_n$, and the dependent variable is the cost, J. Our goal is to get the minimum value of the cost function. The set of values of the parameters associated with this minimum cost is considered optimal. For this set of parameters, the hypothesis provides the best-fitted straight line which maps the features with the target in the best possible way.

In order to get the minimum value of the cost function another function is used which is called the optimization function or optimizer [35]. There are several functions that can be used as an optimizer. A frequently used optimizer is Gradient Descent [36], as shown in (3). The process in (3) starts with some arbitrary values of $\theta_0, \theta_1, \theta_2, \dots, \theta_n$ and keeps changing the values to reduce $J(\theta_0, \theta_1, \theta_2, \dots, \theta_n)$ until it hopefully ends up at a minimum.

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1, \theta_2, \dots, \theta_n) \quad (\text{for } j = 0,1,2, \dots, n) \tag{3}$$

Here, α is the learning rate that determines how fast the intended minimum of the cost function would be approached. Special care should be taken in determining the value of α, because a too small value of α can make the gradient descent process significantly slow. On the other hand, a too large value of α can cause the gradient descent to overshoot the minimum. Therefore, it may fail to converge or even may diverge.

In an ideal case, the gradient descent process is expected to converge to the minimum value of the cost function for which the associated set of parameters $(\theta_0, \theta_1, \theta_2, \dots, \theta_n)$ would be considered optimal. The hypothesis function will then provide the best fit line based on this set of parameters.

Incremental learning deals with the concept drift of the relevant parameters of a system [37]. In contrast to batch learning, the focus of incremental learning is mainly on the adjustment of this concept drift. Therefore, the incremental version of the multiple regression differs from its batch counterpart basically in the formulation of the cost function and optimizer. While batch learning takes all the training samples at once, incremental learning takes only one sample at a time. Hence, the formulation of the cost function and optimizer is done based on only one sample in an incremental mode, instead of all the training samples. As a result, as indicated below in (4), the incremental cost calculation does not consider the overall training sample size.

$$\begin{aligned} &J(\theta_0, \theta_1, \theta_2, \dots, \theta_n) \\ &= \frac{1}{2} (h_\theta(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}, \dots, x_n^{(i)}) - y^{(i)})^2 \end{aligned} \tag{4}$$

The incremental version of the gradient descent method resembles that of the batch version as shown in (3). But unlike the batch version, it optimizes the parameters for only one sample at a time instead of all training samples.

In our research study, a multiple regression model was implemented as a hypothesis or predictor function. The incremental version of the Squared Loss function and Stochastic Gradient Descent (SGD) algorithm were used as the loss function and optimizer respectively. These techniques were applied in accordance with the incremental learning approach depicted in Fig. 1.

The closing price of the next day of a particular stock was predicted and the performance of the prediction was evaluated in terms of rolling Mean Absolute Error (MAE). The rolling window size was twelve, i.e., the average score of MAE of the previous twelve days was considered to update the metric for each sample.

In this research study, the experiment was conducted using Python programming language. To perform Exploratory Data Analysis (EDA) and machine learning related tasks several Python libraries were used. These libraries are Pandas [38], Numpy[39], Matplotlib [40], and River [41].

V. EXPERIMENTAL RESULTS

Experiments have been done multiple times to get an optimal set of hyper-parameters settings. Besides, special care was taken so that no known bias could be introduced. The closing prices, both actual and predicted, with respect to date, are shown in Fig. 2.

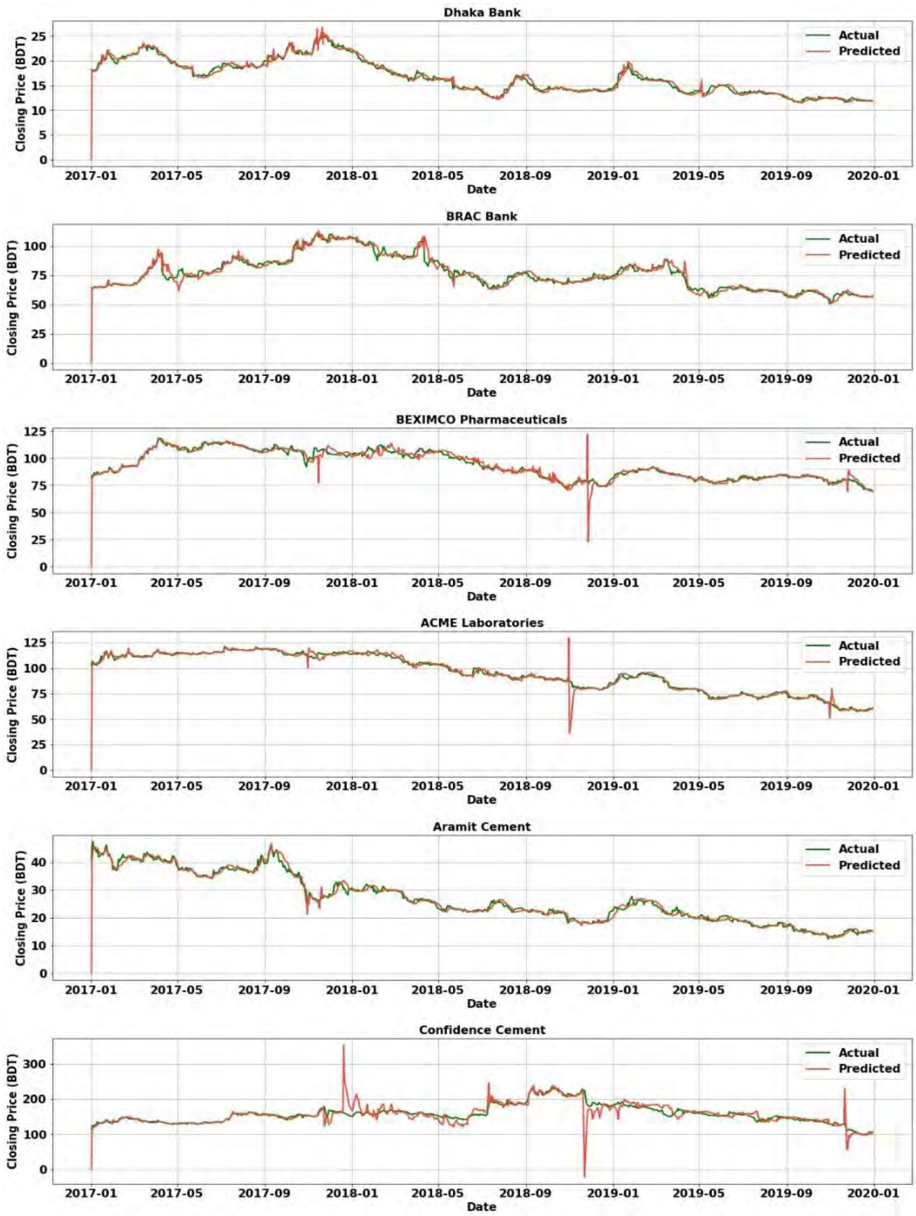

Fig. 2. Actual and predicted closing prices of different companies

From Fig. 2 we can observe that the trend of changes in the actual and predicted closing prices is similar for both banks, i.e., Dhaka Bank, and BRAC Bank. We also observed that the model started performing better from slightly before the year 2018 till the end. Similarly, both pharmaceutical companies, i.e., BEXIMCO and ACME also show a similar trend in terms of the change in the actual and predicted closing prices. Though there are a few abnormal predictions, the overall predictive performance is good.

Unlike the banks and pharmaceuticals, there is no similarity in the change in actual and predicted closing prices between the two cement companies, i.e., Aramit Cement, and Confidence Cement as shown in Fig 2. Even though there are several abnormal predictions, the overall performance is pretty good. The MAE of predictions with respect to date is shown in Fig. 3.

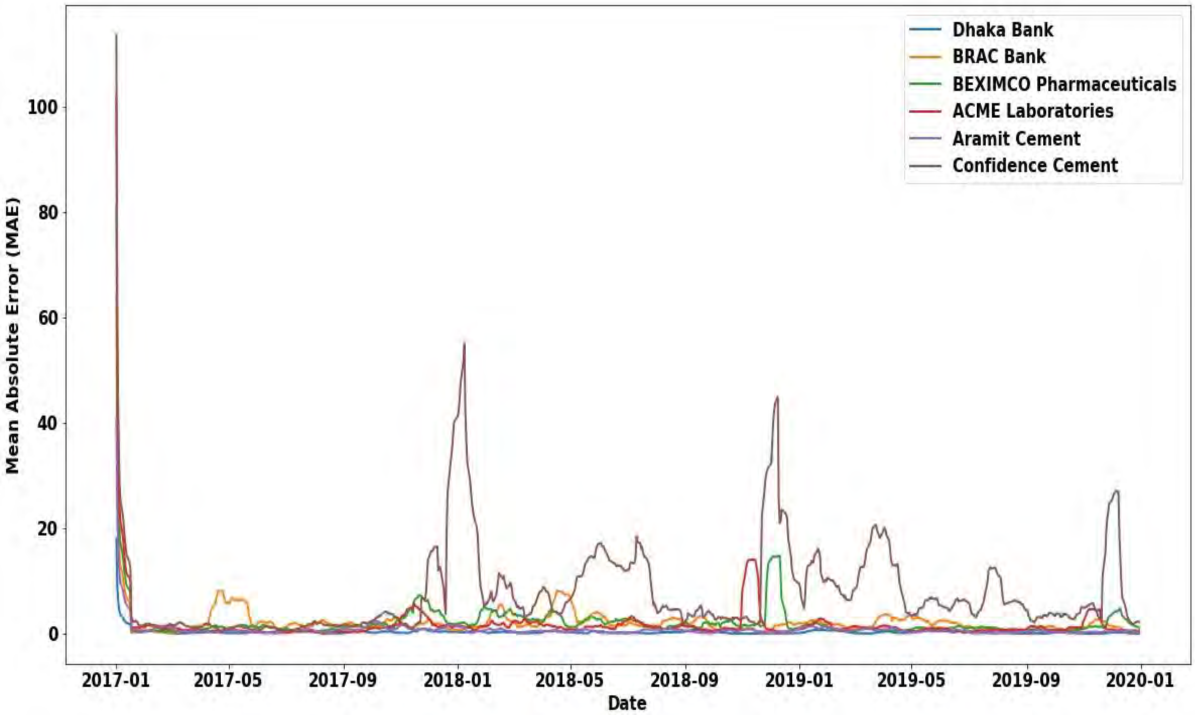

Fig. 3. Trend of the changes of MAE of different companies

The change in MAE of predictions with respect to date reflects the gradual improvement of the model, shown in Fig. 3. Though the MAE suddenly increases several times for BEXIMCO Pharmaceuticals, ACME Laboratories, and Confidence Cement, the overall tendency of MAE is toward the decreasing mode for all the companies.

Table II shows several successive MAE scores with 90 days of time intervals for each of the companies.

TABLE II
SUCCESSIVE MAE SCORES WITH 90 DAYS INTERVAL

Companies	**MAE Score**								
	After 1 day	After 90 days	After 180 days	After 270 days	After 360 days	After 450 days	After 540 days	After 630 days	After 720 days
Dhaka Bank	18.200	0.211	0.573	0.400	0.251	0.145	0.159	0.227	0.221
BRAC Bank	64.200	6.601	1.270	1.443	1.710	0.584	1.343	0.859	0.536
BEXIMCO Pharmaceuticals	81.200	1.505	1.359	5.214	2.319	1.536	1.062	1.068	2.512
ACME Laboratories	102.000	0.779	0.561	1.600	1.083	1.009	1.249	0.765	0.586
Aramit Cement	40.900	0.606	1.082	0.671	0.255	0.680	1.148	0.275	0.685
Confidence Cement	113.600	1.529	2.183	5.265	14.015	2.670	14.951	11.304	3.202

The experimental results as shown in Fig. 2, Fig. 3, and Table II reflect the fluctuating nature of the stock market, which is quite natural. At some points in time, the actual prices for some organizations changed abnormally, possibly due to many other secondary factors that affect stock prices. The abnormal spikes in the graphs of predictions reflect this fact. But the model seems to update itself and comply with the latest trend very rapidly. This is the strength of the model. Apart from this, the changes in MAE scores with respect to date, as shown in Fig. 3, are also satisfactory. Therefore, the performance of the model seems quite promising from an overall perspective.

VI. CONCLUSION

Predicting the movement of the stock market is a very complicated job as there are many hidden parameters associated with it that affect the trend of the market. The impact of intangible parameters which are related to national and international politics, and socio-economic aspects are very hard to study in a quantitative manner. Nevertheless, the attempt of predicting the market can provide some valuable insights to make effective data-driven decisions. This research study investigated the applicability of the multiple linear regression model for predicting stock prices in an incremental learning setting. Based on the study, the gradual improvement and performance of the model seem quite promising.

REFERENCES

[1] Bruno M. Henrique, Vinicius A. Sobreiro, H. Kimura, "Stock price prediction using support vector regression on daily and up to the minute prices," *The Journal of Finance and Data Science 4(3)*, April 2018, DOI: 10.1016/j.jfds.2018.04.003.

[2] Z. Iqbal *et al.*, "Efficient machine learning techniques for stock market prediction," *International Journal of Engineering Research and Applications*, ISSN: 2248-9622, Vol. 3, Issue 6, Nov-Dec 2013, pp.855-867.

[3] S. Nann, J. Krauss, D. Schoder, "Predictive Analytics on Public Data – The Case of Stock Markets," in *Proceedings of the 21st European Conference on Information Systems*, June 2013.

[4] H. He, S. Chen, K. Li and X. Xu, "Incremental Learning From Stream Data," in *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 1901-1914, Dec. 2011, doi: 10.1109/TNN.2011.2171713.

[5] H. Murilo Gomes *et al.*, "Machine learning for streaming data: state of the art, challenges, and opportunities," *ACM SIGKDD Explorations Newsletter 21(2):6-22*, November 2019, DOI: 10.1145/3373464.3373470.

[6] R. Seethalakshmi, "Analysis of stock market predictor variables using Linear Regression," *International Journal of Pure and Applied Mathematics,* vol. 119, pp. 369-377, Jan. 2018.

[7] A. Qomaruddin Munir, R. Wardoyo, "Comparison Analysis of Time Series Data Algorithm Complexity for Forecasting of Dengue Fever Occurrences," *International Journal of Advanced Computer Research 6(7)*, October 2015.

[8] "*Dhaka Stock Exchange Limited*," Accessed on: Feb. 17, 2021. [Online]. Available: https://www.dsebd.org/

[9] V. Murthy, K. Washer, John R Wingender, "Are Stock Prices in the US Nonstationary? Evidence from Contemporary Unit Root Tests," *Applied Financial Economics 21(22):1703-1709*, November 2011, DOI: 10.1080/09603107.2011.591731.

[10] Luke K. Rose, T. Gartner, "Automated Stock Market Trading Using Machine Learning," April 2018, DOI: 10.13140/RG.2.2.19887.59045.

[11] "*How AI Trading Technology is Making Stock Market Investors Smarter*," March 16, 2019. Accessed on: April 17, 2021. [Online]. Available: https://builtin.com/artificial-intelligence/ai-trading-stock-market-tech

[12] M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," *2016 3rd International Conference on Computer and Information Sciences (ICCOINS)*, 2016, pp. 322-327, doi: 10.1109/ICCOINS.2016.7783235.

[13] M. Faruquee, S. Haque, "Impact of fundamental factors on stock price: a case based approach on pharmaceutical companies listed with Dhaka stock exchange," *International Journal of Business and Management Invention*, ISSN (Online): 2319 – 8028, ISSN (Print): 2319 – 801X, vol. 2, Issue 9, September 2013, PP. 34-41.

[14] B. Purcell, "The Emergence of "Big Data" Technology and Analytics," *Journal of Technology Research*, vol. 4, Jan. 2012.

[15] H. Koscielniak, A. Puto, "Big Data in Decision Making Processes of Enterprises," in *Procedia Computer Science*, ISSN 1877-0509, vol. 65, pp. 1052-1058, December 2015.

[16] "*Incremental Learning*," Accessed on: March 21, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Incremental_learning

[17] Y. Luo *et al.*, "An appraisal of incremental learning methods," *Entropy*, ISSN 1099-4300, vol. 22, number 11, article number 1190, October 2020.

[18] V. Losing, B. Hammer, H. Wersing, "Incremental On-line Learning: A Review and Comparison of State of the Art Algorithms," *Neurocomputing*, ISSN 0925-2312, vol. 275, pp. 1261-1274, Sept. 2017, DOI: 10.1016/j.neucom.2017.06.084.

[19] A. Pradhan, "Support vector machine-A survey," *International Journal of Emerging Technology and Advanced Engineering*, ISSN 2250-2459, Vol. 2, Issue 8, Aug. 2012.

[20] "*Ensemble Learning*," Accessed on: March 21, 2021. [Online]. Available: https://en.wikipedia.org/wiki/Ensemble_learning

[21] E. Grossi, M. Buscema, "Introduction to artificial neural networks," *European Journal of Gastroenterology & Hepatology*, vol. 19, Issue 12, pp. 1046-1054. Dec. 2007, DOI: 10.1097/MEG.0b013e3282f198a0.

[22] N. Malik, "Artificial Neural Networks and their Applications," in *National Conference on 'Unearthing Technological Developments & their Transfer for Serving Masses'*, GLA ITM, Mathura, India, 17-18 April 2005.

[23] Z. Zhang, "A gentle introduction to artificial neural networks," *Annals of Translational Medicine, ISSN 2305-5847, vol. 4 – No. 19*, Oct. 2016, DOI: 10.21037/atm.2016.06.20.

[24] S. Banerjee, N. Dabeeru and R. Lavanya, "Stock market prediction," *International Journal of Innovative Technology and Exploring Engineering*, ISSN 2278-3075, vol. 9, Issue 9, July 2020, DOI:10.35940/ijitee.I7642.079920.

[25] Md. Farhad Hossain, P. Chakraborty and S. Islam, "Predicting Daily Closing Prices of Selected Shares of Dhaka Stock Exchange (DSE) Using Support Vector Machines," *Internet of Things and Cloud Computing* , vol. 8, pp. 46-51, Dec. 2020.

[26] N. Hasan and R. I. Rasel, "Artificial Neural Network Approach for Stock Price and Trend Prediction," in *International conference on advanced information & communication technology*, May 2016.

[27] Mustain B. Rubel *et al.*, "Predicting Closing Stock Price using Artificial Neural Network and Adaptive Neuro Fuzzy Inference System (ANFIS): The Case of the Dhaka Stock Exchange," *International Journal of Computer Applications*, vol. 129 – No. 11, Nov. 2015, DOI: 10.5120/ijca2015906952.

[28] D. M. Q. Nelson, A. C. M. Pereira and R. A. de Oliveira, "Stock market's price movement prediction with LSTM neural networks," *2017 International Joint Conference on Neural Networks (IJCNN)*, 2017, pp. 1419-1426, doi: 10.1109/IJCNN.2017.7966019.

[29] G. Montana, F. Parrella, "Learning to Trade with Incremental Support Vector Regression Experts," in *Hybrid Artificial Intelligence Systems, Third International Workshop, HAIS 2008, Burgos, Spain*, September 24-26, 2008, vol. 5271, pp. 591-598, DOI: 10.1007/978-3-540-87656-4_73.

[30] B. Pavlyshenko, "Machine-Learning Models for Sales Time Series Forecasting," *Data*, vol. 4, no. 1, p. 15, Jan. 2019, doi:

10.3390/data4010015. [Online]. Available: http://dx.doi.org/10.3390/data4010015
[31] L. Sindayigaya, "Machine Learning Algorithms: A Review," *Information Systems Journal*, vol. ISJ-RA-3392, sp. 6, Aug. 2022.
[32] M. Haque "*Dhaka Stock Exchange (Historical prices of stocks)*," Version 3, Updated on: Jan. 24, 2021. Accessed on: Feb. 17, 2021. [Online]. Available: https://www.kaggle.com/mahmudulhaque/dsebd
[33] Andrew Ng, "*Machine Learning*," Accessed on: May 24, 2021. [Online]. Available: https://www.coursera.org/lecture/machine-learning/multiple-features-6Nj1q
[34] "*Data Science 101: Introduction to Cost Function*," Mar. 4, 2021. Accessed on: Apr. 17, 2021. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/03/data-science-101-introduction-to-cost-function/
[35] "*Introduction to optimizers*," May 7, 2018. Accessed on: Apr. 24, 2021. [Online]. Available: https://algorithmia.com/blog/introduction-to-optimizers
[36] "*CS 2750 Machine Learning Lecture 8-9*," Accessed on: May 19, 2021. [Online]. Available: https://people.cs.pitt.edu/~milos/courses/cs2750-Spring2020/Lectures/Class8.pdf
[37] "*A Gentle Introduction to Concept Drift in Machine Learning*," Dec. 10, 2020. Accessed on: Apr. 13, 2021. [Online]. Available: https://machinelearningmastery.com/gentle-introduction-concept-drift-machine-learning/
[38] "*Pandas Documentation*," Accessed on: Mar. 11, 2021. [Online]. Available: https://pandas.pydata.org/docs/
[39] "*NumPy v1.20.0*," Accessed on: Mar. 11, 2021. [Online]. Available: https://numpy.org/
[40] "*Matplotlib v3.4.2*," Accessed on: Mar. 13, 2021. [Online]. Available: https://matplotlib.org/
[41] J. Montiel *et al.*,"*River: machine learning for streaming data in Python*," year- 2020, eprint- 2012.04740, archivePrefix- arXiv, primaryClass- cs.LG

Md. Tanvir Mahtab received his B.Sc. degree in Computer Science and Engineering from American International University-Bangladesh (AIUB) in 2021. Currently, he is working as a data analyst. His fields of interest are data science and machine learning.

A. G. M. Zaman is an Assistant Professor of the Computer Science Department, Faculty of Science and Technology, American International University-Bangladesh (AIUB). His research interests include but not limited to Algorithms, Data Mining, Machine Learning, Optimization, Intelligent Traffic System, Autonomous Vehicle, and Robotics.

Montasir Rahman Mahin received his B.Sc. degree in Computer Science and Engineering from American International University-Bangladesh (AIUB) in 2021. Currently, he is working as an associate SQA engineer at Brainstation 23 Limited. His research interests are big data and data analysis.

Mia Mohammad Nazim received his B.Sc. degree in Computer Science and Engineering from American International University-Bangladesh (AIUB) in 2021. He is currently working as an associate software engineer at Star Soft Limited. His research interests are data science and data analysis.

Md. Tanjirul Islam received his B.Sc. degree in Computer Science and Engineering from American International University-Bangladesh (AIUB) in 2021. His research interests are big data and data analysis.