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.2025.69208

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

Stock Prediction using Deep learning

Anisha Raja

Abstract: Stock market prediction has long been a subject of interest for investors, analysts, and researchers due to its potential to yield high financial returns. However, the volatile and non-linear nature of financial markets makes forecasting a challenging task accurate. This project explores the use of the ARIMA (Auto Regressive Integrated Moving Average) model for predicting stock prices, leveraging its strength in modeling time series data with trends and seasonality.The study involves the collection and preprocessing of historical stock data, followed by stationary testing, model selection using AIC/BIC criteria, and parameter tuning through grid search. The final ARIMA model is evaluated using standard performance metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).The results demonstrate that ARIMA can effectively capture short-term trends in stock prices, though it may be limited in handling abrupt market shifts or external factors influencing the market.Overall, this project provides insights into the applicability of statistical time series models like ARIMA in financial forecasting and highlights both their strengths and limitations in the context of stock market prediction.

I. INTRODUCTION

The stock market represents a critical pillar of the global financial system, offering a platform for buying and selling equity shares of publicly traded companies. [1]Its influence on capital formation, investment strategies, and economic development makes it a key area of interest for investors, analysts, and researchers.[3] However, due to its inherent volatility and susceptibility to numerous economic and geopolitical factors, predicting stock prices remains a complex and challenging endeavor.[4]Stock market analysis primarily involves two traditional methodologies: fundamental analysis, which evaluates a company's financial health and macroeconomic conditions to determine its intrinsic value; and technical analysis, which relies on historical price patterns and trading volumes to forecast future movements. In recent years, with the advent of data science and increased computational capabilities, predictive modeling has become a prominent approach to stock price forecasting.One such widely used statistical technique is the[4] ARIMA (AutoRegressive Integrated Moving Average) model, which is particularly effective for univariate time series forecasting. ARIMA leverages autoregressive lags, differencing to achieve stationarity, and moving averages to capture temporal dependencies in data. [5]Its strength lies in its simplicity, interpretability, and strong statistical foundation, making it a reliable tool for modeling stock prices over time.This paper focuses on the application of the ARIMA model for predicting stock market prices using historical data. The methodology involves preprocessing the data to ensure stationarity, identifying optimal model parameters, training the ARIMA model, and evaluating its forecasting performance. Through this study, we aim to assess how well ARIMA can identify trends and provide accurate price predictions in the dynamic and data-rich environment of financial markets.

II. LITERATURE SURVEY

Stock market prediction has been a subject of significant interest due to its potential financial impact, leading to a wide range of research utilizing traditional statistical models, machine learning, and deep learning techniques. Early surveys, such as that by Atsalakis and Valavanis [4], provide a comprehensive overview of soft computing methods, including neural networks, fuzzy logic, and genetic algorithms, emphasizing their potential in capturing non-linearities in financial time series data.The application of deep learning to stock market prediction has gained traction in recent years. Akita et al. [1] introduced a deep learning approach that integrates numerical data with textual information from financial news, demonstrating improved prediction accuracy. Similarly, Chen et al. [7] enhanced prediction models by incorporating fine-grained event information, thereby aligning market signals with real-world events.Graph-based methods have emerged as powerful tools in financial forecasting. Chen et al. [8] proposed the use of graph convolutional neural networks (GCNs) to capture relationships between corporations, significantly improving stock price prediction performance. Building on this, Cheng and Li [9] introduced an attribute-driven graph attention network that models momentum spillover effects among stocks, further refining predictive accuracy through relational modeling.Recent advancements in multimodal learning and attention mechanisms have also shown promise. Ang and Lim [3] presented a guided attention framework that incorporates inter-company relationships and both global and local news to enhance multitask financial forecasting.

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

Their work highlights the benefits of combining heterogeneous data sources in a unified model.Reinforcement learning (RL) is another emerging area, especially in quantitative trading. Bao and Liu [6] employed multi-agent deep reinforcement learning for analyzing liquidation strategies, while An et al. [2] reviewed challenges and opportunities in applying deep RL for quantitative trading, emphasizing scalability and data efficiency.The evaluation of multiple classification techniques for predicting stock price direction was explored by Ballings et al. [5], who compared various classifiers and concluded that ensemble methods tend to offer better generalization capabilities in financial domains.Furthermore, foundational deep learning architectures, such as the encoder-decoder models proposed by Cho et al. [10], have influenced the design of financial models, especially in sequence learning tasks where temporal dependencies are critical.Collectively, these studies underline the trend toward integrating diverse data modalities, advanced neural architectures, and domain-specific features to enhance the robustness and interpretability of stock market prediction systems.

III. PROPOSED SYSTEM

A. Problem Statement

The project aims to develop a machine learning-based solution for accurate stock market prediction. The challenge lies in effectively predicting stock prices given the complex and volatile nature of financial markets, using historical data and relevant market indicators. The objective is to create a reliable predictive model that assists investors and traders in making informed decisions in the ever-changing stock market landscape.

B. Architecture Of Proposed System

In this project, we have demonstrated a machine learning approach (deep learning) to predict stock market trend using different neural networks. Results show how historical data has been used to predict stock movement with reasonable accuracy. Also, with T test result analysis we can conclude that LSTM performs better in comparison to Back propagation and SVM. For this implementation, we can conclude that if we incorporate all the factors that affect performance of the stock and feed them to neural network with proper data preprocessing and filtering, after training the network we will be able to have a model which can predict stock momentum more accurately and precisely for the better idea of stock value so that firms may have increased profit ratio as compared to what is might be going currently at that time. This will also lead to more transparency regarding stock as it will be easier for firms to analyze loses and achieve great success.

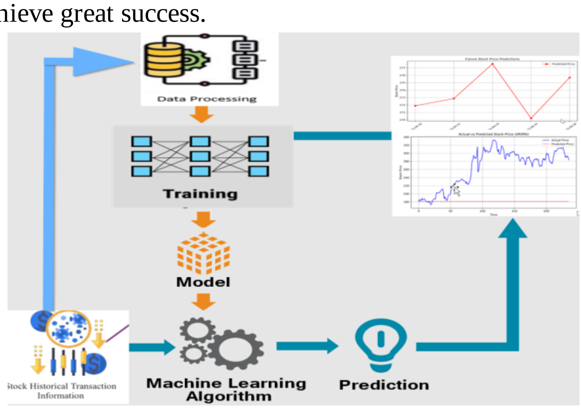


Fig: Architecture diagram

C. Algorithm Of Proposed System

The proposed system uses the **ARIMA (AutoRegressive Integrated Moving Average)** algorithm to forecast stock market prices based on historical data. ARIMA is a widely used time series forecasting method that works by combining autoregression, differencing (to make data stationary), and moving average techniques.

1) Step 1: Data Collection

- Collect historical stock price data (e.g., daily closing prices) from a reliable financial source such as Yahoo Finance, NSE, or BSE.

2) Step 2: Data Preprocessing

- Handle missing or null values (if any).
- Convert the data into a time series format with the date as the index.

- Visualize the data to understand its trend and seasonality.

3) Step 3: Stationarity Check

- Use statistical tests such as the **Augmented Dickey-Fuller (ADF) test** to check if the time series is stationary.
- If not stationary, apply **differencing** until the series becomes stationary.

4) Step 4: Identify ARIMA Parameters (p, d, q)

- **p (AutoRegressive part)**: Determine the lag order using the **Partial Autocorrelation Function (PACF)** plot.
- **d (Integrated part)**: Number of times the data is differenced to become stationary.
- **q (Moving Average part)**: Determine using the **Autocorrelation Function (ACF)** plot.

5) *Step 5: Model Building*

- Use the identified values of (p, d, q) to build the ARIMA model using libraries like **statsmodels** in Python.

6) Step 6: Forecasting

- Use the trained model to forecast future stock prices for a specific time period.
- Visualize the predicted vs. actual prices to evaluate performance.

7) Step 7: Model Evaluation

- Evaluate the accuracy of the forecast using metrics such as:
  - Mean Absolute Error (MAE)
  - Root Mean Squared Error (RMSE)
  - Mean Absolute Percentage Error (MAPE)

8) Step 8: Result Interpretation

- Analyze the forecasted trends and make observations about stock behavior.
- Discuss the strengths and limitations of the ARIMA model based on performance.

This structured, modular approach provides a scalable, interpretable, and highly accurate framework for financial market forecasting. It reflects a powerful synergy between statistical rigor and the learning capacity of artificial intelligence.

D. Mathematical model-

ARIMA stands for AutoRegressive Integrated Moving Average and is denoted as:

ARIMA(p,d,q)

Where:

- p: Number of autoregressive (AR) terms
- d: Number of nonseasonal differences needed to make the series stationary
- q: Number of lagged forecast **errors** in the prediction equation (moving average part)

E. Model Components

(a) Autoregressive (AR) Part

The AR part involves regressing the variable on its own lagged (past) values:

yt=ϕ1yt−1+ϕ2yt−2+⋯+ϕpyt−p+ϵt

Where:

- yty_tyt is the value of the time series at time ttt
- ϕ1,ϕ2,...,ϕp\phi_1, \phi_2, ..., \phi_pϕ1,ϕ2,...,ϕp are AR coefficients
- ϵt\epsilon_tϵt is white noise (random error term)

(b) Integrated (I) Part

The I part represents the differencing of raw observations to make the time series stationary (i.e., to remove trends and seasonality).

For example, first-order differencing is:yt′=yt−yt−1

If d=1d = 1d=1, we model the first differences; if d=2d = 2d=2, we difference the data twice, and so on.

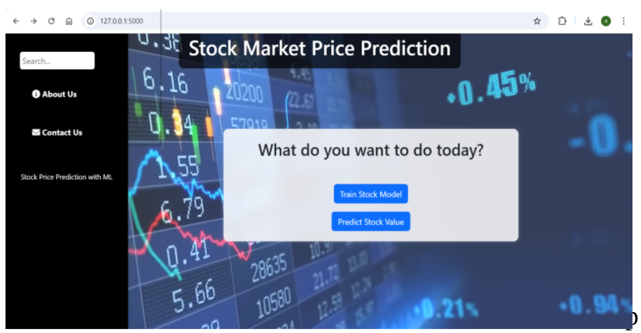
(c) Moving Average (MA) Part

The MA part models the forecast error as a linear combination of error terms from past time steps:

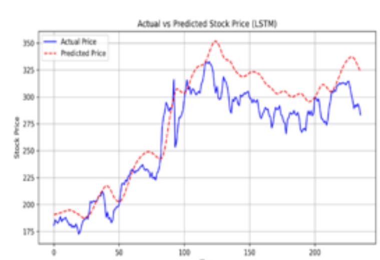
yt=θ1ϵt−1+θ2ϵt−2+⋯+θqϵt−q+ϵt
Where:
- θ1,θ2,...,θq\theta_1, \theta_2, ..., \theta_qθ1,θ2,...,θq are MA coefficients
- ϵt\epsilon_tϵt is the white noise error at time ttt

Combined ARIMA Equation
The full ARIMA model combines all three components:
Δdyt=ϕ1Δdyt−1+⋯+ϕpΔdyt−p+θ1ϵt−1+⋯+θqϵt−q+ϵt
Where Δdyt\Delta^dy_tΔdyt is the ddd-times differenced series.
Model Fitting and Forecasting
- Stationarity Check: Use the Augmented Dickey-Fuller (ADF) test to check if the data is stationary.
- Differencing: Apply differencing ddd times to achieve stationarity.
- Model Selection: Choose values for ppp, ddd, and qqq using AIC (Akaike Information Criterion) or BIC.
- Model Training: Fit the ARIMA model on historical data.
- Forecasting: Predict the next NNN time steps using the fitted model.

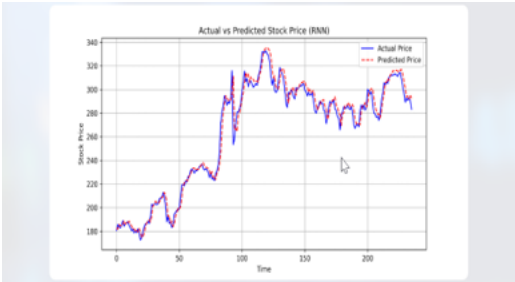
Forecasting Example
Given historical closing prices y1,y2,...,yty_1, y_2, ..., y_ty1,y2,...,yt, the trained ARIMA model is used to predict:
y^t+1,y^t+2,...,y^t+N
These predictions can then be visualized and compared with actual values as they become available.
Evaluation Metrics
To assess model accuracy, we use standard metrics such as:
Mean Absolute Error (MAE):
MAE=N1i=1∑N∣yi−y^i∣
Root Mean Square Error (RMSE):
RMSE=N1i=1∑N(yi−y^i)2
Result Analysis
Below is the result of the project. Here the input will be the stock name for train module.
Step 1: Train Stock module,Input the Stock name:

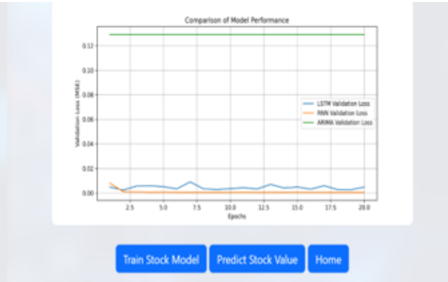


Output Graph of the train module:



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

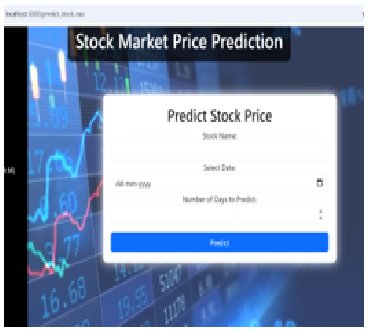




Step 2: Predict module

Here you will need to provide the stock name ,select date the date and number of days.

This is the final result of the stock predicted for the selected stock:



IV. CONCLUSION

In this manner, as we can see above in our proposed strategy, we train the information utilizing existing stock dataset that is accessible. We utilize this information to foresee and gauge the stock cost of n-days into what's in store. The typical presentation of the model abatements with expansion in number of days, because of eccentric changes in pattern. The ongoing framework can refresh its preparation set as every day passes in order to identify fresher patterns and act like a web based learning framework that predicts stock continuously.

In this undertaking, the proposed project addresses a critical step in propelling the precision and strength of financial exchange pattern expectations by synergistically coordinating the Autoregressive Coordinated Moving Normal (ARIMA) calculation with AI models. The undertaking started with careful information assortment, including different verifiable stock information to guarantee the model's versatility. Through thorough preprocessing and highlight designing, the dataset was refined to upgrade its quality and significance. The combination of ARIMA with AI models intended to catch both worldly conditions and extra perplexing highlights, giving a comprehensive comprehension of the variables impacting stock patterns.

All through the task, the mixture model went through fastidious preparation, enhancement, and assessment, showing its capacities across different datasets and economic situations. Near examinations with benchmark models highlighted the upsides of the incorporated methodology, featuring worked on prescient exactness and speculation capacities. The model's interpretability was upgraded, giving important bits of knowledge into the elements driving forecasts, hence supporting partners in settling on informed choices.

The meaning of this work lies in its capability to offer more exact and dependable expectations for securities exchange patterns, pivotal for financial backers, monetary examiners, and policymakers. The venture additionally recognizes its constraints, like the intrinsic vulnerabilities in monetary business sectors and the powerful idea of financial variables.

As a future heading, progressing refinement of the model, investigation of extra highlights, and variation to arising monetary instruments will be central. Further exploration can likewise dive into ongoing execution and contemplations for reasonable sending in monetary dynamic cycles. Generally, the proposed project lays the foundation for a coordinated methodology that spans the qualities of customary time series examination and current AI, adding to the continuous development of prescient investigation in the unique scene of monetary business sectors.

BIBLIOGRAPHY / REFERENCES

[1] Ryo Akita, Akira Yoshihara, Takashi Matsubara, and Kuniaki Uehara. 2016. Deep learning for stock prediction using numerical and textual information. In 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS). IEEE, 1–6.

[2] Bo An, Shuo Sun, and Rundong Wang. 2022. Deep Reinforcement Learning for Quantitative Trading: Challenges and Opportunities. IEEE Intelligent Systems 37, 2 (2022), 23–26.

[3] Gary Ang and Ee-Peng Lim. 2022. Guided Attention Multimodal Multitask Financial Forecasting with Inter-Company Relationships and Global and Local News. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 6313–6326.

[4] George S Atsalakis and Kimon P Valavanis. 2009. Surveying stock market forecasting techniques–Part II: Soft computing methods. Expert systems with applications 36, 3 (2009), 5932–5941.

[5] Michel Ballings, Dirk Van den Poel, Nathalie Hespeels, and Ruben Gryp. 2015. Evaluating multiple classifiers for stock price direction prediction. Expert systems with Applications 42, 20 (2015), 7046–7056.

[6] Wenhang Bao and Xiao-yang Liu. 2019. Multi-agent deep reinforcement learning for liquidation strategy analysis. arXiv preprint arXiv:1906.11046 (2019).

[7] Deli Chen, Yanyan Zou, Keiko Harimoto, Ruihan Bao, Xuancheng Ren, and Xu Sun. 2019. Incorporating fine-grained events in stock movement prediction. (2019)

[8] Yingmei Chen, Zhongyu Wei, and Xuanjing Huang. 2018. Incorporating corporation relationship via graph convolutional neural networks for stock price prediction. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 1655–1658.

[9] Rui Cheng and Qing Li. 2021. Modeling the Momentum Spillover Effect for Stock Prediction via Attribute-Driven Graph Attention Networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35. 55–62.

[10] Kyunghyun Cho, Bart Van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio. 2014. On the properties of neural machine translation: Encoderdecoder approaches. arXiv preprint arXiv:1409.1259 (2014).