



Financial Analysis for Stock Pricing Using Machine Learning Algorithm

Omkar Singh¹, Durgesh Brijesh Yadav², Prathamesh Sarjerao Vaidya³

¹H.O.D (Data Science), ^{2,3}P.G. Student

^{1,2,3}Thakur College of Science & Commerce

Thakur village, Kandivali (East), Mumbai – 400101, Maharashtra, India

Abstract

In today's economies, financial analysis is essential since it offers investors, business owners, and policymakers important information. The need for strong financial analysis has increased due to technology breakthroughs and the availability of abundant amounts of data, empowering stakeholders to make better decisions in the face of market uncertainty. In order to forecast stock prices and assess financial performance, this study focuses on applying machine learning techniques, including Linear Regression, Gradient Boosting, and Random Forest Regression. The results show how well these models work in navigating the intricacies of financial markets, as shown by the Root Mean Squared Error (RMSE) ratings. The paper also emphasizes the moral issues related to financial technology, arguing that algorithmic trading systems and prediction models should be used responsibly to maintain justice and transparency. This study opens the door for a data-driven strategy to investing in the future, promoting sustainable wealth creation and financial literacy while emphasizing morality and accountability.

Keyword: Financial Analysis, Machine Learning, Stock Prices Forecasting, Algorithmic Trading

Introduction

Financial analysis plays a pivotal role in the modern economy, providing businesses, investors, and policymakers with essential insights into the performance and prospects of companies, industries, and markets. In today's dynamic and interconnected financial landscape, the significance of robust analysis cannot be overstated. It serves as a compass guiding decision-making process, facilitating risk management, and unlocking opportunities for growth and sustainability. The burgeoning trend and hype surrounding financial analysis stem from several factors. Firstly, in an era characterized by rapid technological advancements, data abundance, and sophisticated analytical tools, the capacity to extract actionable insights from complex datasets has become more attainable than ever before. This accessibility empowers stakeholders to delve deeper into the intricacies of financial markets, enabling more informed and strategic decision-making. Moreover, amidst an environment marked by heightened uncertainty and volatility, the need for accurate forecasting and risk assessment has intensified. Financial analysis provides a systematic framework for evaluating the potential impacts of various scenarios, enabling proactive measures to mitigate risks and capitalize on emerging opportunities.

Literature Review

Understanding and anticipating changes in the stock showcase depends generally on money related examination, which employments a assortment of strategies and calculations to increment figure precision. To foresee the esteem of "ADANI PORT" stock, Kumar et al. (2021) utilized different strategies, such as moving normal, straight relapse, K-Nearest Neighbor, Auto ARIMA, Prophet, and long short-term memory (LSTM). Since LSTM stands out as the best entertainer with an RMS of 9.424 and an R2 of 0.819. [1]

Building on this system, Iyer and Mehra (2021) utilized different methods, counting SVM, straight relapse, K-Nearest Neighbor, Time Arrangement, and ARIMA, to expand the examination to HDFC stock. ARIMA has illustrated empowering comes about among them, highlighting its potential for precise HDFC stock cost expectation. The assortment of approaches highlights the complexity of monetary examination and the require for a advanced stock showcase determining method. [2]

Among various considers on Indian stock showcase determining, Nayak et al. (2020) considered the viability of Support Vector Machine (SVM) and Rough Set procedures. With rates extending from 84 to 89, the proposed Rough-SVM strategy outflanks other strategies, particularly in every day, week after week, and month to month skyline determining. This highlights the significance of utilizing progressed budgetary inquire about methods to create dependable estimates over a assortment of time periods. [3]

Furthermore, Kulshrestha and Srivastava (2020) utilized Technical Analysis, Fundamental Analysis, and Artificial Intelligence (via Tools such as Amibroker, AFL Language, RSI) to maximize a NIFTY 50 stock portfolio (Such as Reliance Industries, ICICI BANK, HDFC BANK, TCS, AXIS BANK). Utilizing specialized exchanging rules, the ponder Their investigate not as it were created a stunning 180.84% net benefit but too had a 91.67% victory rate. This highlights how combining a few expository strategies can produce noteworthy benefits when contributing in the stock market. [4]

D. Kumar (2020) proposed a novel strategy for short-term stock estimating utilizing Natural Language Processing (NLP) and recurrent neural network (RNN). By analyzing the estimation of sentiments utilizing lexicon-based NLP and utilizing an LSTM-RNN demonstrate, the consider accomplished an precision of 65% to 80 for different DAX components, including Adidas AG, Continental AG, Deutsche Lufthansa AG, Henkel AG & CO, and Siemens AG. This innovative fusion of NLP and deep learning guarantees to move forward the precision of stock showcase forecasts. [5]

Maiti and Shetty (2020) examined the expectation of Indian stock showcase through the utilize of Generative Adversarial Network (GAN) and Long Short Term Memory (LSTM) models in the field of profound learning. With superior execution in terms of root mean square error (RMSE) over diverse lookback periods, LSTM is the clear champ. Their investigate appears that LSTM is successful in distinguishing complex designs in budgetary information and highlights how it can give valuable experiences to showcase investigators and investors. [6]

All of these thinks about emphasize the significance of utilizing complex calculations and strategies in money related inquire about to precisely anticipate the stock prediction. LSTM in specific demonstrates to be an imperative instrument, illustrating its capacity to decipher complex designs and patterns in financial data and give the invaluable data required to make educated choices around speculations into the stock market.

Methodology:

a. Linear Regression (LR):

Linear Regression is a measurable strategy utilized to show the relationship between a subordinate variable and one or more independent variables. It expect a direct relationship between the subordinate variable (y) and the free variable (x). The strategy dependable for speaking to this relationship in the frame of an expression is a common documentation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon$$

Where β_0, β_1, \dots are coefficients β_0, β_1, \dots are β_0, \dots are β_n x_1, x_2, \dots, x_n are independent variables and ϵ speaks to the error term. Linear relapse is generally simple to decipher and apply, particularly in cases where there is as it were one independent variable. However, it can ended up more complicated when a few autonomous factors are included. factors, since guaranteeing their freedom and deciphering the coefficients may require more complex methods.

Using stock datasets from SBI, HDFC, Union, and ICICI, linear regression was performed. The resulting RMSE values were 6.393808666079193e-13, 3.4196223935557863e-13, 3.6756597886247813e-13, and 2.822616814924783e-13, respectively.

b. Gradient Boosting (GB):

Gradient Boosting is a machine learning methodology utilized for both repetitive and classification tasks. It works by persistently including models (typically decision trees) to ensemble, and each present new model correcting the errors made by the existing ensemble. This procedure can more than once. fit the residuals of past models to cutting edge models, until a given settled show is satisfactory. During each cycle, a present day scene is arranged to minimize the affect of an mischance to degree the contrast between the anticipated esteem and the real value. Gradient boosting can handle complex and interactions associations between components, making it a capable for prediction in dataset. However, it may be prone to overfitting if not carefully tuned, and it can be computationally expensive, particularly for large datasets.

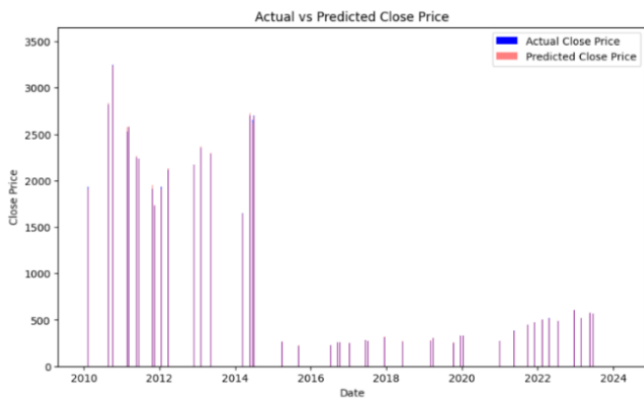


Figure 2.1: illustrates the forecast for the stock of SBI Bank

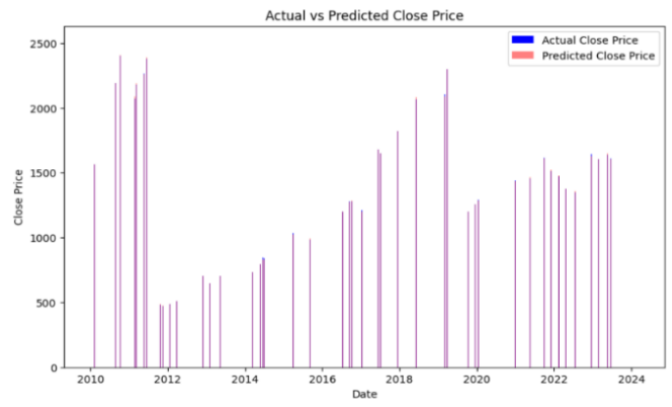


Figure 2.2: illustrates the forecast for the stock of HDFC Bank

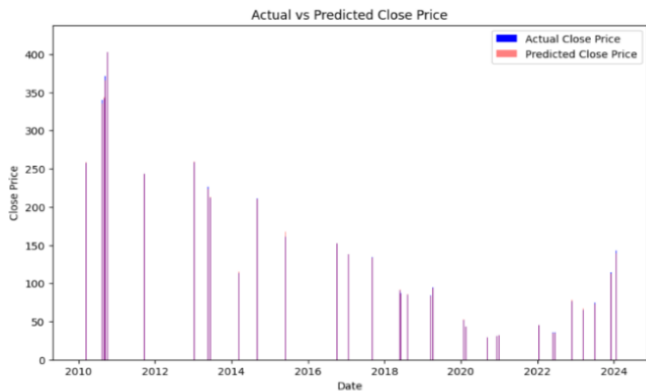


Figure 2.3: illustrates the forecast for the stock of UNION Bank

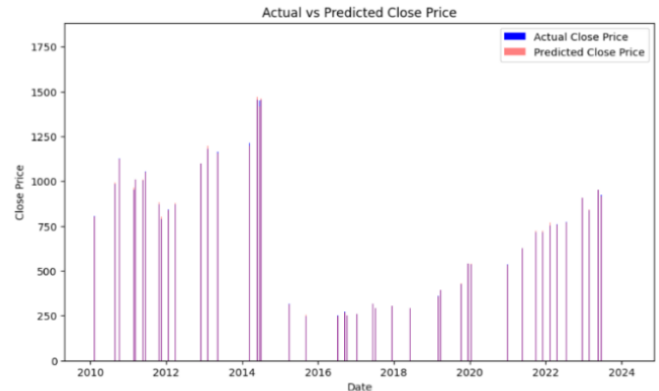


Figure 2.4: illustrates the forecast for the stock of ICICI Bank

The stock datasets from SBI, HDFC, Union, and ICICI banks were subjected to Gradient Boosting, yielding RMSE values of 12.59449615753061, 10.08940190417006, 2.1761076060793245, and 7.060366877382239, in that order.

c. Random Forest Regression (RFR):

Random Forest Regression is a learning strategy in which different decision trees are prepared on distinctive subsets of information and their expectations are found the average value to create a final result. It is viable for regression and classification tasks.

- a. It builds a set of collection of decision trees during the training phase on irregular highlights and subsets of the information. This haphazardness decreases oversampling.
- b. It speaks to the relationship between the independent and dependent variable by averaging the expectations of each tree. In relapse, the last forecast is more often than not the normal of all the expectations in the tree.
- c. It handles large datasets with complex relationships well. It requires minimal feature engineering and is less sensitive to outliers compared to some other regression techniques.

In this study, random forest regression was applied to four different stock datasets from banks (SBI, HDFC, Union, and ICICI), yielding respective RMSE values of 10.895259072330479, 9.483261027264719, 1.979636758241366, and 6.223968736065822.

Results & Discussion

The Root Mean Squared Error (RMSE) of four distinct stocks—including SBIN, HDFCBANK, UNIONBANK, and ICICIBANK—is displayed in Table 1 below.

	SBIN	HDFC	UNION	ICICI
Linear Regression (LR)	6.3938086e-13	3.4196223e-13	3.6756597e-13	2.8226168e-13
Gradient Boosting (GB)	12.5944961	10.0894019	2.1761076	7.0603668
Random Forest Regression (RFR)	10.8952590	9.4832610	1.9796367	6.2239687

Table 1: Results for different algorithms.

Conclusions

In brief, our research on "Financial Analysis for stock pricing using machine learning Algorithms" marks a critical turning point in the field of stock analysis and decision-making. Using different machine learning strategies, such as Linear Regression, Gradient Boosting Method, and Random Forest Regression, we can foresee stock costs. and appraise budgetary execution way better than anticipated with RMSE scores. This accomplishment appears when present day factual models are fitting to explore the complexities of stock markets. In expansion, our investigate highlights the moral suggestions of financial innovation and the require to utilize algorithmic commerce frameworks and prescient models responsibly. Our objective is to minimize potential biases and keep up the integrity of financial decision-making strategies by supporting straightforwardness, reasonableness and accountability. Going forward, our activities will clear the way. more information- and information-driven contributing, driven by mechanical advancement and commitment to speculation principles. Moral principles. Through a more profound understanding of showcase elements and venture behavior, we point to increment money related proficiency and advance maintainable riches. for all financial specialists and organizations.

References

- [1] S. Kumar, R. Gupta, P. Kumar and N. Aggarwal, "A Survey on Artificial Neural Network based Stock Price Prediction Using Various Methods," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 1866-1872.
- [2] Iyer, M., Mehra, R. (2021). Forecasting Price of Indian Stock Market Using Supervised Machine Learning Technique. In: Panigrahi, C.R., Pati, B., Pattanayak, B.K., Amic, S., Li, KC. (eds) Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing, vol 1299. Springer, Singapore.
- [3] Nayak, R.K. et al. (2021). Indian Stock Market Prediction Based on Rough Set and Support Vector Machine Approach. In: Mishra, D., Buyya, R., Mohapatra, P., Patnaik, S. (eds) Intelligent and Cloud Computing. Smart Innovation, Systems and Technologies, vol 153. Springer, Singapore.
- [4] N. Kulshrestha and V. K. Srivastava, "Synthesizing Technical Analysis, Fundamental Analysis & Artificial Intelligence – An Applied Approach to Portfolio Optimisation & Performance Analysis of Stock Prices in India," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2020, pp. 1185-1188.
- [5] D. Kumar, "Stock Forecasting Using Natural Language and Recurrent Network," 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE), Jaipur, India, 2020, pp. 1-5.
- [6] A. Maiti and P. Shetty D, "Indian Stock Market Prediction using Deep Learning," 2020 IEEE REGION 10 CONFERENCE (TENCON), Osaka, Japan, 2020, pp. 1215-1220.