

Stock Price Prediction Using Machine Learning: Research Paper Directions

Shivam Kumar Khirwal

School of Computer Application and Technology, Galgotias University, Greater Noida,

Uttar Pradesh, India

ABSTRACT

Stock market investment strategies are complex and rely on an evaluation of vast amounts of data. In recent years, machine learning techniques have increasingly been examined to assess whether they can improve market forecasting when compared with traditional approaches. The objective for this study is to identify directions for future machine learning stock market prediction research based upon a review of current literature. A systematic literature review methodology is used to identify relevant peerreviewed journal articles from the past twenty years and categorize studies that have similar methods and contexts. Four categories emerge: artificial neural network studies, support vector machine studies, studies using genetic algorithms combined with other techniques, and studies using hybrid or other artificial intelligence approaches. Studies in each category are reviewed to identify common findings, unique findings, limitations, and areas that need further investigation. The final section provides overall conclusions and directions for future research.

Keywords: machine learning, stock market prediction, literature review, research taxonomy, artificial neural network, support vector machine, genetic algorithm, investment decision

INTRODUCTION

The objective for this study is to identify directions for future machine learning (ML) stock market prediction research based upon a review of current literature. A systematic literature review methodology is used to identify relevant peer-reviewed journal articles from the past twenty years, evaluate and categorize studies that have similar methods and contexts, and then compare the studies in each category to identify common findings, unique findings, limitations, and areas that need further investigation. This will provide artificial intelligence and finance researchers with directions for future research into the use of ML techniques to predict stock market index values and trends.

METHOD FOR IDENTIFYING RELEVANT STUDIES

Each researcher involved in this study conducted an independent search for peer reviewed journal articles where some form of machine learning was used to predict a stock market related outcome.

Articles were found using Google Scholar, EBSCO, and EconLit. To identify findings that are relevant to today's IT environment, only studies from the past years were included in the final list. Each study used one, or more, machine learning techniques to predict stock market index values or expectations for whether the future index value will rise or fall.

MACHINE LEARNING STOCK MARKET PREDICTION

SCREENSHOTS

Stock Price Prediction

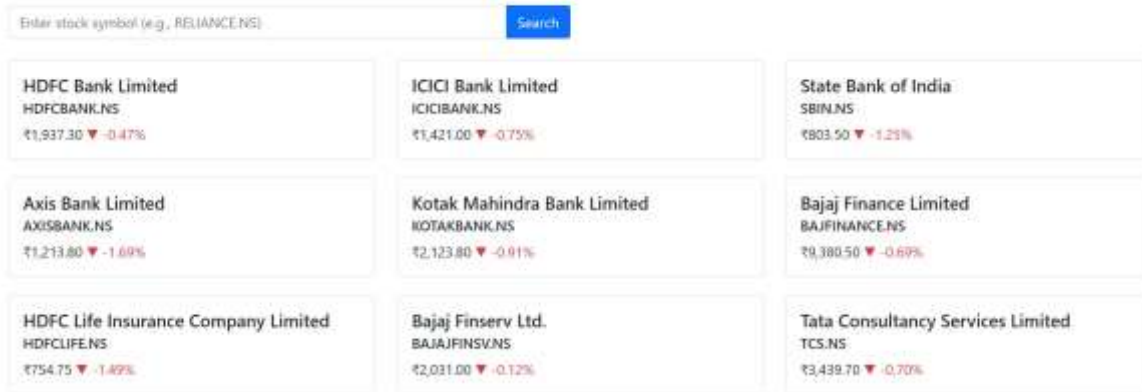


fig: front page of a website

A complete list of reviewed studies is provided in the Appendix. A brief description of each machine learning approach is also provided prior to describing the related studies.

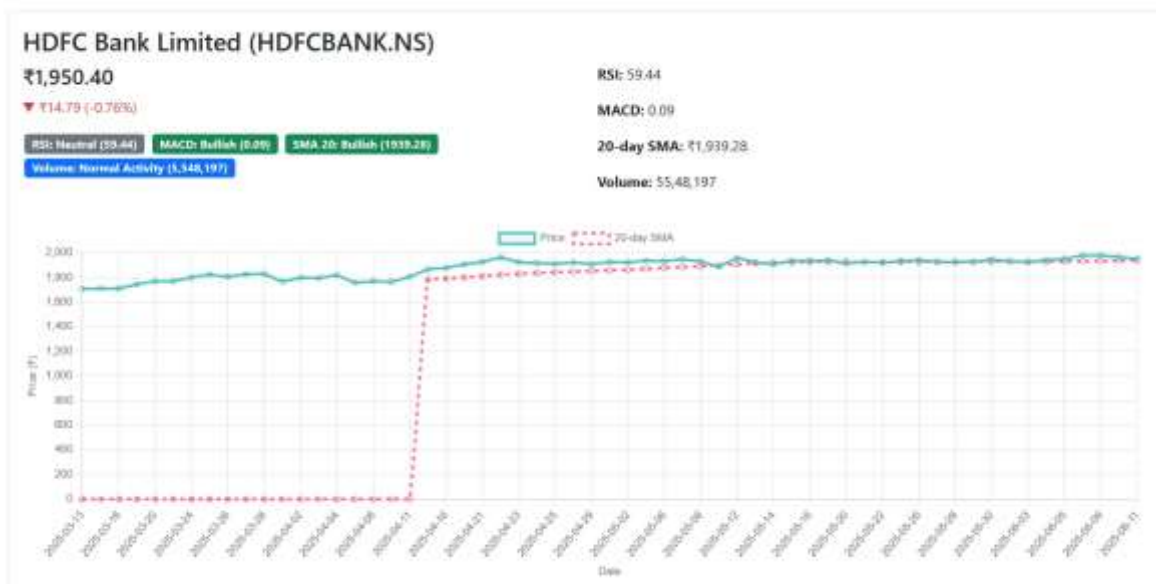


fig: Overview of a particular stock



fig: Shows the predicted price of a particular stock

Above shows the live price of a particular stock with a predicted price and the confidence with predicted price and graph based on the graph and technical indicators the graph is being predicted.

Thus, the above screenshots show the working of the project on how it works

Literature Review

Several studies have applied machine learning techniques for stock price prediction. Linear models like ARIMA have been used in the past but are limited in handling non-linearity. Recent work has shifted towards non-linear models like Random Forests, Support Vector Machines, and LSTM neural networks. Research indicates that deep learning models can capture temporal dependencies better than traditional models.

Methodology

Data Collection

We collected historical stock data using the Yahoo Finance API, including Open, High, Low, Close, and Volume prices.

Data Preprocessing

Handling missing values

Normalization/Scaling

Feature engineering (e.g., Moving Average, RSI)

Models Used

1. Linear Regression

A baseline model that assumes a linear relationship between the features and the stock price.

2. *Random Forest Regressor*

An ensemble learning method based on decision trees that helps reduce overfitting and improve accuracy.

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

A special kind of RNN well-suited for time series data. It remembers previous states and patterns to predict future prices.

Tools and Technologies

Language: Python

Libraries: Pandas, NumPy, Scikit-learn, Keras, TensorFlow, Matplotlib

IDE: Jupyter Notebook / VS Code

Conclusion

This study demonstrates that machine learning, especially LSTM models, can be highly effective in predicting stock prices based on historical data. While no model can guarantee 100% accuracy due to market unpredictability, using ML can give investors a strategic advantage.

The final set of conclusions was also apparent after reflection. Financial investment theory needs to be a stronger driver underlying the ML systems' inputs, algorithms, and performance measures. If this is not the case then results may just be random and not have any practical use. Too many studies use techniques without consideration of the vast amount of financial theory that has been developed over the past centuries. Reporting failures where techniques do not improve predictive performance would also be informative. At this point this rarely occurs so it is impossible to find patterns where there is a mismatch between a particular stock market prediction problem and a machine learning technique. Finally, the irony in this research area is that it is a zero-sum game for investors. If the best machine learning stock market prediction technique is found, and all investors adopt this system, the result is that no one is better off. Large investment firms researching the best machine learning methods have no incentive to share this information with others.

Future Scope

Integration of real-time news sentiment analysis

Incorporation of macroeconomic indicators

Deployment as a live web or mobile application for investors

References

Brownlee, J. (2017). Deep Learning for Time Series Forecasting.

Zhang, G., Eddy Patuwo, B., & Hu, M. Y. (1998). Forecasting with artificial neural networks.

Yahoo Finance API Documentation.