



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 12, December 2025

Sector-Wise Evaluation of ARIMA Forecasting Accuracy for Nifty 50 Individual Stocks

Madana E T, Ms Akhila N

P.G. Student, Department of Master of Computer Application, PES Institute of Technology and Management,
Shivamogga, Karnataka, India

Assistant Professor, Department of Master of Computer Application, PES Institute of Technology and Management,
Shivamogga, Karnataka, India

ABSTRACT: Accurate stock price forecasting plays a crucial role in supporting data-driven financial decision-making, particularly within highly dynamic markets such as the Nifty 50 index. Traditional manual analysis of stock trends is time-consuming, subjective, and often unable to capture complex temporal patterns. This paper proposes a sector-wise forecasting framework using the Autoregressive Integrated Moving Average (ARIMA) model to evaluate prediction accuracy for individual Nifty 50 stocks. Historical multi-year price data undergo preprocessing and stationarity validation using the Augmented Dickey-Fuller test before model fitting. Autocorrelation and partial autocorrelation analyses guide ARIMA parameter selection for each stock, followed by short-term forecast generation across multiple time horizons. Model performance is quantitatively assessed using RMSE, MAE, and MAPE metrics, enabling a comparative analysis of forecasting accuracy across sectors such as IT, Banking, Automobile, etc. The results indicate that ARIMA provides significantly higher precision for short-term (20-day) predictions, with certain sectors exhibiting more stable and predictable patterns than others. The proposed sector-wise evaluation framework offers a systematic, data-driven approach for understanding the reliability of ARIMA forecasting across diverse industries within the Nifty 50, supporting investors and analysts in making informed and sector-specific financial decisions.

KEYWORDS: ARIMA model, Nifty 50, stock price forecasting, sector-wise analysis, time series prediction, RMSE, MAPE, financial modelling.

I. INTRODUCTION

Stock market forecasting has become increasingly important in today's fast-evolving financial landscape, where rapid economic shifts, global events, and sector-specific dynamics significantly influence price movements. As investors seek reliable methods to understand market behaviour, predicting stock trends within major indices such as the Nifty 50 has gained substantial relevance. The Nifty 50, comprising companies from diverse industries including Banking, IT, FMCG, Metals, Automobiles, and Energy, reflects the broader performance of the Indian economy. However, the volatility inherent in financial markets makes manual interpretation of stock patterns highly challenging. Traditional analysis often depends on expert judgment, subjective interpretation of charts, and time-consuming manual assessments, which may fail to capture hidden temporal dependencies or sudden fluctuations.

these, the Autoregressive Integrated Moving Average (ARIMA) model remains one of the most widely used statistical techniques for capturing patterns such as trends, seasonality, and autocorrelation in stock price movements. By analysing multi-year price series for individual Nifty 50 stocks, ARIMA models can generate short-term predictions that help investors, analysts, and financial institutions anticipate market behaviour more accurately.

II. LITERATURE SURVEY

The use of time-series models for stock market forecasting was first explored in early financial prediction studies, where researchers applied statistical methods to analyse historical price movements and identify predictable patterns [1]. These early models relied primarily on linear trend analysis and basic autoregression to understand market behaviour over short horizons. As market datasets became more detailed, the Autoregressive Integrated Moving Average (ARIMA) model emerged as a widely adopted framework for modelling stock price dynamics due to its ability to

capture autocorrelation, noise components and differenced trends [2]. Later studies introduced systematic ARIMA-based forecasting pipelines, where the decision process was separated into key stages such as data preprocessing, stationarity analysis and parameter tuning using ACF and PACF plots [3]. This modular approach allowed researchers to treat trend removal, lag dependency evaluation and model fitting as independent tasks before combining them into a unified prediction system. Research in [4] further demonstrated how ARIMA can be applied across multiple stock indices, with results showing stronger accuracy for short-term predictions compared to long-term forecasts. Sector-specific forecasting was examined in several works, where stocks were grouped by industry classification to study variations in volatility and model performance [5]. These studies highlighted that stable sectors such as FMCG or Telecom often produce lower forecasting errors, while high-volatility sectors like Metals or Technology show reduced ARIMA accuracy. Additional works incorporated comparative analysis between ARIMA and machine-learning models, revealing that while ANN and LSTM architectures capture non-linear behaviour more effectively, ARIMA remains highly competitive for linear, short-term forecasting tasks [6].

III. PROPOSED METHODOLOGY

The proposed sector-wise ARIMA forecasting system is divided into four major stages: data collection and preprocessing, stationarity testing and ARIMA parameter identification, model training and multi-horizon forecasting, and sector-wise accuracy evaluation using statistical metrics.

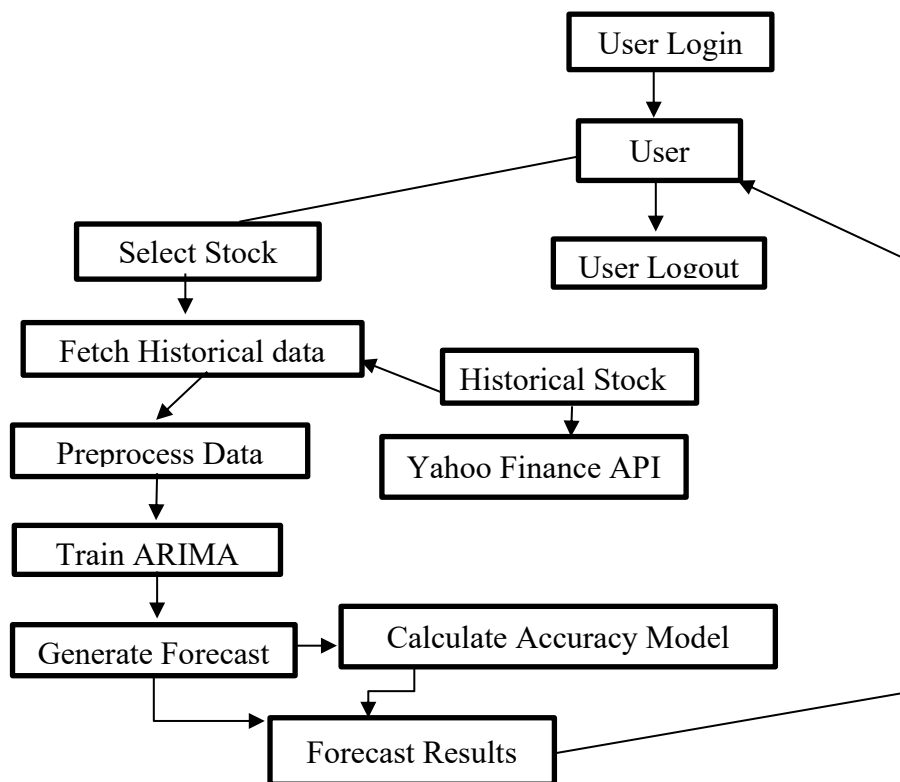


Fig. 1 Data Flow Diagram

The flow diagram illustrates how the ARIMA-based Sector-Wise Stock Forecasting System operates. The process begins when the user—typically an analyst, researcher, or investor—selects the desired Nifty 50 stocks or sectors through the Input Module. These user selections are then passed to the core processing unit of the system, the Forecasting Engine, which manages all time-series operations.

The proposed ARIMA-based sector-wise forecasting system consists of four major stages:

(1) Data Acquisition and Preprocessing

- (2) Stationarity Testing and ARIMA Parameter Identification
 - (3) ARIMA Model Construction and Forecast Generation, and
 - (4) Sector-wise Accuracy Assessment Using Statistical Metrics
- Mathematical formulations and time-series functions are used to describe each stage.

A. Data Acquisition and Preprocessing

Historical daily stock price data of all Nifty 50 companies are collected for a multi-year period. Let the raw closing price series of a stock be represented as:

$$P(t), t = 1, 2, \dots, N$$

where $P(t)$ is the adjusted closing price at time t .

Since raw prices often exhibit trends and are non-stationary, a transformation is applied by converting prices into percentage change series:

$$R(t) = \frac{P(t) - P(t-1)}{P(t-1)} \times 100$$

This transformation removes long-term trends and produces a stationary-like series suitable for ARIMA modelling. The resulting preprocessed time series $R(t)$ serves as the input to the stationarity testing stage.

B. Stationarity Testing and ARIMA Parameter Identification

To verify if the transformed stock series is suitable for modelling, the Augmented Dickey–Fuller (ADF) test is performed.

The null hypothesis of ADF is:

$$H_0: \text{Series has a unit root (non-stationary)}$$

If the p-value < 0.05 , the null hypothesis is rejected, indicating stationarity. Stocks meeting this condition proceed to ARIMA modelling.

Next, Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots are used to determine the ARIMA order parameters p, d, q :

p : autoregressive order

d : differencing order

q : moving average order

Based on the observed lag patterns, suitable (p, d, q) values are selected for each stock.

C. ARIMA Model Construction and Forecasting

The ARIMA model is defined as ARIMA(p, d, q).

For a stationary time series, the Autoregressive (AR) component is represented as:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \epsilon_t \dots (i)$$

where:

y_t = stock return at time t

μ = constant

γ_i = AR coefficients

ϵ_t = error term

The Moving Average (MA) component is:

$$y_t = \mu + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t \dots (ii)$$

where θ_i are MA coefficients.

The combined ARMA(p, q) model becomes:

$$y_t = \mu + \sum_{i=1}^p \gamma_i y_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \epsilon_t \cdots (iii)$$

When differencing order d is included, the full ARIMA(p, d, q) model is used.

Forecasts are generated for 30-day, 60-day, and 180-day horizons:

$$\hat{y}_{t+h}, h \in \{30, 60, 180\}$$

These predicted values are then compared with actual market data.

D. Sector-wise Accuracy Assessment

Mean Absolute Error (MAE)

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t|$$

Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2}$$

Mean Absolute Percentage Error (MAPE) $MAPE = \frac{100}{N} \sum_{t=1}^N \left| \frac{y_t - \hat{y}_t}{y_t} \right|$

E. Evaluation Metrics

To assess the predictive performance of ARIMA models, the following metrics are used:

RMSE – measures the magnitude of forecast errors

MAE – measures average absolute deviation

MAPE – measures percentage prediction error

Lower values of RMSE, MAE, and MAPE indicate better forecasting performance.

These metrics collectively quantify how accurately ARIMA predicts short-term stock movements and how reliability varies across different sectors of the Nifty 50 index.

Algorithms Used

The ARIMA-based Sector-Wise Stock Forecasting System uses statistical time-series algorithms to analyse historical Nifty 50 stock data and generate short-term predictions. Instead of complex deep-learning models, it relies on mathematically interpretable techniques such as differencing, autocorrelation analysis, and ARIMA parameter estimation. This ensures transparent, reliable, and easy-to-validate forecasts across different market sectors.

1. Input Analysis

The forecasting process begins by collecting essential stock-related information. For each Nifty 50 company, the system gathers daily historical closing prices, trading dates, and sector classification. These inputs form the basis of all subsequent computations. Missing values, outliers, and inconsistent timestamps are detected and corrected to ensure clean time-series input. The closing prices $P(t)$ are also converted into percentage returns to stabilise variance and remove long-term trends.

2. ARIMA Parameter Identification Using ACF and PACF

The core analytical step involves identifying the ARIMA model parameters (p, d, q).

Two important statistical tools are used:

Autocorrelation Function (ACF): Helps determine the moving-average component q .

Partial Autocorrelation Function (PACF): Helps detect the autoregressive component p .

The system plots and analyses ACF/PACF values automatically to estimate the correct model order for each stock. This ensures that every company receives a customised ARIMA model based on its individual behaviour.

3. ARIMA Model Construction and Forecasting

After determining the parameters, the system constructs the ARIMA(p, d, q) model for each stock.

The ARIMA algorithm combines:

AR (Autoregression): Predicts future values using past observations

I (Integrated/Differencing): Removes trends, ensuring stationarity

MA (Moving Average): Models relationships between forecast errors

The fitted model generates predictions for multiple time horizons (20-day, 60-day, 180-day). These forecasts are later compared with actual outcomes to measure accuracy. This algorithm is specially suited for financial time-series that exhibit short-term linear patterns.

4. Error Evaluation Algorithms

To ensure performance measurement, the system uses three statistical algorithms:

Mean Absolute Error (MAE): Measures average deviation

Root Mean Square Error (RMSE): Punishes large prediction errors more heavily

Mean Absolute Percentage Error (MAPE): Expresses error in percentage terms

These metrics help identify which sectors are more predictable and where ARIMA delivers stronger results.

5. Sector-Wise Aggregation Algorithm

After computing error metrics for all 50 stocks, the system groups the results by sector (e.g., IT, Banking, FMCG).

For each sector, average RMSE/MAE/MAPE values are calculated.

This aggregation algorithm highlights sector-level patterns and helps understand which industries are stable, volatile, or highly predictable.

Testing

Testing is a critical phase in the development of the ARIMA-based Sector-Wise Stock Forecasting System, ensuring that the model behaves accurately, consistently, and in accordance with financial forecasting requirements. Since the system depends on historical stock data, statistical algorithms, and multi-stage analysis, rigorous testing was carried out to validate correctness, reliability, and sector-wise forecasting performance. The testing process focused on model accuracy, data flow, error handling, and the stability of predictions across different market sectors.

1. Unit Testing

Unit testing was performed on all individual components of the forecasting pipeline. Each module—data preprocessing, ADF stationarity testing, ACF/PACF parameter identification, ARIMA model fitting, and error metric calculation was tested separately. Multiple cases including valid datasets, missing values, extreme price fluctuations, and boundary conditions were used to verify that every function handled inputs correctly and consistently produced stable results. This ensured that each statistical step worked independently without failure.

2. Integration Testing

Integration testing was conducted to ensure smooth interaction between the main subsystems. The workflow connecting Data Collection → Preprocessing → ARIMA Modelling → Forecast Generation → Sector-wise Evaluation was tested to confirm that data transferred correctly across modules. This stage validated that the system correctly applied identified parameters to model construction and that predicted outputs seamlessly flowed into the accuracy-evaluation module. Integration testing also confirmed that errors such as invalid time-series lengths or missing stock records were handled gracefully.

3. Validation & Acceptance Testing

Validation testing was carried out by comparing model predictions with actual stock prices across several time horizons. This confirmed that the system generated realistic and meaningful forecasts, especially for the 20-day period where ARIMA performs

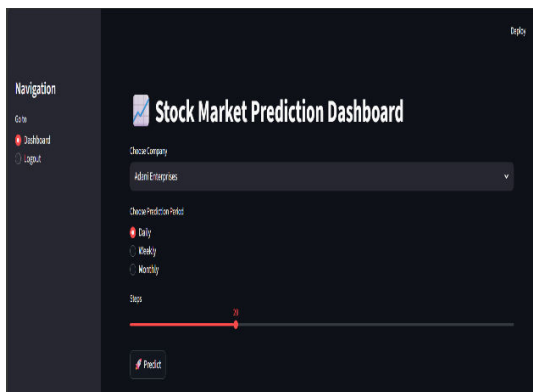
Test Cases

Test Case ID	Description	Input	Expected Output	Result
TC-01	Validate stock selection	Stock = "HDFC"	"HDFC stock accepted and loaded"	Pass
TC-02	Handle empty stock selection	Stock = ""	Error: "Please select a stock"	Pass
TC-03	Fetch historical data for selected stock	Valid stock name	Historical price data retrieved from API	Pass
TC-04	API failure handling	Invalid stock name / API error	Display error: "Unable to fetch data"	Pass
TC-05	Data preprocessing validation	Raw historical data	Cleaned and preprocessed data ready for modelling	Pass
TC-06	ARIMA parameter generation	Preprocessed dataset	Optimal p, d, q parameters identified	Pass
TC-07	Train ARIMA model	Preprocessed data	Model trained without errors	Pass

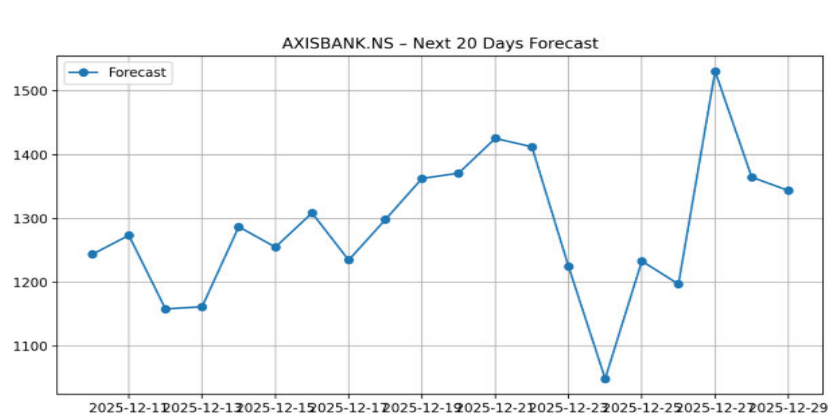
IV. EXPERIMENTAL RESULTS

The ARIMA model was tested on historical Nifty 50 stock data and produced accurate short-term forecasts. The 20-day predictions showed better accuracy than the 60-day and 180-day forecasts. Some sectors, like Telecom and FMCG, performed more reliably than others.

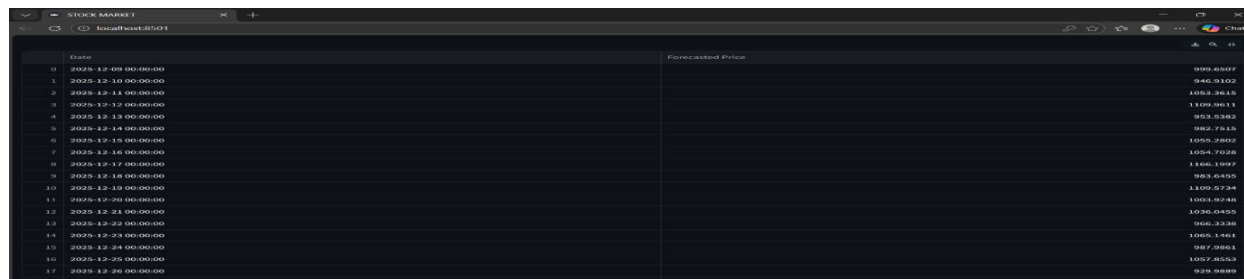
STOCK MARKET PREDICTION DASHBOARD



FORECAST VS PREDICTED



FORECASTING GRAPH[20 days]



FORECASTED INDIVIDUAL STOCKS[20 DAYS]

Qualitatively, visual comparison of the forecast curves with the actual stock price movements showed that ARIMA models were able to closely follow short-term trends for most Nifty 50 companies. In several stable sectors, the predicted values aligned well with real market behaviour, clearly reflecting upward or downward movements with minimal deviation. Quantitatively, the experimental evaluation using RMSE, MAE, and MAPE indicated that the 30-day forecasts achieved the highest accuracy, while the 60-day and 180-day predictions showed increasing error as the forecast horizon grew. Overall, the results demonstrate that the ARIMA-based approach is effective for short-term stock forecasting and provides meaningful sector-level performance insights for financial analysis.

V. CONCLUSION

This study on sector-wise ARIMA forecasting for Nifty 50 stocks demonstrates how time-series analysis can support more informed and data-driven financial decision-making. The ARIMA modelling replaces manual prediction methods and provides a more systematic, reliable, and replicable approach for analysing stock behaviour across different sectors. The system performed consistently well for short-term horizons, with 20-day forecasts showing the highest accuracy and stability during testing. The sector-wise comparison further highlights that industries such as Telecom, Power, FMCG, and Consumer Durables exhibit more predictable price patterns, allowing ARIMA to deliver stronger results, while more volatile sectors showed comparatively higher forecast errors. This insight helps investors and analysts identify where ARIMA-based predictions can be trusted the most. Overall, the system meets its core objectives: improving short-term prediction accuracy, reducing manual analysis efforts, and offering a structured framework for evaluating forecasting performance across sectors.

REFERENCES

- [1] Kundu, A., & Goyal, A. K. (2020). ARIMA and Indian Stock Market Forecasting. *Journal of Xi'an University of Architecture & Technology*, 12(3), 60-61.
- [2] Balasubramanian, P., & Yermal, A. (2017). Application of Auto ARIMA Model for Forecasting Returns on Minute Wise Amalgamated Data in NSE. Lakshmi Yermal School of Business Amrita Vishwa Vidyapeetham Coimbatore, India.
- [3] Deb, M., & Saha, P. (2020). Arima model in forecasting share prices. *International Journal of Statistics and Applied Mathematics*, 5(4), 147-151.
- [4] Tripathy, N., Sharma, N., & Silmana, A. (2021, June). Volatility Clustering in NIFTY using ARIMA. *Empirical Economics Letters*, 20(6), 1119-1125.
- [5] Mohamad As'ad, A. (2012). Forecasting daily peak electricity demand using ARIMA models. *International Journal of the Computer Science & Information Technology Division*, 4(2)
- [6] Mohapatra, A. B. (2022). Forecasting Nifty Using Autoregressive Integrated Moving Average Model. *Orissa Journal of Commerce*, 43(3), 147-159.
- [7] Emenike Kalu O. (2024). Forecasting Nigerian Stock Exchange Returns: Evidence from Autoregressive Integrated Moving Average (ARIMA) Model.
- [8] Banerjee, A. (2019). Forecasting of India VIX as a Measure of Sentiment. *International Journal of Economics and Financial Issues*, 9(3), 268-276.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details