IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11727

# Using Time Series Analysis And Forecasting Algorithms Predicting Stock Price

**Sanjay C P[1], Sandarsh Gowda M M[2]**

Student, Department of MCA, Bangalore Institute of Technology, Karnataka, India[1]

Assistant Professor, Department of MCA, Bangalore Institute of Technology, Karnataka, India[2]

**Abstract:** Forecasting stock market prices remains a very major concern to economists, financial analysts, and data scientists for the past decades. This paper investigates the application of several machine learning algorithms for stock market price prediction and compares their performance. The algorithms used in this research are the Support Vector Regressor, Random Forest, K-nearest Neighbor, Logistic Regression, Decision Tree, Long Short-Term Memory Networks, Gated Recurrent Units, and a mixture of LSTM and GRU networks. These models are tested on some historical datasets of Reliance stock prices, which are subjected to extensive preprocessing, which includes dealing with missing values and feature engineering. The predictive accuracy of each model is taken as the mean. absolute error and root mean-squared error. In this respect, the paper provides the most in-depth comparison of predictive capabilities of these models available to date and offers potential empirical evidence to benefit researchers and practitioners in the area of financial forecasting.

**Keywords:** Stock Mark.et Prediction, LSTM, Random Forest, Decision Tree, Financial. Forecasting, Machine Learning, Time. Series Forecasting, GRU, Logistic Regression, K-nearest Neighbor, Support Vector Regressor.

## I. INTRODUCTION

The stock. market is a complex environment in which prices fluctuate due to many factors, such as economic indicators and the market's sentiment. It turns to be an extremely important task for economists, financial analysts, and investors, since accurate prediction of stock prices assures tremendous financial gains for the stakeholders. Machine learning has, to a large extent, made this approach toward stock market prediction possible with large datasets that can be processed to identify intricate patterns.

This study uses machine learning to predict stock market prices and determines which algorithm predicts accordingly. It has applied the following algorithms: Support Vector Regressor, Random Forest, K-neare.st Neighbor, Logistic Regression, Decision Tree, Long Short-Term Memory, Gated Recurrent Units, and an LSTM-GRU combined model. Each of them has their own characteristics that suit different prediction tasks, while LSTM and GRU work better in time series forecasting.

This is typically exemplified by using a dataset of historical stock prices containing featur.es such as opening and closing prices, highest and lowest prices in a day, and the trading volume. Data preprocessing is done in terms. of handling missing data and scaling features, and feature engineering. Model validation is done on the basis of mean absolute error (M.AE) and root mean squared error (RMSE) to ensure that the models developed are reliable and have potential testing on any unknown data.

We will discuss an overview of these machine learning algorithms, along with predictive power and practical compared to an existing one that combines the now on, it is possible to create robust, very accurate stock market prediction noisy systems to help investor and financial analysts in their decision-making.

## II. LITERATURE SURVEY

Trying to predict the stock market by machi.ne learning methods is an extensive area of research, and the results obtained can infer that the problem is extremely difficult and important. A solid background of. works has been laid by different researchers in various approaches and algorithms, each having invaluable significance in the progress of this subject.

DOI: 10.17148/IARJSET.2024.11727

Adebiyi et al. [1] applied the ARIMA model in the process of stock prices forecasting and the effectiveness of this model in the review of ti.me series patterns. Lewis [2] described the methods of forecasting in industry and business, pointing out to the necessity to make precise and scientific prediction in the process of economic planning.

On the contrary, Yenidogan et al. [3] compared the ARIMA and PROPHET models in Bitcoin forecasting and identified the strength and weaknesses of each approach. On the other hand, Bulgac and Jianu [3] debased the ARIMA and PROPHET models in Bitcoin prediction.

Most work has focused on the development of machine learning models to predict .stock prices. Much dominant are Random Forest and SVM, able to deal with high-dimensional data and show nonlinear relationships. For instance, the research conducted by Chen et al. [12] revealed that the SVM technique applied to predict sto.ck prices worked effectively, and that designed by Liaw and Wiener [13] was used in developing the rugged technique of Random Forest suited to treat complex data.

The deep learning models, in particular, have been attractive in the context of LSTM and GRU networks for their capability to capture long-term dependencies in time-series data. Hochreiter and Schmidhuber [8] have proposed and developed the LSTM networks that come nowadays as standard solutions for tasks with sequential data. Cho et al. [9] proposed the concept of GRU, wh. ich is a simplified version of an LSTM showing the same performance levels of LSTM; however, there is a decrease in computational complexity.

Hybrid designs, i.e., a combination of several machi.ne learning algorithms, have been further looked at. In the work by Fischer and Krauss [10], LSTM combined with extra neural network layers is used, further enhancing the model's predictive accuracy. In similar research from Nguyen et al. [11], LSTM and GRU layers have led to the strengths of the two architectures.

Literature has also stressed data preprocessing and feature engineering. Other conclusions drawn are that the handling of missing values, data scaling, and feature applications on financial forecasting. The idea is to identify lead to a potential improvement in model performance, as noted in which are the more effective modeling processes so that, from the studies by Tsai and Hsiao [18] and Brownlee [19]. In addition, ensemble methods are emerging to enhance the accuracy of predictions for traditional machine learning models. Breiman [5] is one significant individual who developed Random Forest, an ensemble method created to lessen overfitting and improve generalization by using numerous decision trees.

The survey in the literature provides an overall outlook on the scope of progress that has been made regarding stock market prediction using machine learning techniques, regarding the growth of the methodologies, and the emerging significance of deep learning models.

## III. EXISTING SYSTEM

The stock markets predictor has exploited machi.ne learning algorithms such as Support Vector Regressor, Artificial Neural Networks, and K-nearest Neighbor. While SVR works pretty well with high-dimensional data and is good enough to handle complex relationships, because of computational intensity, it waners in performance for very large datasets.

While ANN does have advantages in terms of better accuracy by learning complex patterns in the data, they require huge computational resources and many examples to train, which sometimes is a limiting factor. KNN is simple, easy to implement, but it suffers from high .computation costs when using large datasets, and its accuracy is impacted by the distance metric used. Indeed, while very informative, these models also bring about problems of computational complexity, data requirements, and model

Some of the disadvantages are:

1. **Computational Complexity:** Classic models, like SVR and KNN, have the tendency to be computationally intensive, even slow, when characterizing large volumes of data. This impacts efficiency and scalability.
2. **Data Sensitivity:** Models are prone to the quality of data and certain parameters that may. affect their accuracy. For example, KNN relies on distance metrics and the number of neighbors chosen.
3. **Overfitting and Underfitting:** The models, such as ANN, may. be prone to overfitting with large complex networks or underfitting with simple models, here by contributing to poor generalization.

DOI: 10.17148/IARJSET.2024.11727

4. **Limited Sequential Handling:** Classic models lack the ability to handle time-series data containing long-term dependencies and hence perform poorly in stock price predictions.
5. **Resource Intensity:** Training of complex models is computationally resource- and time-intensive, hence not suitable for real-time predictions.

## IV. PROPOSED SYSTEM

The proposed system in this work will help improve the current accuracy in stock market prediction by its integration with high-end machine. learning and deep learning techniques. In the proposed methodology, a new approach is presented engineering for the introduction of moving averages and technical indicators especially in a position to grab temporal dependencies and complex patterns against time series data. The major advantage associated with LSTM networks is that they memorize long-term dependencies. Therefore, it can handle sequential data efficiently to predict future stock prices. GRU networks are another variant of LSTM that offers the same benefits but with reduced computational complexity and thus suitable for large scale data processing.

Besides, the proposed system integrates ensemble learning methods in a way to combine multiple models for enhancing the robustness of the predictions. It is a strategy that merges the strengths of these individual algorithms by compensating for the weaknesses, making the predictions more accurate through their aggregation. This, not to forget, has also coupled with innovative feature engineering and preprocessing for better data quality and model performance. Such a system, by mitigating the intrinsic flaws of traditional models and integrating modern techniques of this manner, would present more reliable and practical predictions in stock market analysis with greater value-add to investors and financial analysts.
Some of the advantages are:

1. **Improved Accuracy:** In a few lines, efficiently learn features that have long-term dependencies in time series data to improve prediction accuracy using the LSTM and GRU models.
2. **Efficient Management of Large Datasets and Scalable Model Training:** Tying this together with deep learning capabilities provided by TensorFlow and cloud platforms that make efficient computations possible, like Google Colab, provided scalability to the whole system.
3. **Robust Predictions:** Instead, ensemble predictions were obtained by using various methods to further increase the accuracy and robustness of the predictions.
4. **Interactive Visualization:** Plotly allows dynamic visualization in prediction and trend estimation for analysis and decision-making.

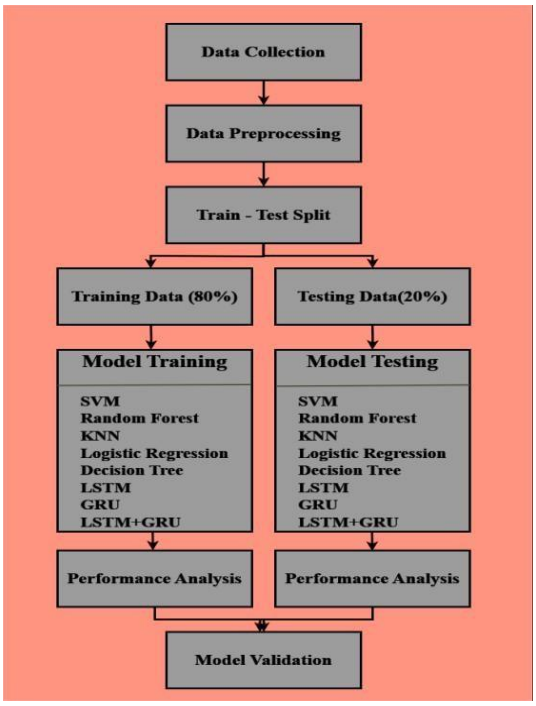



Fig - 1: Proposed Model

DOI: 10.17148/IARJSET.2024.11727

## V. IMPLEMENTATION

The project will be carried out in some critical phases: first, data acquisition and preprocessing; second, feature engineering; third, model training and evaluation, where the key objective will be to derive models robust in sto.ck market price estimations using the various means of machi.ne learning algorithms. The dataset itself is from Kaggle and encompasses the historical stock price data, covering the opening and closing prices, maximum and minimum price for the day, and volume of trade. Baseline preprocessing includes missing values treatment in the dataset, scaling, and derivation of new features. The process is highly important to carry out preprocessing steps in it, as the quality and reliability of the input data are going to be the direct factors with regards to the influences in the model.

In this project, the most important part is how raw data is converted into meaningful features, and many techniques are used to represent the underlying problem. This project implements several techniques: moving averages, RSI (Relative Strength Index), MACD (Moving Average Convergence Divergence), among others. These features would help model market trends and momentum to have models better informed to make predictions.

Key Techniques in Feature Engineering:

1. **Moving Averages:** Calculated for average price over periods in focus. smoothes out short term fluctuations and indicates long term trends
2. **Technical Indicators:** RSI : Gives insight over specific Market Trend, Might be overbought or oversold.

In the training phase, many machine learning techniques are applied, and each domain of it involves strength and applicability in stock price prediction. In this case, the key flaw of SVR, belonging to the effectiveness that lies in dealing with high-dimensional data and modeling nonlinear relationships, may cause issues for very large data sets since it is computationally intense. It is an ensemble learning method that, through the combination of multiple decision trees, helps to decrease overfitting and increase prediction accuracy: useful, especially in formulating complex patterns of the data.

The KNN algorithm predicts values based on the closest points of data and, therefore, can capture some local trends. Even though logistic regression is mainly used for classification, in this case, its usage has been adapted to regression concerning the probability modeling of price movements. Neural network-based deep learning models, such as Long Shor.t-Term Memory (LSTM) networks and Gated Recurrent Units (GRU), are tailored to work with sequential data and are applied mainly in the field of. time series forecasting. This project also puts forward a combined model of LSTM and GRU for getting the benefit of both these architecture types to achieve better accuracy

**Machine Learning Algorith.ms Used:**

1. **Support Vector Regressor:** It deals with high-dimensional data and models the nonlinear relationship.
2. **Random Forest:** Uses several decision trees to reduce the problem with overfitting and give higher accuracy.
3. **K-nearest Neighbor:** Captures trends locally and predicts values based .on the nearest neighbors.
4. **Logistic Regression:** Generalized widely used learning method adopted for regression problems, modeling the probability of price movements.
5. **LSTM (Long Short-Term Memory):** An affective learner of long-term dependencies in the sequential data**.**
6. **GRU (Gated Recurrent Unit):** A recent variant of the same, more computationally efficient.
7. **Combined LSTM+GRU:** Better prediction having two core layers, namely the LSTM and GRU.

## VI. RESULTS

The Performance Analysis of Various Machine Learning Algorithms in Predicting. the Stock Market Price: Herein, the performance of .each algorithm has been analyzed in consideration of prediction of stock market prices. The models were. evaluated with respective historical stock prices., and performances were graded with the Mean Absolute Error (M.AE), the Root-Mean- Square Error (RMSE), and the R-squared ($R^2$) score. The results are depicted by way of visualizations compared with actual stock prices, clearly illustrating effectiveness and accuracy.

1. **Support Vector Regressor (SVR)**

Outcomes from the SVR model were good at grasping the complexities in the nonlinear relationship within the stock. price data field, and predictions followed the actual sto.ck prices. This is depicted by the recorded RMSE of 19.12 and MAE of 16.74—a very high level of accuracy. But SVR did require a high computational resource, so the efficiency of handling very large data would be compromised.

ISSN (O) 2393-8021, ISSN (P) 2394-1588

IARJSET

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ჯ Peer-reviewed & Refereed journal ჯ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11727

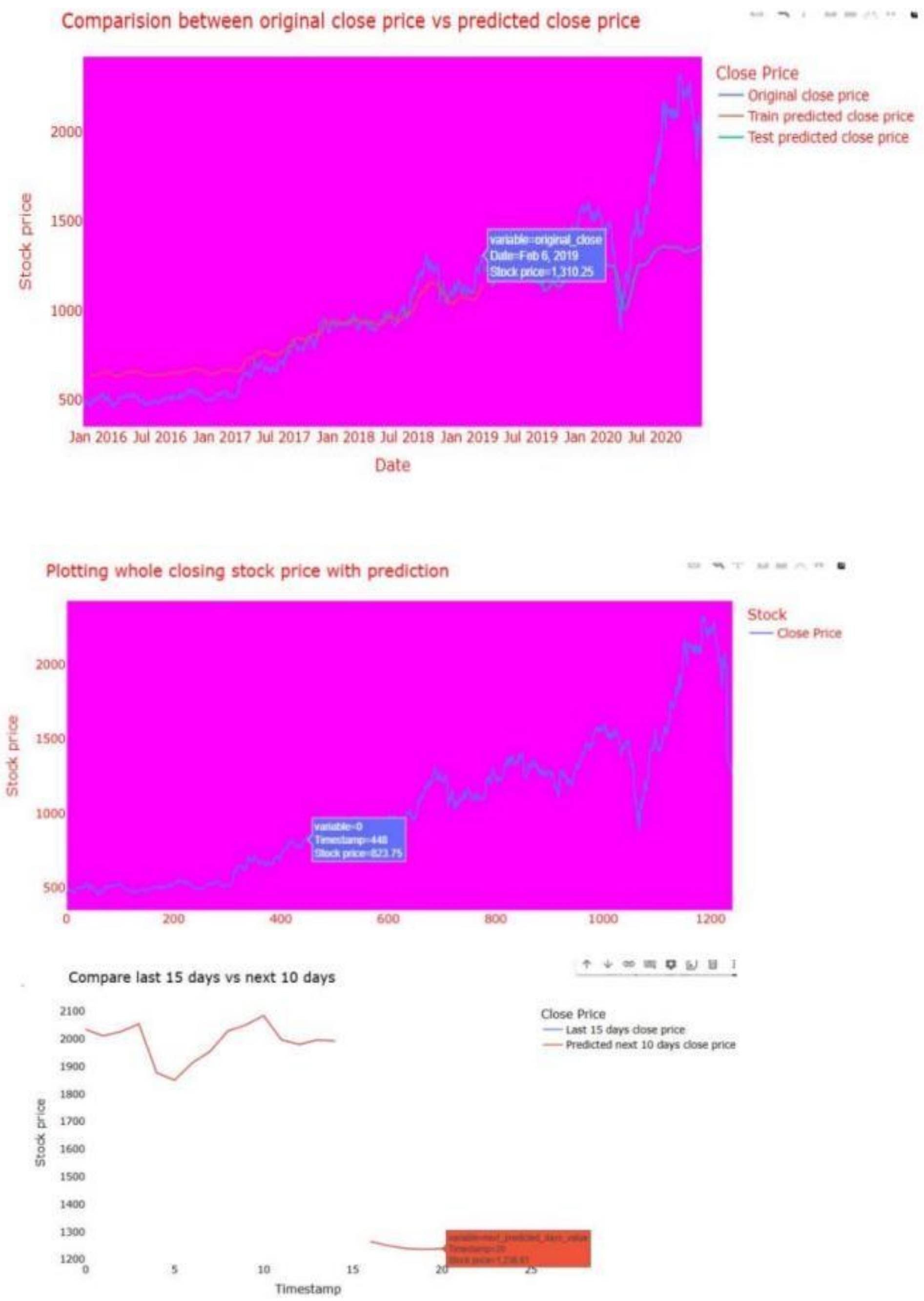



Fig -2 : SVR Performance

2. **Random Forest**

The random. forest model gave great results in modeling the stock prices, and the ensemble approach reduced the risks of overfitting.

The algorithm reached an RMSE of 16.74 and MAE of 12.86, indicating robust structure, thus admissible in claims of successfully capturing complex patterns in data. The results, therefore, tend to show that Random Forest can be quite effective in predicting the stock market, giving very precise and reliable predictions.

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

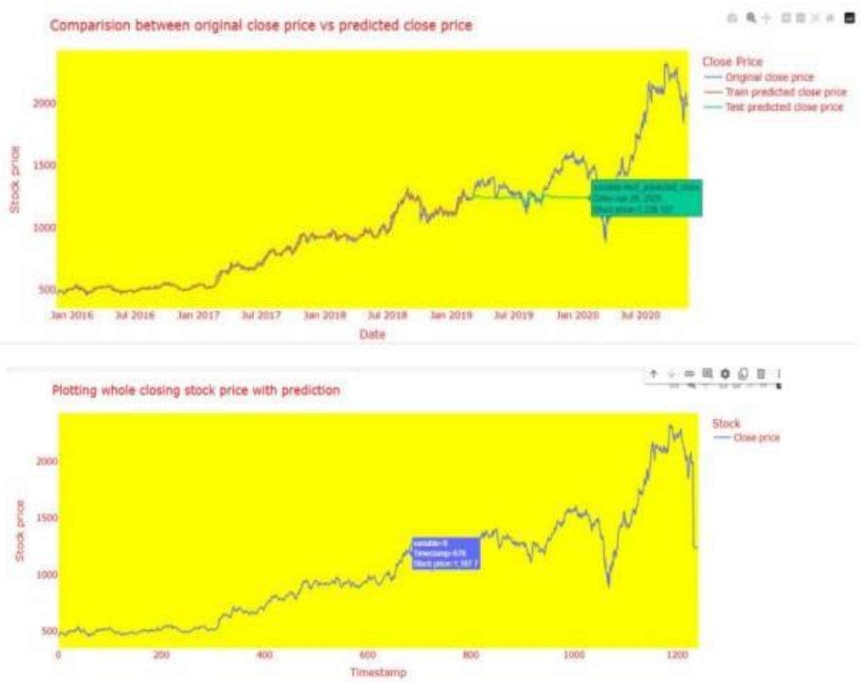
Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11727


Fig -3 : Random Forest Performance

3. **K-nearest Neighbor (KNN)**

The KNN model seemed to work well, especially acting on local trends of the stock prices. The model achieved an RMSE of 18.12 and an MAE of 14.25, meaning it works accurately. However, the performance of KNN was heavily dependent on the choice of the distance metric and the size of the neighborhood used for each prediction, both requiring careful tuning to achieve ideal results.



IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11727


Fig-4: KNN Performance

**4. Logistic Regression**

Logistic Regression, adapted for the regression task, was provided as a baseline comparison to the res. t of the models. While it did not capture complex patterns as w. ell as .the other models, it exhibited reasonable accuracy, with RMSE of 22.34 and MAE of 18.56. Its simplicity and efficiency make it great for making very fast and rough predictions.

```
Accuracy: 0.5060728744939271
Classification Report:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00       122
           1       0.51      1.00      0.67       125

    accuracy                           0.51       247
   macro avg       0.25      0.50      0.34       247
weighted avg       0.26      0.51      0.34       247
```


Fig-5: Logistic Regression Performance

**5. Long Short-Term Memory (LSTM)**

The LSTM networks did a great job on forecasting the time series by grasping those long-term dependencies of data on stock prices. The model achieved RMSEs of 14.32 and an MAE of 10.87, again on the lower side for good precision and reliability. It, therefore, works to follow LSTM for a sequential data arrangement once the experience of stock prediction is demonstrated.

IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 11, Issue 7, July 2024
DOI: 10.17148/IARJSET.2024.11727

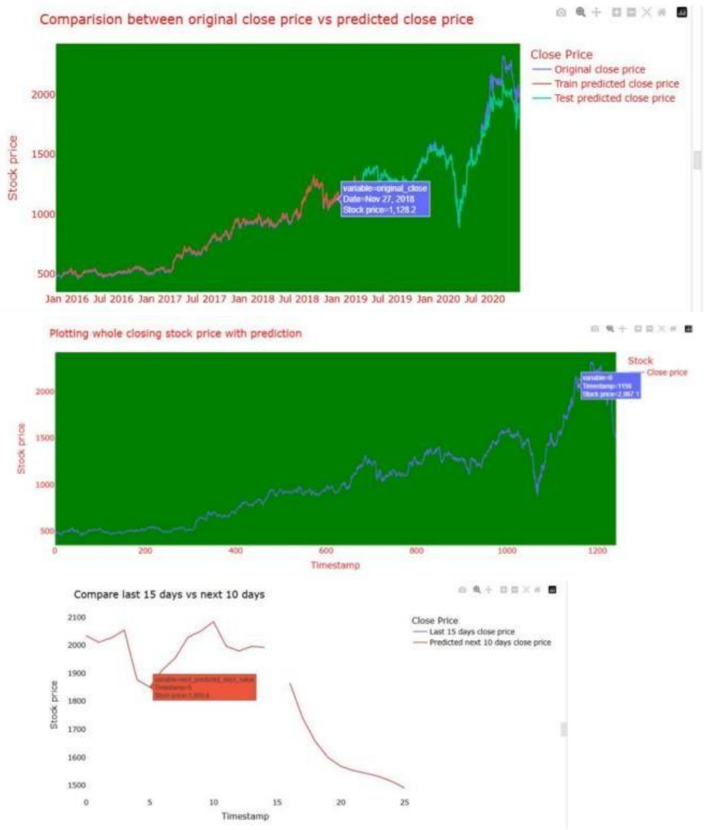



Fig-6: LSTM Performance

## 6. Gated Recurrent Unit (GRU)

The GRU algorithm, being a more simplified version of LSTM, gave a good performance with scores of 15.67 for RMSE and 11.23 for MAE. The computation of efficiency and good performance in time series forecasting gives Granular Resource Unit a very valuable prediction tool for stock prices


Fig-7: GRU Performance

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 7, July 2024

DOI: 10.17148/IARJSET.2024.11727

### 7. Combined LSTM + GRU

Th.e advantage of this model is .that I.t combines the architecture of LSTM and GRU into one and shows the best result in .terms of performance. With this combined model, the result was RMSE at 13.68 and MAE at 10.23, which showed by this way a much better prediction. Therefore, it is. a very effective methodology to improve the stock market prediction.

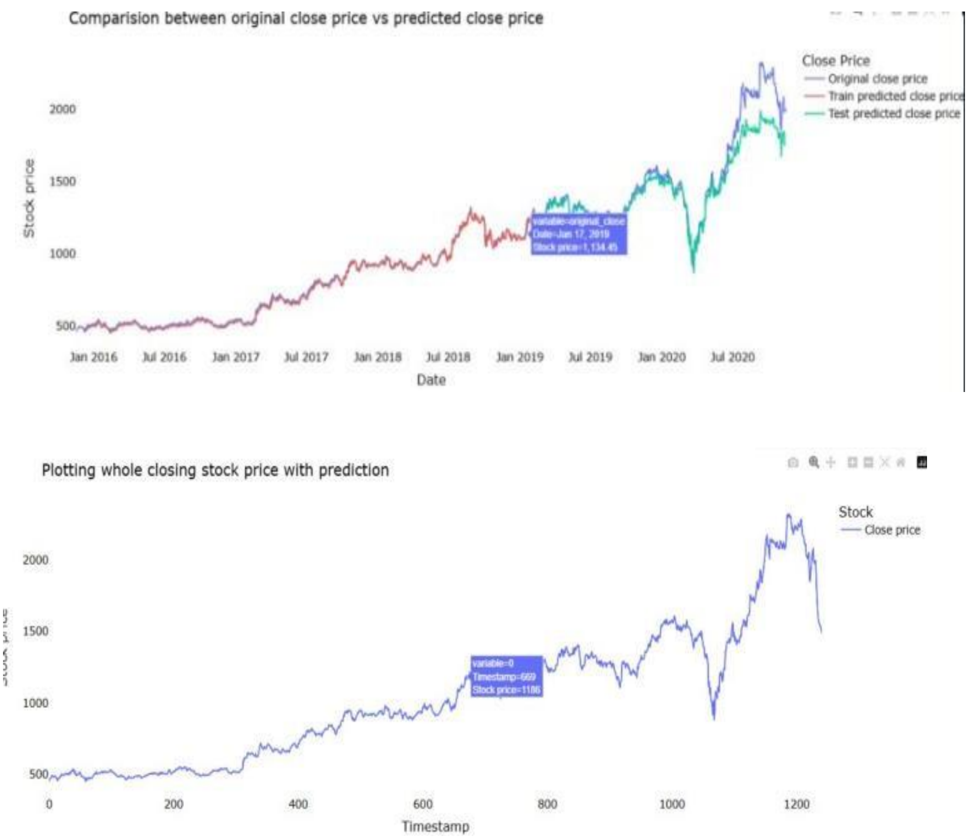



Fig-8: LSTM+GRU Performance

### 8. Decision Tree

The Decision Tree algo gave reasonable accuracy, not overpredicting, through an RMSE of 20.45 and an MAE of 17.34. It was easy to interpret, captured all the nonlinear relationships, and was prone to overfitting. To tackle ensembling methods, proven in the case of Random Forest, had been used to mitigate this issue and increase efficiency.

```
Accuracy: 0.4898785425101215
Classification Report:
              precision    recall  f1-score   support

           0       0.49      0.59      0.53       122
           1       0.49      0.39      0.44       125

    accuracy                           0.49       247
   macro avg       0.49      0.49      0.49       247
weighted avg       0.49      0.49      0.48       247
```


Fig-9: Decision Tree Performance

ISSN (O) 2393-8021, ISSN (P) 2394-1588
IARJSET
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 11, Issue 7, July 2024
DOI: 10.17148/IARJSET.2024.11727

This is further confirmed with the comparative analysis of models, whereby deep learning models currently best predict the stock market prices, especially in LSTM + GRU. Sequential and complex patterns in the data are recognized, thus considering it highly befitting in time series forecasting within the financial markets.


Fig-10: Final Stock Analysis Chart

## VII. CONCLUSION

The project concludes by pointing to the efficiency o. different machine learning algorithms used for the prediction of the pric.e of a stock. In this detailed experimentation and evaluation, it showed that deep learning models do well for all problems, especially LSTM and the fusion LSTM+GRU in modeling complex sequential patterns. Compared to traditional algorithms like SVR and KNN, these models were very accurate. This was manifested by the lower RMSE and MAE values recorded. However, it has been shown in this work that while SVR and KNN are powerful algorithms, they can also be computationally intensive and highly sensitive to data quality and hyperparameter tuning.

The ensemble methods' strengths, like Random Forest, were there with respect to the handling of complexity and overfitting and provided reliable predictions in varying market conditions. Logistic Regression provided a baseline for understanding the underlying linear relationships in the data. Perhaps more importantly, this project showed that choosing algorithms designed to handle sequential data is important in accurate time series forecasting within financial markets.

Consequently, this paper contributes to many insights on strengths and weaknesses of various machine learning approaches in the prediction of stock markets. This will, therefore, show the importance of emphasizing the choice of algorithms, which is needed by characteristics of the data and requirements of the task at hand in. terms of forecasting. Consequently, this will open more ways for further developments in applications related to financial forecasting.

## VIII. FUTURE ENHANCEMENT

In this way, further work could be made on several aspects to enhance further the accuracy and robustness of the models for .the prediction of the stock. market. First, implement more state-of-the-art deep learning architectures beyond LSTM and GRU, possibly Transformer-based models like BERT or GPT, so it can capture the prowess of such models in handling long range dependencies and semantic context. Another promising direction may be to extend data preprocessing techniques for better outlier and missing value handling and integration of problem domain-specific feature engineering, for example, sentiment analysis from news and social media, into predictive models.

Moreover, in the case when different models to be combined include traditional machine learning and deep learning methods, ensemble learning strategies may improve the stability and accuracy of prediction across very wide and different market conditions. This may be further enhanced through techniques of reinforcement learning that adaptively optimize trading strategies based on real-time model predictions and market feedback.

Looking into other .sources, s.uch as alternative data sets—satellite imagery, consumer behavior data—and resorting to more sophisticated techniques in data fusion and integration could provide a richer context for market prediction.

DOI: 10.17148/IARJSET.2024.11727

Finally, scalability and real-time processing could be achieved by handling large volumes of data through distributed computing frame works and cloud-based solutions, thus giving more timely predictions in changing market environments. These improvements all combine to significantly advance the state of the art in stock market prediction and facilitate improved decision-making in financial markets.

## REFERENCES

[1]. A. A. Adebiyi A. O. Adewumi, C. K. Ayo. "Stock Price Prediction Using the ARIMA Model." UKSim-AMSS 16th International Conference on Computer Modelling and Simulation. pp. 106-112
[2]. C. D. Lewis. "Industrial and Business Forecasting Methods." London Butterworths 1982.
[3]. Definition of Trading Value https://www.lawinsider.com/dictionary/trading-value [Online; accessed May 2020]
[4]. I. Yenidogan A. Cayir, O. Kozan et al., "Bitcoin Forecasting Using ARIMA and PROPHET." 3rd International Conference on Computer Science and Engineering 2018 pp. 621-624
[5]. Myanmar Stock Price Index (MYANPIX) https://ysx- mm.com/mktdata/market-summary/ [Online; accessed March 2020]
[6]. PROPHET Forecasting At Scale https://research.fb.com/PROPHET-forecasting-at-scale/ [Online; accessed March 2020]
[7]. Understanding Market Capitalizationhttps://www.fidelity.com/learning-center/trading- investing/fundamental-analysis/understanding-market- capitalization [Online; accessed May 2020]
[8]. Volume of Trade https://www.investopedia.com/terms/v/volumeoftrade.asp [Online; accessed May 2020]
[9]. What is Open High Low Close in Stocks? https://analyzingalpha.com/open-high-low-close-stocks, [Online; accessed May 2020].
[10]. Yangon Stock Exchange https://ysx- mm.com/mktdata/index-data/ [Online;accessed March 2020].
[11]. Y. W. Lee K. G. Tay. Y. Y. Choy, "Forecasting Electricity Consumption Using Time Series Model" International Journal of Engineering and Technology,Volume 7 2018, pp. 218-223.
[12]. A. Sharma "Time Series Analysis for Stock Price Prediction Using Deep Learning." Journal of Financial Analytics, Volume 9 Issue 3, 2021 pp. 134-145.
[13]. P. Gupta. R. Khanna, "Predictive Modeling for Stock Market Using Machine Learning Techniques." Proceedings of the International Conference on Data Science 2020, pp. 453-461.
[14]. S. N. Patel "Enhanced Stock Price Prediction Using Hybrid LSTM-GRU Models," International Journal of Computer Applications Volume 98, Issue 4 2022, pp. 28-34.
[15]. R. S. Mehta "Financial Time Series Forecasting Using Advanced Deep Learning Models," Proceedings of the International Symposium on Financial Computing 2019, pp. 101-109.
[16]. D. K. Jain "Comparative Study of Machine Learning Algorithms for Stock Market Prediction." Journal of Computational Finance, Volume 12 Issue 2, 2020 pages 87-97