e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707203 |

Comparative Study of Indian Stock Market using Different Forecasting Techniques

Anantha Krishna Thalithya, Dr. H. K. Madhu

PG Student, Department of MCA, Bangalore Institute of Technology, Karnataka, India

Associate Professor, Department of MCA, Bangalore Institute of Technology, Karnataka, India

ABSTRACT: Predicting the values of stocks is a critical challenge in financial analysis and trading, driven by the need to forecast future changes in price and optimize investment strategies. This abstract explores the integration of the AutoRegressive Integrated Moving Average (ARIMA) model with moving average techniques to enhance stock market forecasting accuracy. The ARIMA model, a widely utilized time series forecasting method, captures the underlying patterns and trends in in stock price data from the past through its autoregressive, differencing, and moving average components. By identifying temporal dependencies and incorporating past observations, ARIMA provides a robust framework for predicting future price movements. However, while ARIMA excels in trend analysis, it may not fully capture market noise or sudden shifts in trends.

To complement ARIMA, moving average techniques—specifically the moving average crossover strategy—offer a practical approach to identifying buy and sell signals based on the intersection of short-term and long-term moving averages. This technique helps to detect potential market trends and reversals, providing actionable trading signals. When combined with ARIMA forecasts, the moving average crossover strategy enhances the decision making process by leveraging both statistical predictions and technical analysis signals.

I.INTRODUCTION

The stock market is a dynamic and integral component of the global economy, where securities such as stocks, bonds, and other financial instruments are bought and sold. It serves as a platform for companies to raise capital by offering shares to investors, and for investors to purchase ownership in these companies with the expectation of earning returns. Stock markets operate through exchanges like the National Stock Exchange (NSE) and Bombay Stock Exchange (BSE) in India, where transactions are facilitated in a regulated and transparent manner.

One of the primary advantages of the equity market is its potential for substantial returns on investment. Historically, stock markets have provided higher returns compared to other investment vehicles such as bonds or savings accounts. This potential for growth attracts individual and institutional investors looking to build wealth over time. Additionally, the stock market offers liquidity, allowing investors to easily buy and sell shares, making it a versatile choice for individuals who may need to access their funds quickly.

However, the stock market is not without its drawbacks. One significant disadvantage is its inherent volatility. Stock prices can fluctuate widely due to factors such as economic indicators, political events, and company-specific news. This volatility can lead to substantial financial losses, especially for individuals who are not well versed in market dynamics. Additionally, the stock market can be influenced by speculative behaviour, where investors make decisions based on short-term trends rather than long-term fundamentals, which can further contribute to market instability.

II. LITERATURE REVIEW

The paper **"A Comparative Study and Analysis of Time Series Forecasting Techniques" by Srihari Athiyarath, Mousumi Paul, and Srivatsa Krishnaswamy[1]**, published in SN Computer Science, explores different techniques for time series forecasting, particularly focusing on their applicability and effectiveness in different scenarios.

The study provides an extensive review of traditional and modern forecasting techniques, including statistical models like ARIMA (AutoRegressive Integrated Moving Average) and more sophisticated methods of machine learning such as Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks. The authors compare these methods based on their accuracy, computational complexity, and the types of data they handle best. For instance,

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707203 |

ARIMA is highlighted for its robustness in linear data patterns, whereas ANNs and LSTMs are acknowledged for their capacity to model complex, non-linear relationships and capture long-term dependencies in data.

In conclusion, the comparative analysis in the paper underscores the evolving nature of time series forecasting, advocating for a hybrid approach that leverages the strengths of both statistical and machine learning methods to achieve the most accurate and reliable predictions. This comprehensive examination helps practitioners and researchers decide using knowledge regarding which forecasting techniques to employ in various scenarios.

Dr. Vincent Cho's paper, "A Comparison of Three Different Approaches to Tourist Arrival Forecasting,"[2] evaluates three time-series forecasting techniques: exponential smoothing, univariate ARIMA (AutoRegressive Integrated Moving Average), and Elman's Model of Artificial Neural Networks (ANN). The study aims to determine the best technique for forecasting tourist arrivals. The results show that even though all three techniques have their strengths, Elman's ANN model outperforms the others With regard to precision and handling complex, non-linear data patterns. However, ARIMA remains valuable for simpler, linear data, and exponential smoothing is noted for its simplicity and efficiency in certain contexts. The study underscores the importance of selecting the appropriate forecasting method based on specific data characteristics and forecasting requirements.

The paper **"stock market forecasting technique using the ARIMA model" by Bijesh Dhyani, Manish Kumar, Poonam Verma, and Abhisheik Jain**[**3**] proposed a stock market forecasting model using the AutoRegressive Integrated Moving Average (ARIMA) technique, which is a widely used statistical technique for time series forecasting. The ARIMA model was applied to the daily closing prices of the S&P BSE SENSEX index, an extensively tracked stock market index in India. The model was developed and verified with the use of historical data from April 2014 to March 2019.

The findings indicated that the ARIMA model performed well in predicting future stock prices, with a high accuracy rate and minimal error. The authors also compared the ARIMA model's performance in comparison to other forecasts techniques, such as Exponential Smoothing (ES) and Moving Average (MA), and found that ARIMA outperformed these methods.

The study **"Forecasting stock market prices using mixed ARIMA model: a case study of Indian pharmaceutical companies" by Bharat Kumar Meher, Iqbal Thonse Hawaldar, Cristi Spulbar, and Ramona Birau**[**4**] focuses on the application of mixed ARIMA models to predict stock prices within the Indian pharmaceutical sector. The authors aim to enhance forecasting accuracy by combining multiple ARIMA models, thereby leveraging the strengths of each to capture the complexities of stock market behavior.

ARIMA models are well known for having the capacity to analyze and predict time series data by identifying underlying patterns through autoregression, differencing, and moving averages. The mixed ARIMA approach used in this study integrates various ARIMA configurations to better capture both short-term fluctuations and long-term trends in stock prices. By applying this technique to Indian pharmaceutical companies, the authors demonstrate its efficacy in providing more accurate and reliable forecasts compared to single ARIMA models. Their results imply that a combination of ARIMA models can serve as a valuable tool for investors and analysts in the pharmaceutical sector, aiding in more informed decision-making and strategic planning.

In the study **"Time Series Data Mining: Comparative Study of ARIMA and Prophet Methods for Forecasting Closing Prices of Myanmar Stock Exchange," Wint Nyein Chan from the Faculty of Information Science at the University of Computer Studies (Hpaan)**[**5**] explores the effectiveness of ARIMA and Prophet models in predicting stock prices. The research aims to identify which method offers superior accuracy in estimating the final prices of stocks listed on the Myanmar Stock Exchange.

ARIMA (AutoRegressive Integrated Moving Average) is a well-established statistical model that excels in capturing linear patterns and trends in time series data through autoregression, differencing, and moving averages. Conversely, Prophet, developed by Facebook, is designed to manage data in time series that exhibit seasonality and trends, and is particularly robust to missing data and outliers. By comparing these two methods, the study finds that while ARIMA provides reliable forecasts for linear trends, Prophet's flexibility in accommodating seasonality and handling irregularities makes it a strong contender for more complex and noisy datasets. The findings propose that the choice between ARIMA and Prophet should depend on the specific characteristics of the time series data, with Prophet offering advantages in scenarios involving seasonal and irregular patterns.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707203 |

III.METHODOLOGY OF PROPOSED SYSTEM

The proposed system tries to give the reference to the user by generating the buy and sell signal which was not present in the above existing system. Along with this it tries to forecast the stock market using the ARIMA model which tries to boost the confidence of the user and make an well informed logical decision based upon this.

Data preprocessing: Data preprocessing is a critical step in the data science workflow where raw data is transformed into a clean and structured format that is suitable for analysis. It typically involves several steps, including:

Data Cleaning: Removing or correcting any errors or inconsistencies in the data. This may involve handling missing values, correcting typos, or dealing with outliers.

Data Integration: Combining data from multiple sources if necessary, to create a unified dataset.

Data Transformation: Converting the data into a suitable format. This could include normalization (scaling data to a standard range), encoding categorical variables, or transforming variables to meet modeling assumptions.

Feature Selection/Engineering: Selecting the most relevant features (variables) for analysis or creating new features that may be more informative for modeling.

Dimensionality Reduction: Techniques like PCA (Principal Component Analysis) or feature extraction methods to reduce the number of variables in the dataset while preserving important information.

Data Splitting: Splitting the dataset into training and test sets for model training and evaluation. The goal of data preprocessing is to prepare the data in a way that maximizes the performance and reliability of the machine learning models that will be applied to it. Each step in preprocessing is aimed at ensuring that the data is accurate, complete, and ready for analysis.

Moving Average calculation: Calculating a moving average is a common technique used in data analysis and time series forecasting. It helps smooth out short-term fluctuations and highlights longer-term trends or cycles in data.

MODEL BUILDING USING ARIMA

ARIMA (AutoRegressive Integrated Moving Average) is a popular time series forecasting model that combines autoregressive (AR) and moving average (MA) components with differencing (integration) to handle non-stationary data. The model is denoted by **ARIMA(p,d,q)** where:

p: Number of lag observations included in the model (AR order).

d: Number of times that the raw observations are differenced (I order).

q: Size of the moving average window (MA order).

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707203 |

IV. RESULT

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                  Close   No. Observations:                 5968
Model:                 ARIMA(5, 1, 5)   Log Likelihood              -35234.424
Date:                Fri, 19 Jul 2024   AIC                          70490.848
Time:                        12:04:03   BIC                          70564.482
Sample:                             0   HQIC                         70516.426
                               - 5968
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.9336      0.090    -10.397      0.000      -1.110      -0.758
ar.L2         -0.1215      0.123     -0.988      0.323      -0.362       0.120
ar.L3          0.2933      0.116      2.529      0.011       0.066       0.521
ar.L4          0.5274      0.115      4.597      0.000       0.303       0.752
ar.L5          0.6423      0.053     12.136      0.000       0.539       0.746
ma.L1          0.9867      0.091     10.792      0.000       0.807       1.166
ma.L2          0.1574      0.130      1.208      0.227      -0.098       0.413
ma.L3         -0.3055      0.122     -2.514      0.012      -0.544      -0.067
ma.L4         -0.5316      0.123     -4.326      0.000      -0.772      -0.291
ma.L5         -0.6178      0.057    -10.906      0.000      -0.729      -0.507
sigma2      7881.5736     56.654    139.117      0.000    7770.533    7992.614
===================================================================================
Ljung-Box (L1) (Q):                   1.34   Jarque-Bera (JB):             51593.01
Prob(Q):                              0.25   Prob(JB):                         0.00
Heteroskedasticity (H):              12.17   Skew:                            -0.72
Prob(H) (two-sided):                  0.00   Kurtosis:                        17.33
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

Fig 1 ARIMA result

The table presents the results of a SARIMAX model fitted to a time series dataset with the dependent variable "Close." The model is specified as ARIMA(5, 1, 5), indicating five autoregressive (AR) terms, one differencing (I) term, and five moving average (MA) terms. The coefficients, standard errors, z-values, p-values, and confidence intervals for each AR and MA term are listed. Significant terms (with pvalues < 0.05) include ar.L1, ar.L3, ar.L4, ar.L5, ma.L1, ma.L3, ma.L4, and ma.L5. The model's fit statistics include the Log Likelihood (-35234.424), AIC (70490.848), BIC (70564.482), and HQIC (70516.426). Additionally, diagnostic tests like the Ljung-Box test, Jarque-Bera test, and heteroskedasticity test are provided, indicating model adequacy and potential issues with residuals.

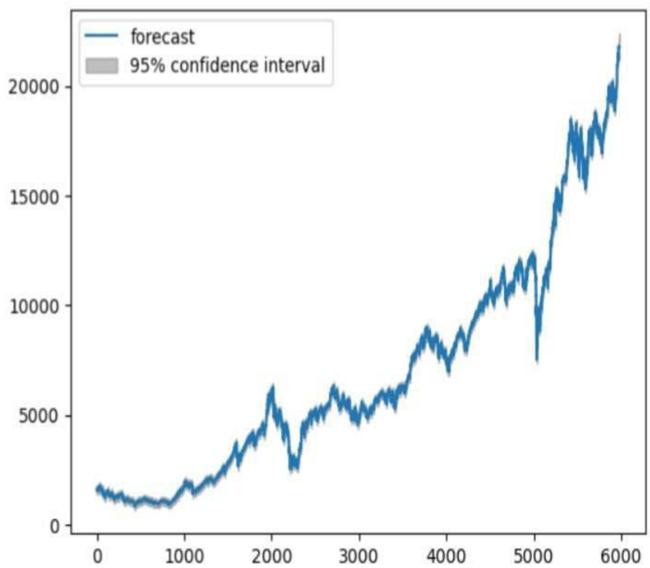

Fig 2 Forecasted graph

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707203 |

An "Actual vs Fitted" plot is used to assess the accuracy of a statistical or machine learning model by comparing the observed (actual) values to the model's predicted (fitted) values.

Fig 3 Generating the signals

Here we are generating the buy and sell signal on the basis of the moving average crossovers. Whenever we are getting an bullish crossover we get an green triangle as the buy signal and when there is an bearish crossover we get an red inverted triangle as the sell signal. The accuracy for the Moving Average crossover strategy is 76.92 %.

V.CONCLUSION AND FUTURE WORK

Integrating the ARIMA (AutoRegressive Integrated Moving Average) model with a moving average crossover strategy can offer a robust framework for forecasting and trading in financial markets. The ARIMA model excels in capturing the underlying patterns and trends in time series data, providing accurate forecasts for future price movements. When these forecasts are combined with a moving average crossover strategy, traders can enhance their decision-making process by relying on both statistical forecasting and technical analysis signals.

The moving average crossover strategy, a widely used technical analysis tool, identifies buy and sell signals based on the crossing of short-term and long-term moving averages. This method helps in recognizing potential market trends and reversals. By applying the ARIMA model to generate forecasts, traders can anticipate future price movements and use these predictions to inform the moving average crossover signals. This combination can potentially increase the strategy's effectiveness, allowing for more timely and accurate trade entries and exits.

To further enhance the effectiveness of combining the ARIMA model with the moving average crossover strategy, incorporating machine learning techniques and advanced analytics can be a significant step forward. For instance, integrating neural networks or support vector machines

(SVMs) to complement ARIMA's predictions can help capture non-linear patterns and complex dependencies in the data that traditional ARIMA models might miss. Additionally, adaptive moving averages, which adjust their parameters based on market volatility and trends, can improve the responsiveness and accuracy of crossover signals, reducing the risk of false positives in volatile markets.

REFERENCES

[1].A COMPARISON OF THREE DIFFERENT APPROACHES TO TOURIST ARRIVAL FORECASTING REF. NO. PQ-4556 By Dr. Vincent Cho Department of Management The Hong Kong Polytechnic University

[2].Improved Analysis of Stock Market Prediction (ARIMA-LSTM-SMP) Asha Ashok, Prathibhamol C.P Department of Computer Science and Engineering Amrita Vishwa Vidyapeetham Amritapuri, India

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707203 |

[3].Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison UPPALA MEENA SIRISHA, MANJULA C. BELAVAGI , AND GIRIJA ATTIGERI Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal 576104, India Corresponding author: Manjula C. Belagavi (manjula.cb@manipal.edu)

[4].Stock Market Forecasting Technique using Arima Model Bijesh Dhyani, Manish Kumar, Poonam Verma, Abhisheik Jain

[5].Time Series Forecasting using LSTM and ARIMA Khulood Albeladi, Bassam Zafar, Ahmed Mueen Department of Information Systems-Faculty of Computing & Information Technology, King Abdulaziz University, Jeddah, KSA

[6].Time Series Data Mining: Comparative Study of ARIMA and Prophet Methods for Forecasting Closing Prices of Myanmar Stock Exchange Wint Nyein Chan Faculty of Information Science, University of Computer Studies (Hpa-an)

[7].A. A. Adebiyi, A. O. Adewumi, C. K. Ayo "Stock Price Prediction Using the ARIMA Model", UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, pp.106-112.

[8].Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "On the difficulty of training recurrent neural networks." International conference on machine learning. PMLR, 2013.

[9].Kirchgässner, Gebhard, Jürgen Wolters, and Uwe Hassler. Introduction to modern time series analysis. Springer Science & Business Media, 2012.

[10].S. Siami-Namini, N. Tavakoli, and A. Siami Namin, ''A comparison of ARIMA and LSTM in forecasting time series,'' in Proc. 17th IEEE Int. Conf. Mach. Learn.
Appl. (ICMLA), Dec. 2018, pp. 1394–1401

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com