iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68945

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Multi-Factor Influence on NIFTY50: A Machine Learning Approach

Siddhi Mehta[1], Nimesh Vaidya[2], Dr. Vijay Gadhavi[3]
[1]PG Scholar – Faculty of Engineering, Computer Engineering Department Swaminarayan University
[2]Assistant Professor & HOD-Faculty of Engineering, Computer Engineering Department Swaminarayan University
[3]Associate Professor & Dean –Faculty of Engineering (I/C), Computer Engineering Department Swaminarayan University

Abstract: A wide range of economic, technical, and sentiment-driven factors shape the Indian financial market, particularly benchmark indices like the NIFTY50. Using an analytical method that is based on machine learning, this study looks into how all of these factors affect the NIFTY50 index. By employing statistical techniques alongside modern algorithms, we examine a diverse set of macroeconomic variables, technical indicators, and global financial signals to identify the most significant contributors to NIFTY50 fluctuations.[1]

We use correlation analysis, mutual information metrics, and recursive feature elimination (RFE) to evaluate the relationships between variables. These methods enable us to get rid of features that are redundant or less important. Not only are predictive but also feature importance assessments made using machine learning models like Random Forest, XGBoost, and LASSO regression. Model clarity and computational efficiency are also improved by employing dimensionality reduction methods like Principal Component Analysis (PCA).[2]

The results show that the models' predictive accuracy and overall performance are both improved and their complexity is reduced by selecting a refined set of influential features. This research contributes to a deeper understanding of market dynamics and offers practical insights for investors, analysts, and policymakers. In addition, it lays the groundwork for the creation of machine learning-powered intelligent financial decision-making systems that operate in real time.[1]

I. INTRODUCTION

1) Numerous internal and external factors play a complex role in stock indices like the NIFTY50's performance. As one of India's most prominent equity indices, the NIFTY50 reflects the overall health of the Indian economy and serves as a benchmark for investors and analysts. However, its movements are not dictated by a single factor, but rather by a combination of macroeconomic indicators, technical signals, market sentiment, and global financial trends.
2) Traditional financial models often struggle to effectively capture the nonlinear and interdependent nature of these factors. Machine learning techniques, on the other hand, are a powerful alternative that can handle a lot of data, find hidden patterns, and change with the market. Beyond straightforward linear correlations, these instruments offer a chance to gain a deeper understanding of how multiple features simultaneously influence index behavior. Utilizing a method that is based on machine learning, the primary objective of this study is to locate and evaluate the most significant factors that have an effect on the NIFTY50 index. In order to select the most relevant variables from a wide range of economic, technical, and global indicators, the research employs feature correlation and selection techniques like Pearson correlation, mutual information, and recursive feature elimination (RFE). Machine learning models including Random Forest, XGBoost, and LASSO regression are employed not only to predict index movements but also to evaluate the importance of each feature.
3) In addition, principal component analysis (PCA) and other dimensionality reduction methods are incorporated to improve interpretability and lessen model complexity. This study's ultimate objective is to develop a predictive model for the NIFTY50 index that is both more accurate and more effective. It will also provide investors, traders, and policymakers who are looking for data-driven strategies in a market that is becoming more and more dynamic with time.[4]
4) As machine learning (ML) techniques and computer resources have become more widely available, numerous statistical, ML, and deep learning (DL) methods have been deployed in stock market forecasting (Gandhmal and Kumar 2019; Shah et al. 2019).

II. LITERATURE SURVEY

During a literature survey, we collected some of the information about Stock market prediction mechanisms currently being used.

1) Survey of Stock Market Prediction Using Machine Learning Approach

The stock market prediction has become an increasingly important issue in the present time. One of the methods employed is technical analysis, but such methods do not always yield accurate results. So it is important to develop methods for a more accurate prediction. Generally, investments are made using predictions that are obtained from the stock price after considering all the factors that might affect it. The technique that was employed in this instance was a regression. Since financial stock marks generate enormous amounts of data at any given time a great volume of data needs to undergo analysis before a prediction can be made. Each of the techniques listed under regression hasits own advantages and limitations over its other counterparts. One of the noteworthy techniques that were mentioned was linear regression. The way linear regression models work is that they are often fitted using the least squares approach, but they may alternatively be also be fitted in other ways, such as by diminishing the "lack of fit" in some other norm, or by diminishing a handicapped version of the least squares loss function. Conversely, the least squares approach can be utilized to fit nonlinear models. [7]

2) Impact of Financial Ratios and Technical Analysis on Stock Price Prediction Using Random Forests

The use of machine learning and artificial intelligence techniques to predict the prices of the stock is an increasing trend. More and more researchers invest their time every day in coming up with ways to arrive at techniques that can further improve the accuracy of the stock prediction model. Due to the vast number of options available, there can be n number of ways on how to predict the price of the stock, but all methods don’t work the same way. The output varies for each technique even if the same data set is being applied. In the cited paper the stock price prediction has been carried out by using the random forest algorithm is being used to predict the price of the stock using financial ratios form the previous quarter. This is just one way of looking at the problem by approaching it using a predictive model, using the random forest to predict the future price of the stock from historical data. However, there are always other factors that influence the price of the stock, such as sentiments of the investor, public opinion about the company, news from various outlets, and even events that cause the entire stock market to fluctuate. By using the financial ratio along with a model that can effectively analyze sentiments the accuracy of the stock price prediction model can be increased. [8]

3) Stock Market Prediction via Multi-Source Multiple Instance Learning

Accurately predicting the stock market is a challenging task, but the modern web has proved to be a very useful tool in making this task easier. Due to the interconnected format of data, it is easy to extract certain sentiments thus making it easier to establish relationships between various variable and roughly scope out a pattern of investment. Investment pattern from various firms show sign of similarity, and the key to successfully predicting the stock market is to exploit these same consistencies between the data sets. The way stock market information can be predicted successfully is by using more than just technical historical data, and using other methods like the use of sentiment analyzer to derive an important connection between people’s emotions and how they are influenced by investment in specific stocks. One more important segment of the prediction process was the extraction of important events from web news to see how it affected stock prices. [9]

A. Block Diagram Of Ai Prediction Model

A block diagram for a Nifty 50 stock market predictor using AI and deep learning can help visualize the flow of data and processes involved in the prediction model. is a refined block diagram for the Nifty 50 predictor using AI and deep learning, along with explanations for each block.[5]

This flowchart represents the **workflow of a machine learning model for stock price prediction**, specifically applied to financial data such as that used in predicting the NIFTY50 index. Here's a step-by-step explanation of the components in the diagram:

- Obtain Financial Information Raw financial data are gathered at this initial step. It could include: Historical stock prices
- Technical indicators (e.g., moving averages, RSI)
- Indicators of the larger economy, such as interest rates and inflation, Global market indices (e.g., S&P 500, crude oil prices)
- Sentiment data (e.g., news headlines, social media sentiment)

B. Extraction of Features In this stage, relevant features are extracted from the raw data. This includes:

- Cleaning and preprocessing calculating brand-new features like volatility and percentage changes, among others. Normalization or standardization of values

- A selection of variables that have the potential to predict how prices will move in the future The extracted features are then used to prepare both the training and testing datasets.

C. Training Data

The machine learning model is trained with a portion of the data, typically historical data. The model can learn about patterns, trends, and relationships between features and the stock price thanks to this dataset.

D. Testing Data

Another subset of the data, typically the most recent ones, is not presented to the model during training. It is put to the test to see how well the model predicts unknown outcomes. 5. Result of Trained Data A result (the trained model) is produced by the model after it has been trained with the training data. This model encapsulates the learned relationships between the features and the target (stock price or trend).

E. Predicted Stock Price

Using the trained model and the testing data, the final output is a predicted stock price or index value. This prediction can be used for decision-making or further analysis.

Loop of Feedback (Implicit) The procedure is iterative, despite the fact that labeled arrows do not explicitly show this. The model can be retrained with new data as it becomes available, and predictions can be updated accordingly.

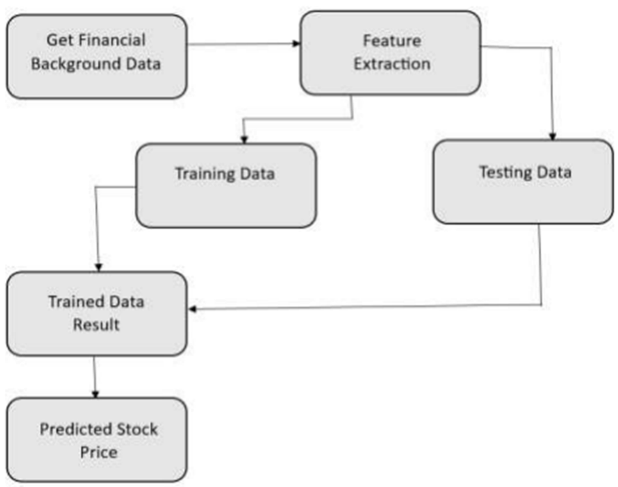

Fig 1: Block Diagram of AI Prediction Model

F. Future Research Agenda

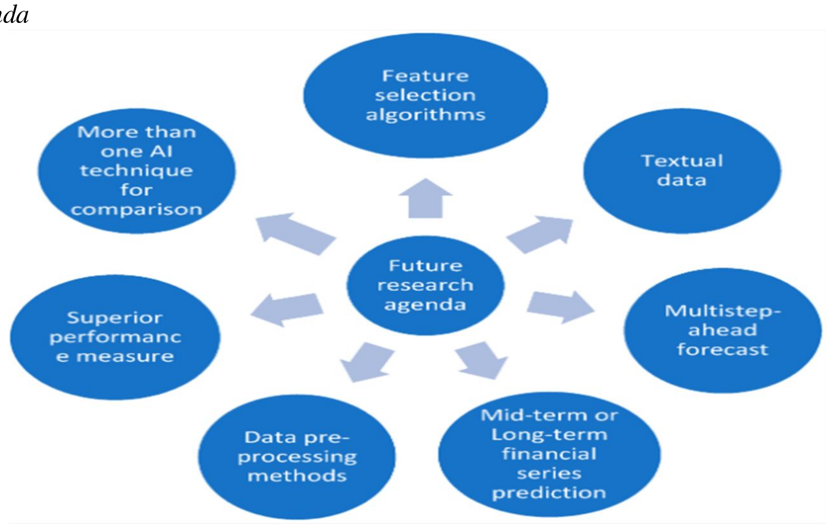

Fig 2. Future research agenda.

This diagram represents a **"Future Research Agenda"** in the field of **financial market prediction using AI and machine learning**. It outlines the key areas where further exploration, innovation, or improvement is needed to enhance the performance, reliability, and practical application of predictive models—especially in contexts like stock price forecasting (e.g., NIFTY50).

The future of financial market forecasting using artificial intelligence (AI) and machine learning (ML) offers a wide range of promising research opportunities. The diagram shows that, in order to improve the accuracy, efficacy, and applicability of predictive models, particularly for stock indices like NIFTY50, several key areas require more in-depth research and development. One of the primary directions is the enhancement of feature selection algorithms. In order to reduce noise, avoid overfitting, and enhance model performance, selecting the features that are most relevant and informative becomes more and more important as the volume of financial data grows. For financial contexts, embedded methods in ensemble models, mutual information analysis, and recursive feature elimination (RFE) can be improved. Incorporating textual data is another crucial aspect. Financial news articles, analyst reports, and social media platforms are rich sources of market sentiment. Natural Language Processing (NLP) can be employed to extract sentiment scores, event indicators, or volatility signals from unstructured data, which can be integrated with structured numerical data for improved forecasting.

Moreover, most existing models focus on single-step prediction, which limits their usefulness for long-term investment strategies. As a result, there is a pressing need to develop models that are capable of multistep-ahead forecasting, in which the system can anticipate market trends for a number of different time periods in the future. This would benefit traders and institutions looking for broader planning horizons.

The focus on long-term and mid-term financial series prediction is related to this. While short-term forecasting is common, models should also try to predict price behavior over longer periods (weeks, months, or even years), taking into account market structural shifts and macroeconomic cycles. The role of data preprocessing methods cannot be overstated. Financial datasets often suffer from missing values, outliers, and non-stationarity. Advanced preprocessing methods—such as time-series normalization, interpolation techniques, and decomposition (e.g., STL, wavelets)—can enhance model stability and accuracy.

Another area of improvement is the development of superior performance measures. It is possible that conventional metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) do not accurately reflect the financial utility of a model. It is necessary to incorporate domain-specific metrics that are more in line with investor interests, such as the Sharpe ratio, maximum drawdown, or directional accuracy.

Future research should also compare multiple AI techniques to validate model reliability. No single model performs best under all market conditions. Therefore, using a mix of algorithms—such as LSTM, XGBoost, ARIMA, and hybrid models—can provide comparative insights and help in building more robust ensemble frameworks.

These areas of study make up a comprehensive plan for future research with the goal of improving AI-driven financial forecasting. The interpretability, dependability, and practical value of machine learning systems in real-world financial decision-making will also rise as a result of addressing these directions.

III. CONCLUSION

The multi-factor influence on the NIFTY50 index is examined using a machine learning-based framework that incorporates macroeconomic indicators, technical parameters, and global financial signals. The research successfully identifies key variables that have a significant impact on the movements of the index by employing feature correlation and selection methods like mutual information, Pearson correlation, and recursive feature elimination (RFE).

Prediction accuracy and feature importance were evaluated using cutting-edge machine learning techniques like Random Forest, XGBoost, and LASSO regression.

Additionally, principal component analysis (PCA) and other dimensionality reduction methods helped to keep the most important information while simultaneously reducing model complexity. Prediction accuracy and model efficiency are both enhanced by concentrating on a more refined set of influential features, as demonstrated by the experiments. In addition to providing a data-driven foundation for the creation of forecasting systems that are more accurate and intelligent, this research contributes to a deeper comprehension of the factors that influence NIFTY50 fluctuations. The findings can aid investors, analysts, and policymakers in making informed decisions in volatile market conditions.

The incorporation of textual sentiment data, the investigation of deep learning architectures, and the expansion of prediction to multi-step and long-term horizons are all promising future research topics in this field.

REFERENCES

[1] Stock Market Prediction of NIFTY 50 Index Applying Machine Learning Techniques Zahra Fathali ,Zahra Kodia & Lamjed Ben Said Article: 2111134 | Published online: 16 Sep 2022

[2] A multi-feature selection fused with investor sentiment for stock price prediction Author links open overlay panelKehan Zhen [a b], Dan Xie [a b c], Xiaochun Hu

[3] Forecasting of NIFTY 50 Index Price by Using Backward Elimination with an LSTM Model by Syed Hasan Jafar [1]Shakeb Akhtar [1],Hani El-Chaarani [2,*],Parvez Alam Khan [3] andRuaa Binsaddig

[4] Survey of feature selection and extraction techniques for stock market predictionHtet Htet Htun, Michael Biehl & Nicolai Petkov volume 9, Article number: 26 (2023)

[5] NIFTY 50 PREDICTIONUSING AI AND MACHINE LEARNING: A REVIEW Ms. Wayal Priti*1, Ms. Wani Aasavari*2, Ms. Mulik Siddhi*3, Dr. S.B. Lande*4 *1,2,3,4Department Of Electronics And Telecommunication VPKBIET Baramati Savitribai Phule Pune University, Maharashtra, India

[6] Stock Market Prediction Using Machine Learning Algorithms K. Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal

[7] Ashish Sharma, Dinesh Bhuriya, Upendra Singh. "Survey of Stock Market Prediction Using Machine Learning Approach", ICECA 2017

[8] Loke.K.S. “Impact Of Financial Ratios And Technical Analysis On Stock Price Prediction Using Random Forests”, IEEE, 201

[9] Xi Zhang1, Siyu Qu1, Jieyun Huang1, Binxing Fang1, Philip Yu2, “Stock Market Prediction via Multi-Source Multiple Instance Learning.” IEEE 2018

[10] Application of Artificial Intelligence in Stock Market Forecasting: A Critique, Review, and Research Agenda by Ritika Chopra andGagan Deep Sharma University School of Management Studies, Guru Gobind Singh Indraprastha University, Dwarka Sector 16-C, New Delhi 110078, India