Prediction of Wall Street Through AI Approach

V. Rajesh, Udimudi Satish Varma, Ch. S. Sashidhar Reddy, V. Dhiraj, Udimudi Prasanna Swetha, and S. K. Hasane Ahammad

Abstract Financial exchange is one of the main areas of a nation's economy. These are hard to foresee due to their unpredictable framework elements, making it an intriguing industry to study. The present target is to discover an ideal regressor algorithm that forecasts future "Close Price" of stocks through a Comparative report between various machine learning, Deep learning, and Time-series forecasting techniques like ARIMA, Random Forest Regressor, Linear Regression, LSTM (which is said to be the particular type of RNN), SVM Regressor is performed including analysis. Then, a Comparative report between these techniques has been done regarding its prediction accuracy and performance. After analyzing all the models separately, the LSTM is the most precise for the stock value forecast.

Keywords Financial exchange · Stock prediction · Regressor model · Support · Vector machine · Random forest · Linear regression · ARIMA · LSTM

1 Introduction

Stock exchanges are places where individuals and institutions can trade stocks. These transactions are carried out in a public setting [1]. It is believed that the volatility of the markets can affect the nation's development [2]. The goal of this work is to predict the closing prices of the stock market records for the period 2012–2020. The predictions are made using the various regressor models that were used [3]. This work aims to analyze the performance of AI models when it comes to estimating the closing price of stocks.

V. Rajesh · S. K. Hasane Ahammad
Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

U. S. Varma · Ch. S. Sashidhar Reddy · V. Dhiraj (✉) · U. P. Swetha
Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India

D. Bhattacharyya et al. (eds.), *Machine Intelligence and Soft Computing*,
Advances in Intelligent Systems and Computing 1419,
https://doi.org/10.1007/978-981-16-8364-0_3

2 Literature Survey

Nonita Sharma [4] and her colleagues discussed the main idea of using past market data to predict the prices of financial exchange. The proposed model is more robust than SVM Regressor and can be utilized to design effective algorithms for financial market forecasts. Aakanksha Sharaff [3] and her colleagues [5] studied various stochastic models such as the Recurrent Neural Network, the Artificial Neural Network, and the ARIMA algorithm for stock market prediction. They concluded that the ANN has robust prediction capability.

Honghai Yu et al. [6] proposed a dual-stage ANN design that combines SVM and Empirical mode Decomposition algorithms. This model was suggested for stock price predictions. This model has been developed to predict the stocks accurately using the nonlinear stock value pattern prediction techniques. It is more advantageous to develop an integrated model that takes into account different features determination.

Srinath Ravikumar [7] and his coauthors [8] proposed a system that can be used for both classification and regression techniques. The former proposes that the system should be able to predict the price movements of stocks based on the previous day's data.

Ze Zhang et al. [9] developed a self-adaptive PSO (price prediction algorithm) based on the Elman [10] RNN framework to forecast the beginning of the financial market. The model is compared with the backpropagation network and their performance is evaluated. ARIMA and ANN models were used for the forecasting of the Indonesian stock market. The former is faster to forecast and has better learning ability [2, 11].

3 Methodology

The proposed methodology can be divided into various sections. The datareader, datapath, and scikit-learn are some of the essential Python libraries that are used to perform various tasks and the complete methodology of this research is explained in Fig. 1.

3.1 Yahoo Finance

Yahoo finance is a website that provides various types of financial news and information, including stock statements, reports, and official statements.

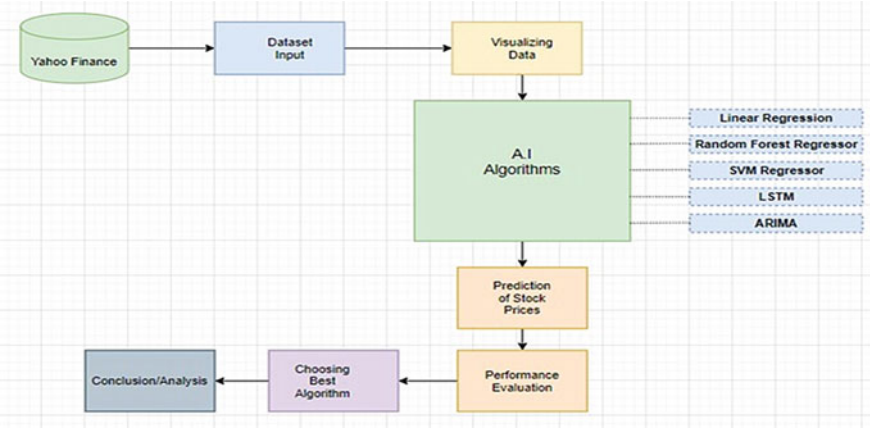

Fig.1 Proposed methodology

3.2 Dataset Input

The data are taken from yahoo finance and are used for training different machine learning models. The models were trained using the data to forecast the close price of the apple stocks.

3.3 Data Visualization

The stock prices of apple have been retrieved from the web resources since the 8th of November 2007. The data includes the date of the stocks, the high and low open, and the close prices.

3.4 Algorithms Used for Prediction

3.4.1 Random Forest

There is no legitimate method for pruning information in Decision Trees. Random Forest was introduced to avoid this issue. Formula for the Random forest can be given as follows. Figure 2 denotes the prediction of apple stocks.

$$s(x) = h_0(x) + h_1(x) + \ldots + h_n(x) \tag{1}$$

In (1)

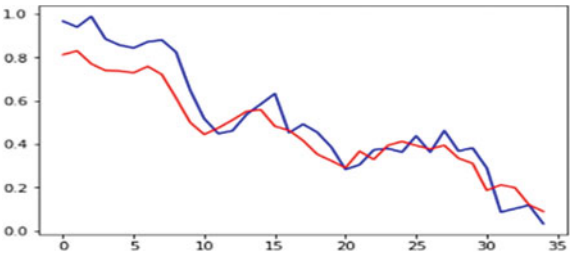

Fig. 2 Stock prediction by Random forest regressor

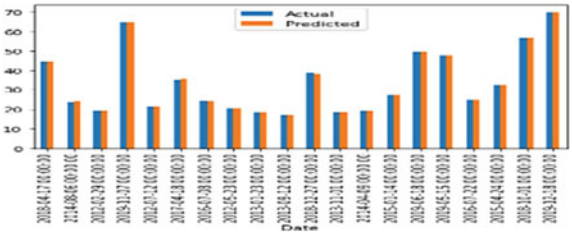

Fig. 3 Prediction of apple stock closing prices

$S(x)$ = Final Prediction.
$h_n(x)$ = nth decision tree's decision function.

3.4.2 Linear Regression

Linear regression is normally utilized in prediction and in stock analysis. Figure 3 denotes the prediction of apple stock closing prices with linear data. Linear regression generally draws a line that was close to all the data points considering the square of distance between all the data points is minimal with respect to the line drawn.

3.4.3 LSTM

LSTM is a special type of RNN utilized in deep learning in light of the fact that enormous architectures can be effectively prepared. Prediction accuracy is more in the LSTM model compared to other models which can be seen in Fig. 4 which denotes the prediction of apple stocks. LSTM is trained using the backpropagation

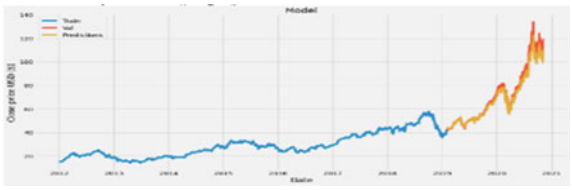

Fig. 4 Predicted versus actual closing prices of apple stock

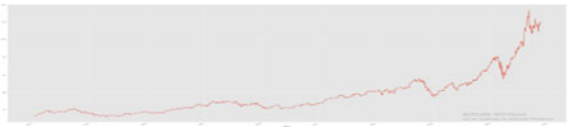

Fig. 5 Prediction of apple stocks by SVR

Fig. 6 Actual versus predicted price graph by ARIMA

through time which helps to overcome the problem of vanishing gradient (shrinking of the gradient as it back propagates) which is a drawback of RNN.

3.4.4 SVM Regressor

Both classification tasks and regression tasks can be implemented using the SVM model, keeping up all the principle includes that portray the model (maximal edge). Figure 5 denotes the prediction of apple stocks by SVR. SVM regressor utilizes almost the same principle as the SVM model but with a small set of changes to it.

3.4.5 ARIMA

ARIMA algorithm converts the non-static data into static data before it starts training from the data. ARIMA model has been utilized widely in the field of account and financial aspects as it is said to be robust, effective, and has a solid potential for forecasting the short-term stock market. Figure 6 denotes the prediction of apple stock prices.

3.5 Performance Evaluation

The performance of the models that are utilized in financial market price forecasting can be given by the accuracy measure, MSE measures normal of the squares of the errors (obtained value to the original value). MSE can be expressed as follows:

$$\text{Meansquarederror} = \frac{1}{\text{s}} \sum\nolimits_{k=1}^{s} \left(x\frac{1}{k} - x\frac{2}{k} \right)^2 \tag{2}$$

Table 1 Prediction analysis

S.no	Algorithm name	MSE (Apple stock)	MAPE(Apple stock)
1	Random Forest Regressor	6.5189	0.9843
2	Linear Regression	4.1724	0.8324
3	Long short term memory	3.2485	0.6523
4	SVM regressor	4.196	0.8564
5	ARIMA	4.161	0.8278

In (2)
S=total number of data points (number of days predicted)
x_k^1 = values observed.
x_k^2 = values predicted.
MAPE is an accuracy prediction measurement for a prediction model in statistics and it is given by:

$$\text{Mean Absolute Percentage Error} = \frac{1}{\text{s}}\sum\nolimits_{k=1}^{s}\left|\frac{O_k^1 - P_k^1}{O_k^1}\right| \tag{3}$$

In (3)
S = Total number of summation iterations.
O_k^1 = original value.
P_k^1 = Predicted value.
The performance measures of the algorithms that are used for predicting the stock price values are given in the following Table 1.

The models performed well in terms of both the stationary and non-stationary sets of data. The ARIMA is also accurate for short-term predictions.

4 Conclusion

AI algorithms are used for stock market price prediction. A comparative analysis is performed among the various algorithms. The ARIMA and LSTM models had the highest precision among the models.

References

1. Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Stock Price prediction using the arima model. In *2014 UKSim-AMSS 16th International Conference on Computer Modelling and*

Simulation, Cambridge (pp. 106–112). https://doi.org/10.1109/UKSim.2014.67.
2. Sharaff, A., & Choudhary, M. (2018). Comparative analysis of various stock prediction techniques. In *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli (pp. 735–738). https://doi.org/10.1109/ICOEI.2018.8553825.
3. Zhang, Z., Shen, Y., Zhang, G., Song, Y., & Zhu, Y. (2017). Short-term prediction for opening price of stock market based on self-adapting variant PSO-Elman neural network. In *2017 8th IEEE International Conference on Software Engineering andService Science (ICSESS)*, Beijing (pp. 225–228). https://doi.org/10.1109/ICSESS.2017.8342901.
4. Sharma, A., Bhuriya, D., & Singh, U. (2017). Survey of stock market prediction using machine learning approach. In 2017 International conference of (ICECA), Coimbatore, (pp. 506–509). https://doi.org/10.1109/ICECA.2017.8212715.
5. Sharma, N., & Juneja, A. (2017). Combining of random forest estimates using LSboost for stock market index prediction. In *2017 2nd International Conference for Convergence in Technology (I2CT),* Mumbai (pp. 1199–1202). https://doi.org/10.1109/I2CT.2017.8226316.
6. Ravikumar, S., & Saraf, P. Prediction of stock prices using machine learning (Regression, Classification) agorithms. In *2020 International Conference for Emerging Technology (INCET).* https://doi.org/10.1109/INCET49848.2020.9154061.
7. Yuan, X., Yuan, J., Jiang, T., & Ain, Q. U. (2020). Integrated long-term stock selection models based on feature selection and machine learning algorithms for China stock market. *IEEE Access, 8*, 22672–22685. https://doi.org/10.1109/ACCESS.2020.2969293
8. Bhattacharjee, I., & Bhattacharja, P. (2019). Stock price prediction: a comparative study between traditional statistical approach and machine learning approach. In *2019 4th International Conference on Electrical Information and Communication Technology* (EICT), Khulna, Bangladesh (pp. 1–6). https://doi.org/10.1109/EICT48899.2019.9068850.
9. Zhang, Y., & Yang, S. (2019). Prediction on the highest price of the stock based on PSO-LSTM Neural Network. In *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, Xiamen, China (pp. 1565–1569). https://doi.org/10.1109/EITCE47263.2019.9094982.
10. Wang, J., Sun, T., Liu, B., Cao, Y., & Wang, D. (2018). Financial markets prediction with deep learning. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Orlando, FL (pp. 97–104). https://doi.org/10.1109/ICMLA.2018.00022.
11. Wijaya, Y. B., Kom, S., & Napitupulu, T. A. (2010). Stock price prediction: Comparison of Arima and artificial neural network methods - An Indonesia Stock's Case. In *2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies*, Jakarta (pp. 176–179). https://doi.org/10.1109/ACT.2010.45.