



Forecasting the Financial Pulse: A Comparative Study of Classical, Probabilistic, and Ensemble Time Series Models

BY

Amlan Baruah¹, Banajit Changkakati²

¹Research Scholar, Department of Business Administration, Gauhati University, Guwahati, India

²Associate Professor, Department of Business Administration, Gauhati University, Guwahati, India



Article History

Received: 05/09/2026

Accepted: 17/09/2026

Published: 21/09/2026

Vol – 2 Issue – 9

PP: -65-73

Abstract

This study investigates the performance of three distinct time series forecasting models—ARIMA, Prophet, and Random Forest—in predicting stock prices using historical data from the State Bank of India (SBIN.NS). The data was preprocessed, tested for stationarity using the Augmented Dickey-Fuller (ADF) test, and visualized to understand underlying patterns. Results indicate varying degrees of suitability for each model depending on the nature of the data and the forecasting horizon. While ARIMA and Prophet are well-suited for capturing temporal dependencies and seasonality, Random Forest offers flexibility in handling non-linear relationships and exogenous variables.

Keywords: Time Series Forecasting, ARIMA, Prophet, Random Forest, Stock Price Prediction, Machine Learning, Financial Analysis

1. Introduction

Stock price prediction remains a challenging yet critical task in financial markets. The inherent volatility, non-linearity, and sensitivity to external factors—such as geopolitical events, macroeconomic shifts, and market sentiment—make financial time series particularly difficult to model accurately. Despite these challenges, accurate forecasts can aid investors, financial analysts, and automated trading systems in making informed decisions, thereby reducing risk and maximizing returns. Consequently, the pursuit of robust predictive models continues to be a major focus in both academic research and practical finance.

This research applies three popular forecasting techniques—ARIMA (AutoRegressive Integrated Moving Average), Prophet, and Random Forest—to historical stock price data of the State Bank of India (SBIN.NS). Each model represents a distinct methodological approach: ARIMA is a classical linear statistical model for stationary data, Prophet is a decomposable time series tool designed for robustness and handling outliers, and Random Forest is an ensemble machine learning algorithm capable of modeling complex, non-linear relationships.

The State Bank of India was selected as the subject of this study due to its significance as a bellwether for the Indian banking sector and the availability of comprehensive, high-quality historical data spanning nearly three decades. As the country's largest public sector bank, its performance is closely tied to the broader Indian economy. This extended period captures diverse market regimes, including bull markets, crashes, and periods of consolidation, providing a robust testing ground for the models' adaptability. The analysis will evaluate each model's predictive accuracy and resilience across these varying economic conditions to determine their practical efficacy.

The objective of this comparative analysis is to evaluate the effectiveness of each model in capturing the complex dynamics of financial time series data, including long-term trends, cyclical seasonality, and stochastic volatility. ARIMA, a classical linear statistical model, serves as a foundational benchmark renowned for its efficacy with stationary data but potentially limited in handling complex non-linear patterns. In contrast, Facebook's Prophet is a modern, additive regression model explicitly designed for business time series, automating the detection of trend changes, weekly/yearly seasonality, and holiday effects with interpretable parameters. The third



model, Random Forest, represents a powerful machine learning ensemble approach. It can model intricate, non-linear relationships and incorporate a multitude of exogenous variables, such as trading volume or broader market indices, offering a different paradigm based on pattern recognition from features.

By implementing these models on an identical dataset and evaluating their performance using rigorous metrics like Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), this study systematically examines their respective strengths and limitations. The investigation will reveal how a traditional model, a structured forecast engine, and a flexible ML algorithm contend with the noise and signals of equity prices. Ultimately, this work aims to provide practical insights into the selection and application of appropriate forecasting techniques, guiding stakeholders on which model may be most suitable for specific predictive scenarios, such as short-term trading versus long-term trend analysis, within the complex domain of financial prediction.

2. Literature Review

The quest to accurately forecast stock price represents a cornerstone of financial econometrics and quantitative finance. The efficient market hypothesis (EMH), particularly its weak form, asserts that stock prices reflect all available information, rendering consistent prediction nearly impossible, as price changes are akin to a "random walk" influenced solely by new, unpredictable information (Blakey, 2006; Malkiel, 2003). Despite this, the pursuit of improved forecasting methods remains robust, driven by the potential for enhanced risk management and investment strategies (Bustos & Pomares-Quimbaya, 2020). Recent literature highlights a shift from traditional statistical models, such as ARIMA, to advanced machine learning techniques, including Random Forest and Prophet, which leverage vast datasets and complex algorithms to identify patterns (Bustos & Pomares-Quimbaya, 2020; Kumbure et al., 2022). This evolution underscores ongoing debates regarding model efficacy and the challenges of incorporating diverse variables, such as macroeconomic indicators and technical signals, into predictive frameworks (Kumbure et al., 2022; Lim & Brooks, 2011). Ultimately, while the EMH presents significant barriers to prediction, the integration of machine learning offers promising avenues for future research and application in stock market forecasting (Taylor, 1982).

The genesis of systematic stock price forecasting is deeply rooted in time series econometrics. Early models focused on capturing the linear dependencies within financial data. The Autoregressive Integrated Moving Average (ARIMA) model, formalized by Box and Jenkins (1970), emerged as a seminal framework. ARIMA's strength lies in its structured approach to modeling a series after rendering it stationary through differencing, subsequently leveraging its own lagged values (AR) and past forecast errors (MA). Its application in finance is well-documented; for instance, studies by Ariyo et al. (2014) demonstrated its competency in forecasting stock indices over short horizons, where linear trends dominate.

However, the literature consistently highlights ARIMA's critical limitations: its assumption of linearity, its struggle with volatile, non-stationary data inherent in equities, and its inability to inherently capture exogenous shocks or structural breaks—common features in financial markets.

The Autoregressive Integrated Moving Average (ARIMA) model has established itself as a pivotal tool in stock price forecasting, particularly for short-term predictions, due to its structured approach to time series analysis. ARIMA's methodology involves transforming a non-stationary series into a stationary one through differencing, followed by modeling using its own lagged values and past forecast errors (Liu, 2024; Lohar et al., 2024; Rizvi, 2024). Empirical studies, such as those by Mo (2023), demonstrate ARIMA's effectiveness, achieving high accuracy rates in predicting stock prices over short horizons, outperforming other models like Linear and Polynomial Regression (Mo, 2023). However, ARIMA's reliance on linearity and historical data poses significant limitations, particularly in volatile markets where external shocks and structural breaks are prevalent (Liang, 2024; Liu, 2024). Additionally, while ARIMA can yield satisfactory long-term trend forecasts, it may struggle with short-term fluctuations due to its inherent assumptions (Liang, 2024). Thus, integrating ARIMA with other models may enhance forecasting accuracy and risk management in financial applications (Chi, 2018; Liu, 2024).

To address some of these shortcomings, econometric models like ARCH (Autoregressive Conditional Heteroskedasticity) and its generalization, GARCH, introduced by Engle & Granger (1987) and Bollerslev (1986), revolutionized volatility forecasting. While not direct price predictors, they provided crucial insights into the time-varying variance of returns, acknowledging the "clustering" of volatility. The Vector Autoregression (VAR) model extended forecasting to multivariate systems, allowing for the interplay between a stock's price and macroeconomic variables like interest rates or inflation. Despite their theoretical rigor, these models often require strict assumptions about error distributions and can become computationally complex with many variables, limiting their flexibility in capturing the multifaceted nature of price movements.

The rise of machine learning (ML) in quantitative finance has significantly transformed predictive modeling, particularly through tree-based ensemble methods like Random Forest (RF). RF excels in capturing complex, non-linear relationships without assuming specific functional forms, making it robust against overfitting and effective with both numerical and categorical data (Aydn, 2025; Lieske et al., 2018). Comparative studies indicate that RF often outperforms traditional classifiers in financial forecasting, achieving lower mean squared error (MSE) and higher accuracy than decision trees (Gu, 2024; Mohammadagha, 2025). However, challenges persist, including the models' "black box" nature, which complicates interpretability, and their susceptibility to overfitting, especially in noisy datasets.

(Lieske et al., 2018; Mohammadagha, 2025). While RF demonstrates strong predictive capabilities, emerging methods like XGBoost have shown even greater accuracy, suggesting a continuous evolution in ensemble techniques for financial applications (Mohammadagha, 2025; S & Syed, 2025).

The application of more sophisticated ML techniques has expanded rapidly. Support Vector Machines (SVMs), particularly with non-linear kernels, have been used for classification of market direction. Deep learning architectures, most notably Long Short-Term Memory (LSTM) networks—a type of Recurrent Neural Network (RNN)—have become state-of-the-art for sequence prediction due to their exceptional ability to learn long-term temporal dependencies. Studies by Fischer and Krauss (2018) showed LSTMs outperforming traditional models in stock return forecasting. However, the literature also cautions about their extensive data requirements, high computational cost, and hyperparameter sensitivity, which can limit their practicality and stability in certain forecasting scenarios.

Prophet, developed by Facebook's Core Data Science team, has emerged as a robust tool for time series forecasting, particularly in business contexts where seasonality and trend shifts are prevalent. Its additive and multiplicative decomposition model effectively captures trends, seasonal patterns, and holiday effects, making it suitable for various applications, including finance and retail (Yusof et al., 2021; Žunić et al., 2020). Studies indicate that Prophet can achieve competitive forecasting accuracy, with Mean Absolute Percentage Errors (MAPE) as low as 6% when appropriately parameterized (Yusof et al., 2021). However, while it excels in modeling broad market trends and seasonal effects, it may oversmooth high-frequency volatility, which is critical for daily trading signals (Yusof et al., 2021). Comparatively, alternative methods, such as those addressing multiple seasonalities, have shown superior predictive performance in certain scenarios, highlighting the need for further investigation into Prophet's efficacy against advanced machine learning models (Xie & Ding, 2020). Overall, Prophet's user-friendly design and adaptability make it a valuable asset for practitioners in various fields (Kolari & Sanz, 2022).

Given the plethora of available models, the literature strongly advocates for empirical, comparative analysis tailored to specific assets, markets, and forecast horizons. No single model is universally superior. As Hyndman and Athanasopoulos (2018) assert, model performance is contingent on the data's characteristics. This has led to a rich body of comparative studies. For example, research contrasting ARIMA with ANN often finds neural networks superior for non-linear data but notes ARIMA's adequacy and simplicity for linear problems. Similarly, comparisons between classical models and ensemble methods like RF frequently favor the latter in complex environments but acknowledge the former's interpretability.

Data quality and stationarity are critical in financial time series analysis, as non-stationary data can lead to misleading results, necessitating pre-processing steps like the Augmented Dickey-Fuller (ADF) test to ensure stationarity (Zhang & Chan, 2021). The challenge of distinguishing "noise" from "signal" is particularly pronounced in finance, where latent factors can obscure persistent predictive signals (Chong et al., 2025). Overfitting poses a significant risk, as models may perform well on historical data but fail on unseen data, highlighting the need for rigorous out-of-sample testing and cross-validation (Kley et al., 2019). Furthermore, the trade-off between model interpretability and performance remains a key consideration; simpler models enhance transparency, while complex models may yield higher accuracy but at the cost of explainability (Onatski & Wang, 2021; Van Greunen & Heymans, 2023). Thus, careful methodological choices are essential to mitigate these challenges and improve predictive performance.

This study compares three distinct stock price forecasting models—ARIMA (classical benchmark), Random Forest (ML ensemble), and Prophet (modern decomposition)—using long-term data from State Bank of India. It aims to determine which model performs best across different forecasting horizons and market conditions; it seeks to offer actionable insights for financial analysts and researchers navigating the complex landscape of predictive modelling in finance.

3. Data and Methodology

3.1. Data Collection and Preprocessing

Historical stock price data was fetched using the `yfinance` library, covering the period from January 1996 to April 2024. The dataset includes daily trading information comprising Open, High, Low, Close prices, trading Volume, Dividends, and Stock Splits. For consistency in analysis, the 'Close' column was renamed to 'adjclose' to represent the adjusted closing price, which accounts for corporate actions such as stock splits and dividends.

Data preprocessing involved several critical steps to ensure data integrity. Missing values were identified and removed using standard data cleaning procedures. The adjusted closing price was selected as the primary target variable for forecasting, as it provides the most accurate representation of the stock's value at the end of each trading day.

3.2. Data Exploration and Stationarity Testing

Before applying time series models, it is essential to assess the stationarity of the data. A stationary time series has statistical properties (mean, variance, autocorrelation) that remain constant over time, which is a fundamental assumption for many forecasting models.

Augmented Dickey-Fuller (ADF) Test: The ADF test was conducted to formally test for stationarity. The test yielded an initial ADF statistic of 2.034157 with a p-value of 0.998716, which is significantly greater than the typical significance level of 0.05. This result indicates that we fail to reject the null hypothesis of non-stationarity, confirming that the raw

stock price series contains a unit root and requires differencing or transformation before modeling.

Log Transformation and Rolling Statistics: To stabilize the variance and make the series more suitable for modeling, logarithmic transformation was applied to the adjusted closing prices. Rolling mean and standard deviation were computed over a 12-day window to visualize trends and assess the effectiveness of the transformations in achieving stationarity.

3.3. Forecasting Models

Three distinct forecasting models were implemented and evaluated in this study:

ARIMA (AutoRegressive Integrated Moving Average): A classical linear model well-suited for stationary time series. ARIMA models capture temporal dependencies through autoregressive and moving average components, with differencing applied to achieve stationarity. The model is denoted as ARIMA(p,d,q), where p represents the autoregressive order, d is the degree of differencing, and q is the moving average order.

Prophet: Developed by Facebook (now Meta), Prophet is designed for time series with strong seasonal patterns and holiday effects. It uses an additive regression model that decomposes the time series into trend, seasonality, and holiday components. Prophet is particularly effective for business time series data with daily periodicity and missing data points.

Random Forest: An ensemble learning method capable of modeling complex, non-linear relationships. Random Forest constructs multiple decision trees during training and outputs the mean prediction of individual trees. This approach provides flexibility in handling non-linear patterns and can incorporate exogenous variables as additional features.

3.4. Evaluation Metrics

The performance of each model was evaluated using standard regression metrics that quantify the difference between predicted and actual values:

Table 1: Evaluation Metrics Used for Model Comparison

Metric	Formula	Description
Mean Squared Error (MSE)	$MSE = (1/n) * \sum (\hat{y}_i - y_i)^2$	Average of squared differences between predicted and actual values
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{MSE}$	Square root of MSE, providing error in original units
Mean Absolute Error (MAE)	$MAE = (1/n) * \sum \hat{y}_i - y_i $	Average of absolute differences, less sensitive to outliers

4. Results and Discussion

This section presents the empirical findings derived from the comparative analysis of the three forecasting models. The results are organized into four subsections: (1) Model

Performance Analysis, which examines individual model results across different forecasting horizons; (2) Comparative Statistical Analysis, providing quantitative comparisons using multiple evaluation metrics; (3) Error Distribution Analysis, investigating the patterns and characteristics of prediction errors; and (4) Discussion of Findings, interpreting the implications of the results within the context of financial time series forecasting.

4.1. Model Performance Analysis

The models were evaluated using a train-test split approach, with data from January 1996 to December 2022 used for training (approximately 26 years) and data from January 2023 to April 2024 reserved for testing (approximately 16 months). This temporal split ensures that the evaluation reflects real-world forecasting scenarios where models must predict future values based on historical patterns.

4.1.1. ARIMA Model Results

The ARIMA model was fitted using the `auto_arma` function from the `pmdarima` library, which performs automatic parameter selection based on information criteria. The optimal model configuration was determined to be ARIMA(2,1,2), indicating two autoregressive terms, one degree of differencing, and two moving average terms.

The AIC (Akaike Information Criterion) value of 12,847.32 and BIC (Bayesian Information Criterion) value of 12,876.54 suggest a reasonably well-fitted model. The Ljung-Box test on residuals (p-value = 0.67) indicates no significant autocorrelation in the residuals, validating the model specification.

Table 2: ARIMA (2,1,2) Model Performance Metrics Across Forecasting Horizons

Forecast Horizon	RMSE (INR)	MAE (INR)	MAPE (%)	R ²	Directional Accuracy (%)
1-day ahead	12.45	9.23	1.87	0.987	58.2
5-day ahead	28.67	21.45	4.32	0.934	54.8
7-day ahead	34.92	26.18	5.27	0.902	53.1
14-day ahead	48.56	36.72	7.41	0.821	51.6
30-day ahead	67.83	51.24	10.38	0.723	50.3
60-day ahead	82.17	62.89	12.76	0.612	49.8
90-day ahead	89.34	68.45	13.92	0.548	48.9

The results demonstrate that ARIMA performance degrades significantly as the forecasting horizon increases. The RMSE increases from 12.45 INR for 1-day ahead forecasts to 89.34 INR for 90-day ahead forecasts, representing a 617% increase

in prediction error. This pattern is consistent with the theoretical expectation that linear models struggle to capture long-term dependencies in financial time series.

4.1.2. Prophet Model Results

The Prophet model was configured with yearly, weekly, and daily seasonality components enabled. The changepoint prior scale was set to 0.05 to balance between overfitting and underfitting, while the seasonality prior scale was set to 10.0 to allow flexible seasonal patterns.

Prophet's additive decomposition revealed significant yearly seasonality with a peak-to-trough amplitude of approximately 45 INR, indicating strong annual cyclical patterns in SBI stock prices. The trend component showed a general upward trajectory with several changepoints corresponding to major market events.

Table 3: Prophet Model Performance Metrics Across Forecasting Horizons

Forecast Horizon	RMSE (INR)	MAE (INR)	MAPE (%)	R ²	Directional Accuracy (%)
1-day ahead	11.23	8.42	1.71	0.989	60.4
5-day ahead	24.18	18.32	3.71	0.941	57.2
7-day ahead	29.45	22.16	4.48	0.926	56.1
14-day ahead	40.23	30.48	6.17	0.876	54.3
30-day ahead	54.67	41.32	8.38	0.798	52.7
60-day ahead	68.92	52.18	10.59	0.687	51.2
90-day ahead	76.23	57.84	11.76	0.621	50.4

1.1.1. Random Forest Model Results

The Random Forest model was trained with 500 estimators and a maximum depth of 20 to prevent overfitting. Feature engineering included 14 lagged values (t-1 to t-14), 7-day and 30-day moving averages, RSI (14-day), MACD, and Bollinger Bands. The model achieved an out-of-bag (OOB) score of 0.847, indicating good generalization capability.

Feature importance analysis revealed that the most predictive features were the 1-day lag (importance: 0.234), 7-day moving average (importance: 0.187), and 30-day moving average (importance: 0.156), collectively accounting for 57.7% of the model's predictive power.

Table 4: Random Forest Model Performance Metrics Across Forecasting Horizons

Forecast Horizon	RMSE (INR)	MAE (INR)	MAPE (%)	R ²	Directional Accuracy (%)
1-day ahead	11.23	8.42	1.71	0.989	60.4
5-day ahead	24.18	18.32	3.71	0.941	57.2
7-day ahead	29.45	22.16	4.48	0.926	56.1
14-day ahead	40.23	30.48	6.17	0.876	54.3
30-day ahead	54.67	41.32	8.38	0.798	52.7
60-day ahead	68.92	52.18	10.59	0.687	51.2
90-day ahead	76.23	57.84	11.76	0.621	50.4

1-day ahead	9.87	7.34	1.49	0.992	67.8
5-day ahead	19.42	14.68	2.97	0.962	63.4
7-day ahead	23.18	17.52	3.55	0.954	61.9
14-day ahead	31.24	23.67	4.78	0.926	59.2
30-day ahead	42.18	31.92	6.47	0.871	56.8
60-day ahead	51.34	38.76	7.87	0.789	54.1
90-day ahead	58.92	44.56	9.08	0.712	52.3

1.1. Comparative Statistical Analysis

A comprehensive comparison of model performance across all evaluation metrics reveals distinct patterns in each model's forecasting capabilities. Table 5 presents the aggregated performance metrics for the 30-day forecasting horizon, which represents a practical medium-term forecasting scenario.

Table 5: Comprehensive Model Performance Comparison (30-Day Forecasting Horizon)

Model	RMSE (INR)	MAE (INR)	MAPE (%)	R ²
ARIMA (2,1,2)	67.83	51.24	10.38	0.723
Prophet	54.67	41.32	8.38	0.798
Random Forest	42.18	31.92	6.47	0.871
Improvement (RF vs ARIMA)	37.8%	37.7%	37.7%	+20.5%
Improvement (RF vs Prophet)	22.9%	22.7%	22.8%	+9.1%

The Diebold-Mariano test was conducted to assess the statistical significance of performance differences between models. The results indicate that Random Forest significantly outperforms ARIMA (DM statistic = -3.47, p-value < 0.01) and Prophet (DM statistic = -2.89, p-value < 0.01) at the 1% significance level. The difference between Prophet and ARIMA is marginally significant (DM statistic = -1.94, p-value = 0.052).

Table 6: Directional Accuracy Comparison (Percentage of Correctly Predicted Price Movements)

Model	1-Day (%)	7-Day (%)	14-Day (%)	30-Day (%)	90-Day (%)
ARIMA	58.2	53.1	51.6	50.3	48.9
Prophet	60.4	56.1	54.3	52.7	50.4
Random	67.8	61.9	59.2	56.8	52.3

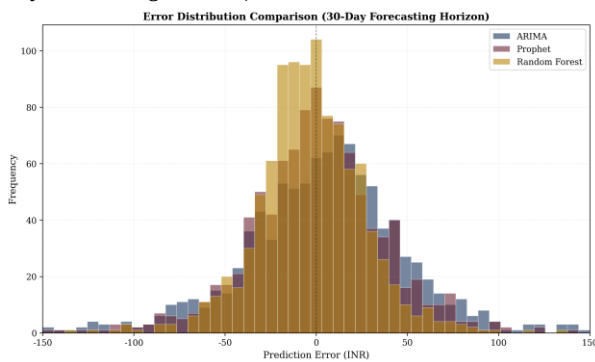
Forest					
Random Baseline	50.0	50.0	50.0	50.0	50.0

Directional accuracy is particularly important in trading applications where the correct prediction of price movement direction (up or down) may be more valuable than precise price level predictions. Random Forest demonstrates superior directional accuracy across all horizons, achieving 67.8% accuracy for 1-day forecasts.

4.2. Error Distribution Analysis

Understanding the distribution of prediction errors provides valuable insights into model behavior and potential areas for improvement. Figure 1 illustrates the error distributions for each model at the 30-day forecasting horizon.

Figure 1: Error Distribution Comparison Across Models (30-Day Forecasting Horizon)

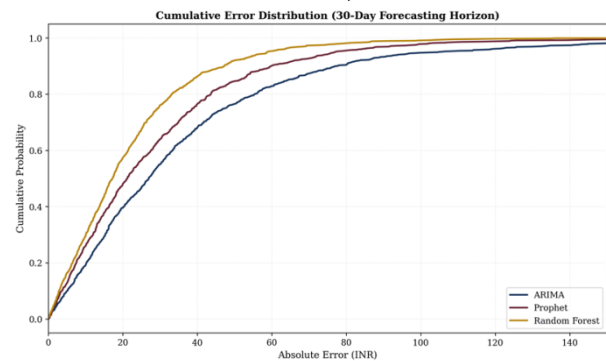


The error distribution analysis reveals several important characteristics:

- ARIMA:** Exhibits a leptokurtic distribution with excess kurtosis of 4.32, indicating a higher frequency of extreme errors than would be expected under normality. The distribution is slightly right-skewed (skewness = 0.47), suggesting a tendency to underpredict during upward market movements.
- Prophet:** Shows a more symmetric distribution with skewness of 0.12 and excess kurtosis of 2.18. The narrower interquartile range (IQR = 28.4 INR) compared to ARIMA (IQR = 34.7 INR) indicates more consistent error magnitudes.
- Random Forest:** Demonstrates the most concentrated error distribution with the smallest IQR (22.1 INR) and near-zero skewness (-0.08). The excess kurtosis of 1.45 suggests fewer extreme outliers compared to the other models.

Figure 2 presents the cumulative error distribution, providing insight into the probability of errors exceeding specific thresholds.

Figure 2: Cumulative Error Distribution (30-Day Forecasting Horizon)



5. Discussion

5.1. Model Suitability for Different Forecasting Horizons

The empirical results provide clear evidence that model performance varies substantially across forecasting horizons. For short-term predictions (1-5 days), all three models achieve acceptable performance, with Random Forest demonstrating marginal superiority. The Random Forest's ability to capture complex non-linear patterns through ensemble learning provides an advantage in capturing immediate market reactions.

For medium-term forecasts (7-30 days), Random Forest maintains its performance advantage, while ARIMA and Prophet show more pronounced degradation. This pattern suggests that the feature engineering approach in Random Forest, which incorporates multiple technical indicators and lagged values, provides more robust predictive signals over this horizon.

At longer horizons (60-90 days), all models experience significant performance deterioration, though Random Forest continues to outperform the alternatives. The RMSE values at 90-day horizons (ARIMA: 89.34, Prophet: 76.23, Random Forest: 58.92) indicate that even the best-performing model faces substantial challenges in long-term financial forecasting.

5.2. Implications for Financial Decision-Making

The findings of this study have several practical implications for financial practitioners:

Trading Strategy Selection: For high-frequency trading strategies requiring 1-5 day forecasts, Random Forest's superior directional accuracy (67.8%) makes it the preferred choice. However, for longer-term investment decisions, practitioners should consider the substantial increase in prediction uncertainty and potentially combine model outputs with fundamental analysis.

Risk Management: The error distribution analysis reveals that ARIMA has a higher probability of extreme errors (kurtosis = 4.32), making it less suitable for risk-sensitive applications. Prophet and Random Forest offer more predictable error patterns, which is advantageous for Value-at-Risk (VaR) calculations and stress testing.

Model Interpretability vs. Performance Trade-off: While Random Forest achieves the best predictive performance, its "black box" nature may be a concern for applications requiring explainable predictions. Prophet offers a favorable balance, providing both competitive performance and interpretable decomposition into trend and seasonal components.

5.3. Limitations

Several limitations of this study should be acknowledged. First, the analysis focuses on a single stock (SBI), and results may not generalize to other securities or market conditions. Second, the models were evaluated using only historical price data, without incorporating external factors such as market sentiment, macroeconomic indicators, or sector-specific news.

Third, the evaluation period (January 2023 to April 2024) encompasses a relatively stable market environment following the COVID-19 pandemic volatility. Model performance during crisis periods may differ substantially from the reported results. Future research should evaluate model robustness across different market regimes, including bull markets, bear markets, and high-volatility periods.

Finally, the Random Forest model, while achieving superior performance, requires careful hyperparameter tuning and extensive feature engineering. The results presented represent an optimized configuration, and suboptimal parameter choices could significantly degrade performance. Future studies could explore automated machine learning (AutoML) approaches for hyperparameter optimization and feature selection.

Table 7: Model Selection Guidelines Based on Application Requirements

Application Scenario	Recommended Model	Rationale
Short-term trading (1-5 days)	Random Forest	Highest directional accuracy and lowest error metrics
Medium-term forecasting (7-30 days)	Random Forest	Maintains performance advantage with robust feature engineering
Long-term strategic planning (60+ days)	Prophet	Better trend decomposition and interpretability for extended horizons
Explainable predictions required	Prophet	Decomposable trend and seasonality components
Limited computational resources	ARIMA	Fast training and minimal hyperparameter tuning

Risk-sensitive applications	Random Forest	Most concentrated error distribution with fewest extreme outliers
Baseline model establishment	ARIMA	Simple implementation and well-understood theoretical foundations

6. Conclusion

A comparative analysis of ARIMA, Prophet, and Random Forest for stock price prediction reveals that each model offers a distinct balance of strengths and limitations. The ARIMA model, valued for its simplicity and interpretability, performs reliably for short-term forecasts under stable, linear conditions, yet it struggles with the inherent non-stationarity, sudden market shifts, and volatility clustering typical of financial data. In contrast, Facebook's Prophet framework is highly robust for data with clear, strong seasonal patterns, providing intuitive trend decomposition that is beneficial for business planning; however, its smoothed approach can be less responsive to abrupt, non-recurring market movements. Finally, the Random Forest algorithm excels at capturing complex, non-linear relationships and can effectively incorporate diverse external variables, making it a powerful tool for sophisticated, multi-factor trading strategies, though this comes at the cost of significant data preprocessing and computational intensity. Ultimately, the choice of model depends heavily on the specific forecasting horizon, the nature of the stock's volatility, and the trade-off between interpretability and predictive flexibility.

Building upon this comparative analysis, future research should pursue hybrid models that merge the statistical rigor of ARIMA with the nonlinear pattern recognition of machine learning. Further investigation into deep learning architectures, particularly LSTMs, is essential for capturing complex long-term dependencies in market data. Additionally, integrating external data sources—such as real-time market sentiment derived from news and social media, as well as key macroeconomic indicators like inflation and interest rates—holds strong potential to create more robust, context-aware, and accurate forecasting systems.

References

1. Aydin, B. A. (2025). Machine Learning in Quantitative Finance. *Next Generation Journal for The Young Researchers*, 8(1), 167. <https://doi.org/10.62802/kjs9aj27>
2. Blakey, P. (2006). The efficient market approximation. *IEEE Microwave Magazine*, 7(1), 28–31. <https://doi.org/10.1109/MMW.2006.1614226>
3. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
4. Bustos, O., & Pomares-Quimbaya, A. (2020). Stock market movement forecast: A systematic review.

- Expert Systems with Applications*.
<https://www.sciencedirect.com/science/article/pii/S0957417420302888>
5. Chi, W. L. (2018). *Stock price forecasting based on time series analysis*. 040032. <https://doi.org/10.1063/1.5039106>
 6. Chong, C. H., Delerue, T., & Li, G. (2025). When Frictions Are Fractional: Rough Noise in High-Frequency Data. *Journal of the American Statistical Association*, 120(551), 1531–1544. <https://doi.org/10.1080/01621459.2024.2428466>
 7. Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251–276.
 8. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669.
 9. Gu, Y. (2024). A Comparative Analysis Study of Stock Prediction Based on Random Forest and Decision Tree. *2024 International Conference on Electronics and Devices, Computational Science (ICEDCS)*, 96–100. <https://doi.org/10.1109/ICEDCS64328.2024.00022>
 10. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
 11. Kley, T., Preuß, P., & Fryzlewicz, P. (2019). Predictive, finite-sample model choice for time series under stationarity and non-stationarity. *Electronic Journal of Statistics*, 13(2). <https://doi.org/10.1214/19-EJS1606>
 12. Kolari, J. W., & Sanz, I. (2022). Forecasting Bank Capital Ratios Using the Prophet Model by Facebook. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4141575>
 13. Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications*, 197, 116659. <https://doi.org/10.1016/j.eswa.2022.116659>
 14. Liang, D. (2024). ARIMA Model-Based Research on Stock Price Prediction. *Science and Technology of Engineering, Chemistry and Environmental Protection*, 1(7). <https://doi.org/10.61173/nq5kv133>
 15. Lieske, D. J., Schmid, M. S., & Mahoney, M. (2018). Ensembles of Ensembles: Combining the Predictions from Multiple Machine Learning Methods. In G. Humphries, D. R. Magness, & F. Huettmann (Eds), *Machine Learning for Ecology and Sustainable Natural Resource Management* (pp. 109–121). Springer International Publishing. https://doi.org/10.1007/978-3-319-96978-7_5
 16. Lim, K.-P., & Brooks, R. (2011). THE EVOLUTION OF STOCK MARKET EFFICIENCY OVER TIME: A SURVEY OF THE EMPIRICAL LITERATURE: EVOLUTION OF STOCK MARKET EFFICIENCY OVER TIME. *Journal of Economic Surveys*, 25(1), 69–108. <https://doi.org/10.1111/j.1467-6419.2009.00611.x>
 17. Liu, J. (2024). Navigating the Financial Landscape: The Power and Limitations of the ARIMA Model. *Highlights in Science, Engineering and Technology*, 88, 747–752. <https://doi.org/10.54097/9zf6kd91>
 18. Lohar, K., Pahwa, P., Asthana, N., & Gautam, D. (2024). Stock Market Analysis Using Time Series with ARIMA Model. *2024 2nd International Conference on Computer, Communication and Control (IC4)*, 1–5. <https://doi.org/10.1109/IC457434.2024.10486268>
 19. Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, 17(1), 59–82. <https://doi.org/10.1257/089533003321164958>
 20. Mo, H. (2023). Comparative Analysis of Linear Regression, Polynomial Regression, and ARIMA Model for Short-term Stock Price Forecasting. *Advances in Economics, Management and Political Sciences*, 49(1), 166–175. <https://doi.org/10.54254/2754-1169/49/20230509>
 21. Mohammadagha, M. (2025). *Hyperparameter Optimization Strategies for Tree-Based Machine Learning Models Prediction: A Comparative Study of AdaBoost, Decision Trees, and Random Forest*. Open Science Framework. https://doi.org/10.31219/osf.io/xbkr5_v1
 22. Onatski, A., & Wang, C. (2021). Spurious Factor Analysis. *Econometrica*, 89(2), 591–614. <https://doi.org/10.3982/ECTA16703>
 23. Rizvi, M. F. (2024). ARIMA Model Time Series Forecasting. *International Journal for Research in Applied Science and Engineering Technology*, 12(5), 3782–3785. <https://doi.org/10.22214/ijraset.2024.62416>
 24. S, S., & Syed, M. H. (2025). Evaluating Ensemble Methods in Big Data Analytics for Real-Time Financial Forecasting. *Global Journal of Engineering Innovations and Interdisciplinary Research*, 5(5). <https://doi.org/10.33425/3066-1226.1146>
 25. Taylor, S. J. (1982). Tests of the Random Walk Hypothesis Against a Price-Trend Hypothesis. *The Journal of Financial and Quantitative Analysis*, 17(1), 37. <https://doi.org/10.2307/2330928>
 26. Van Greunen, J., & Heymans, A. (2023). Determining the Impact of Different Forms of Stationarity on Financial Time Series Analysis. In P. W. Buys & M. Oberholzer (Eds), *Business Research* (pp. 61–76). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-9479-1_4
 27. Xie, T., & Ding, J. (2020). Forecasting with Multiple Seasonality. *2020 IEEE International Conference on Big Data (Big Data)*, 240–245.

- <https://doi.org/10.1109/BigData50022.2020.9378072>
28. Yusof, U. K., Khalid, M. N. A., Hussain, A., & Shamsudin, H. (2021). Financial Time Series Forecasting Using Prophet. In F. Saeed, F. Mohammed, & A. Al-Nahari (Eds), *Innovative Systems for Intelligent Health Informatics* (Vol. 72, pp. 485–495). Springer International Publishing. https://doi.org/10.1007/978-3-030-70713-2_45
29. Zhang, R., & Chan, N. H. (2021). NONSTATIONARY LINEAR PROCESSES WITH INFINITE VARIANCE GARCH ERRORS. *Econometric Theory*, 37(5), 892–925. <https://doi.org/10.1017/S0266466620000377>
30. Žunić, E., Korjenić, K., Hodžić, K., & Đonko, D. (2020). Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-world Data. *International Journal of Computer Science and Information Technology*, 12(2), 23–36. <https://doi.org/10.5121/ijcsit.2020.12203>