A Survey on Machine Learning and Deep Learning Models for Predicting Stock Movements

Priya Dohre[1], Prof. Pankaj Rahguwanshi[2]

Abstract: The stock prices depend both on time as well as associated variables and finding patterns in among the variables aid forecasting future stock prices which is often termed as stock market forecasting. Stock market prediction extremely challenging due to the dependence of stock prices on several financial, socio-economic and political parameters etc. For real life applications utilizing stock market data, it is necessary to predict stock market data with low errors and high accuracy. This needs design of appropriate artificial intelligence (AI) and machine learning (ML) based techniques which can analyze large and complex data sets pertaining to stock markets and forecast future prices and trends in stock prices with relatively high accuracy. This paper presents a comprehensive review on the various techniques used in recent contemporary papers for stock market forecasting.

Keywords: Time Series Models, Stock Market Forecasting, Artificial Intelligence, Artificial Neural Networks, Forecasting accuracy.

I. INTRODUCTION

The stock market movement is extremely volatile and dependent on a multitude of variables. The unpredictability in the nature of the stock markets make it a risky proposition [3]. Stock market prediction is fundamentally a regression problem in which patterns in previous data and its associated variables need to be found. Stock prices can be mathematically modelled as a time series function as:

$$\boldsymbol{Prices = function\ (time, variables)}$$

The stock prices depend both on time as well as associated variables and finding patterns in among the variables aid forecasting future stock prices which is often termed as stock market forecasting. Stock market prediction extremely challenging due to the dependence of stock prices on several financial, socio-economic and political parameters etc. For real life applications utilizing stock market data, it is necessary to predict stock market data with low errors and high accuracy. This needs design of appropriate artificial intelligence (AI) and machine learning (ML) based techniques which can analyze large and complex data sets pertaining to stock markets and forecast future prices and trends in stock prices with relatively high accuracy.

II. MACHINE LEARNING MODELS FOR STOCK MARKET FORECASTING

Machine Learning models are employed for data analysis where the data to be analyzed is extremely large and complex or both. Primarily, Artificial Intelligence and Machine Learning (AI and ML) have been extensively used for financial and business applications where large data has to be analyzed. One such major area is investment banking [1]-[2]. In such applications, it is necessary to estimate the future movement in stock market prices. Several decisions pertaining to investments, shares etc. depend on the behavior of the stocks of a company. The stock price values are often leveraged by financial and investment firms for gaining profits and investing.

III. LITERATURE REVIEW

The contemporary work in the domain and the noteworthy contributions is cited in this section.

S. Kim et al. in [3] developed a technique termed as effective transfer entropy (ETE) to be used in conjugation with existing ML algorithms such as LR, MLP, LSTM etc. The ETE metric served as an exogenous feature which helped the training performance of the standard training models based on the entropy measure of the data set which is a stochastic variable of the training data set, while the data set used for the study was the US stick market dataset.

B. Bouktif et al. in [4] designed an N-grams based approach utilizing the semantic analysis of data related to stock movement for the prediction problem. The sentiment polarity was utilized to predict the impact of the users of different social media platforms on the stock prices. The polarities used were positive, negative and neutral which served as tokenized impacts on the feature values of the dataset.

X.Li et al. in [5] devised a deep learning model employing sentiment analysis results to predict the stock market behaviour. The individual and cumulative impact of the sentiment features were used for designing the sentiment vector for the forecasting model. Textual normalization and opinion mining techniques were incorporated as features to gauge the sentiments of the common public regarding reputations of the firms since previous prices alone may not always render the moving trends in the market

Gaurang Bansal et al. in [6] proposed a de-centralized forecasting model incorporating block chain which acts as a distributed ledger for stock market behaviours. The use of block chain was done to relate the variables or features for training the system model to find trends or visible patterns in the data blocks. The performance evaluation of the system was done in terms of the accuracy of prediction.

Jithin Eapen et al. in [7] proposed a pipeline approach of CNNs along with a bi-directional LSTM model. The authors were able to gain significant performance improvement in the prediction accuracy of the system using the pipeline CNN model as compared to the baseline regression models for the same S & P dataset. The bidirectional LSTM model was also tested for prediction accuracy of the system for the same data base. It was shown that the pipelined CNN based approach outperformed the conventional techniques.

Min Wen et al. in [8] proposed a stacked CNN based approach for the analysis of noisy time series data for stock market behavioural patterns. The stacked CNN structure was able to extract different levels of features for the different layers of the data set. The proposed technique was shown to perform better than existing techniques for temporal stock market behavioural data patterns.

Y Guo et al. in [9] proposed a modified version of the support vector regression (SVR) model with weight updating mechanism based on evolutionary algorithms such as the PSO. The inclusion of PSO helped in finding the local and global best feature values while optimizing the objective function simultaneously. It was shown that the proposed approach could outperform the existing regression or backpropagation models.

MS Raimundo et al. in [10] proposed a technique that was the amalgamation of the wavelet transform and the support vector regression. The technique used the DWT as a data processing tools and retained the approximate co-efficient values of the multi-tree level DWT analysis of the raw data thereby enabling more noise immunity for the SVR algorithm. The DWT-SVR hybrid was shown to perform better in terms of performance accuracy w.r.t. to SVR alone.

Y Baek et al. in [11] designed a deep neural network named ModAugNet. The deep neural network was again an amalgamation of two LSTM layers. The first layer avoided the chances of overfitting while the second LSTM block was used purely for prediction. The approach was novel in the sense that a similar network with different Hyperparameters were used for the optimization and prediction purposes.

S.Selving et al. in [12] utilized different data fitting algorithms for stock movement estimates. The data fitting approaches utilized were both linear and non-linear in approach such as ARIMA, GARCH etc. The exogenous input feature vector was the closing price of the day which served as a separate feature value. The performance for the same data set with and without the closing price as an exogenous input was tested on the system performance.

Z. Zhao et al. in [13] developed an approach which utilized different time-weighted feature vectors to train an LSTM neural net. The essence of the proposed approach was the fact that recent time or temporal sample has different weighted values compared to the generalized weighted values of a normal feature vector. The performance of the system was evaluated in terms of the accuracy achieved. The designed system achieved an accuracy score of 83.91% while feeding the system with refined feature values.

DMQ Nelson et al. in [14] proposed an LSTM based model for stock market prediction along with technical analysis indicators. The fundamental approach of the system was to find the correlation among different variables for stock market movement. The price indicators of a particular company in a specific stock market were linked to the stock market in other stock markets listed globally. The effect of closing prices of one stock in a particular stock exchange was linked to the opening price of the same stock in some other stock exchange. Thus along with the historical data, the correlation among other variables was also evaluated.

M. Billah et al. in [15] designed a back-propagation based neural network training algorithm with data structuring. The Levenberg Marquardt (LM) weight updating rule was used to forecast closing stock prices of stocks for the Dhaka Stock Exchange. It was shown that the LM algorithm needed lesser memory consumption as well as iterations compared to conventional neural networks and ANFIS systems. The performance evaluation metric was the accuracy of the system.

H.J. Sadei et al in [16] proposed a fuzzy time series predictor based on the concept of fuzzy expert systems. The fuzzy set creation based on temporal data was dine followed by the design of membership functions. Finally, the fuzzy relationships were computed and the defuzzification block was used to predict future trends in the stock prices. Different membership functions were used for the purpose of designing the fuzzy sets.

G.R.M.Lincy et al. in [17] proposed a model based on multiple fuzzy inference systems and applied it to the NASDAQ stock exchange data. The proposed system was pitted against the conventional ANFIS systems and it was shown that the proposed system outperformed the conventional techniques based on expert systems.

Table.1 Summary of Previous Work

S.No.	Authors	Findings	Research Gaps
1	Kim et al.	Effective transfer entropy (ETE) used in conjugation with existing ML algorithms such as LR, MLP and LSTM.	Lack of Data Optimization and Dimensional Reduction.
2	Bouktif et al.	Opinion Mining and Sentiment Analysis was used along with historical stock prices for market prediction.	No data filtering or optimization approach used.
3	Li et al.	Textual normalization and opinion mining techniques were incorporated as features to gauge the sentiments of the common public regarding reputations of stocks.	No data filtering approach
4	Bansal et al.	Model proposed de-centralized forecasting model incorporating block chain which acts as a distributed ledger for stock market behaviours. The use of block chain was done to relate the variables or features for training the system.	No data optimization
5	Eapen et al.	A pipelined approach of CNNs along with a bi-directional LSTM model.	Dimensional reduction and data optimization not used. Opinion mining not employed.
6	Wen at al.	Stacked CNN based approach for the analysis of noisy time series data for stock market behavioural patterns.	No estimation of overfitting for the CNN model.
7	Guo et al.	Adaptive Support vector regression (SVR) model with weight updating mechanism based on Particle Swarm Optimization.	Opinion mining and filtering of data not employed. Performance SVR generally saturates after which adding more training data doesn't increase accuracy of the system.
8	Raimundo et al.	Combination of discrete wavelet transform (DWT) and the support vector regression. The DWT was used as a data filtering tool.	SVR'r performance doesn't improve above a threshold. Opinion mining data not considered.
9	Baek et al.	An amalgamation of two LSTM layers performed. The first layer avoided the chances of overfitting while the second LSTM block was used for prediction.	No data optimization and sentiment analysis data used.
10	Selving et al.	The approach used made predictions based on the daily closing price. The models used ARIMA, GARCH, LSTM, RNN and Sliding window CNN.	Only daily closing price chosen as the time dependent feature. Very restricted feature set.
11	Zhao et al.	Proposed model used time-weighted feature vectors to train an LSTM neural net	No estimates of overfitting or vanishing gradient.

IV. PERFORMACNE METRICS

The parameters which can be used to evaluate the performance of the ANN design for time series models is given by:
1) Mean Absolute Error (MAE)
2) Mean Absolute Percentage Error (MAPE)
3) Mean square error (MSE)

The above mentioned errors are mathematically expressed as:

$$MAE = \frac{1}{N}\sum_{t=1}^{N} |V_t - \hat{V}_t|$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N} |e_t|$$

$$MAPE = \frac{100}{N}\sum_{t=1}^{N} \frac{|V_t - \hat{V}_t|}{V_t}$$

$$MSE = \frac{1}{N}\sum_{t=1}^{N} {e_t}^2$$

Here,

N is the number of predicted samples
V is the predicted value
$\hat{V}_t$ is the actual value
e is the error value

CONCLUSION

This paper presents a comprehensive review and taxonomy on the use of machine learning based approaches for stock market prediction or forecasting. As the data to be analysed is extremely large and complex, hence it is mandatory to employ machine learning for regression analysis. The multiple machine learning and deep learning approaches used in contemporary work have been cited and the related research gaps have been identified. It is expected that this paper puts future research directions in better stead with an aim to enhance the forecasting accuracy.

References

1. Martin T. Hagan, Howard B. Demuth, Mark H. Beale, Orlando De Jesus, "Neural Network Design", 2nd edition, Cengage Publications.
2. Shalev-Shwartz, Shai, Ben-David, "Understanding Machine Learning: From Theory to Algorithms", Cambridge University Press.
3. S Kim, S Ku, W Chang, JW Song, "Predicting the Direction of US Stock Prices Using Effective Transfer Entropy and Machine Learning Techniques", IEEE Access 2022, Vol-8, pp: 111660 – 111682. DOI: 10.1109/ACCESS.2020.3002174
4. S Bouktif, A Fiaz, M Awad, Amir Mosavi, "Augmented Textual Features-Based Stock Market Prediction", IEEE Access 2021, Volume-8, PP: 40269 – 40282. DOI: 10.1109/ACCESS.2020.2976725
5. X Li, P Wu, W Wang, "Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong", Information Processing & Management, Elsevier 2020. Volume 57, Issue 5, pp: 1-19. https://doi.org/10.1016/j.ipm.2020.102212
6. Gaurang Bansal; Vikas Hasija; Vinay Chamola; Neeraj Kumar; Mohsen Guizani, "Smart Stock Exchange Market: A Secure Predictive Decentralized Model", 2019 IEEE Global Communications Conference (GLOBECOM), IEEE 2019 pp. 1-6. DOI: 10.1109/GLOBECOM38437.2019.9013787
7. Jithin Eapen; Doina Bein; Abhishek Verma, "Novel Deep Learning Model with CNN and Bi-Directional LSTM for Improved Stock Market Index Prediction", 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), IEEE 2019 pp. 0264-0270. DOI: 10.1109/CCWC.2019.8666592
8. Min Wen; Ping Li; Lingfei Zhang; Yan Chen, "Stock Market Trend Prediction Using

High-Order Information of Time Series", IEEE Access 2019, Volume 7, pp : 28299 – 28308.
DOI: 10.1109/ACCESS.2019.2901842

9. Y Guo, S Han, C Shen, Y Li, X Yin, Y Bai, "An adaptive SVR for high-frequency stock price forecasting", Volume-6, IEEE Access 2018, pp: 11397 – 11404.
DOI: 10.1109/ACCESS.2018.2806180
10. MS Raimundo, J Okamoto, "SVR-wavelet adaptive model for forecasting financial time series", 2018 International Conference on Information and Computer Technologies (ICICT), IEEE 2018, pp. 111-114.
DOI: 10.1109/INFOCT.2018.8356851
11. Y Baek, HY Kim, "ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module" Journal of Expert System and Applications, Elsevier 2018, Volule-113, pp: 457-480.
https://doi.org/10.1016/j.eswa.2018.07.019
12. S Selvin, R Vinayakumar, E. A Gopalakrishnan ; Vijay Krishna Menon; K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model", 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE 2017, pp. 1643-1647.
13. Z Zhao, R Rao, S Tu, J Shi, "Time-weighted LSTM model with redefined labeling for stock trend prediction", 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), pp. 1210-1217.
DOI: 10.1109/ICTAI.2017.00184
14. DMQ Nelson, ACM Pereira, Renato A. de Oliveira , "Stock market's price movement prediction with LSTM neural networks", 2017 International Joint Conference on Neural Networks (IJCNN), IEEE 2017, pp. 1419-1426
DOI: 10.1109/IJCNN.2017.7966019
15. M Billah, S Waheed, A Hanifa, "Stock market prediction using an improved training algorithm of neural network", 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), IEEE 2016, pp. 1-4,
doi: 10.1109/ICECTE.2016.7879611
16. HJ Sadaei, R Enayatifar, MH Lee, M Mahmud, "A hybrid model based on differential fuzzy logic relationships and imperialist competitive algorithm for stock market forecasting", Journal of Applied Soft Computing, Elsevier 2016, Volume 40, pp: 132-149.
https://doi.org/10.1016/j.asoc.2015.11.026
17. GRM Lincy, CJ John, "A multiple fuzzy inference systems framework for daily stock trading with application to NASDAQ stock exchange", Journal of Expert Systems with Applications, Volume-44, Issue-C, ACM 2016.
DOI: 10.1109/ICACCI.2017.8126078