International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

USER PERCEPTION TOWARDS STOCK PRICE PREDICTION USING LINEAR REGRESSION MODEL

Submitted By

NIKHIL KUMAR

22042010474

Galgotias University

ABSTRACT

Stock market prediction is a critical area of research with profound implications for investors, financial analysts, and policymakers. This study explores the efficacy of linear regression analysis within the domain of business analytics for forecasting stock market movements. Leveraging historical stock prices and relevant economic indicators, we employ a rigorous methodology to construct and evaluate predictive models using linear regression techniques. Technological advancement increases the study on stock and share market industry. Decision making is enhanced by various statistical and machine learning algorithms. Enormous research work has been concentrated on the feature prediction of stock prices based on historical prices and volume. Performance measures are analyzed in this work with S&P 500 Index using statistical methods in Python environment. Results obtained in this study are superior than the existing methods. The conventional methods for financial market analysis are based on linear regression. This paper focuses on the best independent variables to predict the closing value of the stock market. This study is used to determine specific factors which are providing the most impact on the prediction of the closing price.

INTRODUCTION

Trade in stock market deals the movement of money of a security or stock from a trader to a buyer. This require these two parties to have the same opinion on a price. Equities (stocks or shares) present a rights interest in a specific company. Stock market participants range from small individual stock investors to larger traders investors, who can be based wherever in the world, and may contain insurance companies or pension funds, banks and hedge funds. Their buy or sell orders may execute on their behalf by a stock exchange dealer. Stock trading volume includes the number of lots bought and sold which is express in daily basis . The more trading volume of a stock is higher, the more the stock is active. Trading volume is an appreciative to price patterns in practical testing and it's additional vital than stock price. Stock market contribution refers to the number of agents who buy and sell equity backed securities either directly or indirectly in a financial trade. Participants are normally subdivided into three distinct sectors; households, institutions, and foreign traders. Direct participation occur when any of the above entities buys or sells securities on its own behalf on a trade. Indirect participation happens when an institutional investor exchanges a stock on behalf of an individual or household. Indirect investment takes in the form of pooled investment accounts, retirement accounts, and other managed financial accounts

The stock market serves as a cornerstone of global economies, facilitating capital allocation, investment, and economic growth. Predicting stock market movements has long been a pursuit of investors, analysts, and researchers alike, driven by the desire to anticipate market trends and make informed investment decisions. In recent years, the advent of advanced data analytics techniques, coupled with the proliferation of vast datasets, has opened new avenues for enhancing the accuracy and reliability of stock market predictions.

LITERATURE REVIEW

The literature review dives into different ways people try to predict stock prices. Some folks use old-school methods like Box-Jenkins' time series analysis, which has been around forever and helps forecast how stocks might behave. Then there are newer tricks, like Neural Networks (NN), which are like computer brains learning from past stock movements, as seen in White's study on IBM stocks. But there's a catch—stock prices are tricky and can be super unpredictable, especially when things get nonlinear over time. So, researchers are mixing and matching methods to make better predictions, although it can get pretty complicated and might need some expert know-how.

In finance, other methods like Logistic Regression (LR) and Multivariate Discriminant Analysis (MDA) are still pretty popular. They crunch numbers to guess stuff like whether a company might default on its debts. And there's this cool new thing called binary classification, which is all about sorting stuff into two groups, like predicting if a company might go bankrupt. It's like using high-tech puzzles to understand the financial world better. Basically, this review shows how the finance world is always trying out new ways to predict the future of stocks, and it's a mix of old tricks and shiny new ones.

RESEARCH OBJECTIVE

1) To Assess the Efficacy of Linear Regression in Stock Market Prediction this objective involves evaluating the performance of linear regression models in forecasting stock market movements based on historical data and relevant economic indicators.
2) To Identify Key Predictive Factors this objective entails identifying and analyzing the key variables or features that significantly influence stock market trends and fluctuations.
3) By conducting rigorous validation and testing procedures, we aim to determine the reliability and stability of the models in real-world scenarios.
4) To Provide Insights for Investment Decision-Making which aims to translate the research findings into actionable insights for investors, financial analysts, and policymakers
5) By highlighting limitations, challenges, and opportunities encountered during the study, we aim to guide future research endeavors aimed at advancing predictive modelling techniques and enhancing decision-making processes in finance.

RESEARCH DESIGN AND METHODOLOGY

i. Research Design:

For this study, a mixed-methods research design is employed. It combines quantitative analysis of historical stock market data obtained from Kaggle with a qualitative survey conducted via Google Forms. The quantitative analysis will utilize descriptive statistics and linear regression techniques to examine the relationship between historical stock prices and future stock movements. The qualitative survey will gather insights and opinions from participants regarding their perceptions of stock market prediction and the usefulness of linear regression models in forecasting.

ii. Data Collection Method and Forms:

1. Data Collection Medium:

- Historical stock market data will be collected from Kaggle, a reputable platform for datasets and data science competitions. These datasets provide comprehensive information on historical stock prices, trading volumes, and other relevant variables.
- A Google Form questionnaire will be used for the survey component of the study. The questionnaire will be designed to gather qualitative insights from participants regarding their views on stock market prediction and the effectiveness of linear regression models.

2. Questionnaire:

- The Google Form questionnaire will consist of open-ended and closed-ended questions designed to explore participants' perceptions, opinions, and experiences related to stock market prediction and linear regression modeling.
- Questions will cover topics such as participants' level of confidence in stock market predictions, their familiarity with linear regression models, and their preferences for predictive modeling techniques.

3. Logic of Choice:

Kaggle is selected as the data collection platform for historical stock market data due to its reliability and accessibility. Google Forms is chosen for the survey component for its ease of use, flexibility in questionnaire design, and ability to collect responses efficiently.

iii. Sampling Design and Plan:

1. Target Population:

The target population for the survey includes individuals with an interest in finance, investing, or data science. Considering the clarification, the revised target population is 33 instead of 40.

2. Sampling Frame:

The sampling frame consists of individuals who have access to the internet and are willing to participate in the survey.

3. Sample Units Used:

The sample units are individual survey respondents who omplete the Google Form questionnaire.

4. Methods for Selecting Sample Units:

Convenience sampling will be utilized to select survey participants. The questionnaire will be shared with relevant online communities, social media groups, and professional networks.

5. Sample Size:

The sample size will be determined based on the desired level of statistical power and the feasibility of reaching a sufficient number of respondents.

6. Response Rate:

The response rate will be monitored throughout the survey period to assess the effectiveness of the recruitment strategy and ensure a representative sample. Since the revised target population is 33, the response rate will be calculated accordingly based on the number of completed responses received compared to this revised target.

iv. Fieldwork:

1. Conduct of Fieldwork:

The fieldwork for data collection will involve accessing historical stock market datasets from Kaggle and distributing the Google Form questionnaire to potential participants.

LIMITATION

I. .Linear regression assumes a linear relationship between independent and dependent variables. However, stock price movements may exhibit non-linear patterns, which can limit the accuracy of linear regression models in capturing complex market dynamics.

II. Linear regression models are sensitive to outliers, which can distort the relationship between variables and lead to biased parameter estimates. In the context of stock market data, outliers such as extreme price fluctuations or market anomalies may adversely impact the performance of linear regression models.

III. While linear regression models can provide valuable insights into historical price trends, their predictive power may be limited when it comes to forecasting future stock prices accurately. Factors such as sudden market shocks, changes in investor sentiment, and unforeseen events can introduce uncertainty and volatility, making it challenging to predict stock prices with high precision.

IV. Effective implementation of linear regression models for stock price prediction relies heavily on the availability and quality of data. Challenges such as missing data, data lags, and inaccuracies in financial data sources can undermine the reliability and robustness of the models, limiting their effectiveness in capturing underlying market trends.

V. Overfitting occurs when a model captures noise in the data rather than the underlying relationships, leading to poor generalization performance on unseen data. Linear regression models, particularly when applied to large datasets with many predictor variables, are susceptible to overfitting, which can result in inflated model performance metrics and unreliable predictions.

FINDINGS

The study findings suggest that while many individuals are familiar with using linear regression for stock market analysis, not all of them feel entirely confident in their comprehension of it. Although a majority have experimented with these models for predicting stock prices, there exists a broad spectrum in how frequently people engage in actual trading or investment activities, with some opting out altogether. It's intriguing to note that most individuals rely on trading platforms for market information, and while they generally trust in the predictive capabilities of linear regression models, they also acknowledge challenges such as overfitting and data quality issues.

Moreover, the study highlights the significance people place on company fundamentals when attempting to forecast stock prices, alongside their recognition of the complexities inherent in effectively utilizing linear regression. Despite these hurdles, there persists a prevailing belief that linear regression models outperform other prediction methods in stock market analysis. Looking forward, a substantial number of respondents express openness to integrating these models into their investment strategies, demonstrating a willingness to adapt and explore new approaches within the constantly evolving landscape of stock trading.

The findings encapsulate a blend of awareness, usage patterns, beliefs, and considerations surrounding linear regression models in stock market analysis. They underscore both the potential benefits of these models and the challenges that need to be addressed, indicating a nuanced perspective on their role in navigating the intricacies of stock market dynamics

CONCLUSION

In conclusion, the study sheds light on the perceptions, practices, and challenges associated with using linear regression models for stock market analysis. The findings reveal a varied landscape, with respondents demonstrating a moderate level of familiarity and confidence in linear regression techniques. While a majority have utilized these models for predicting stock prices, there exists a diverse range of trading frequencies among respondents, suggesting differing levels of engagement with the stock market. Trading platforms emerge as the primary source of stock market information, and there is a prevailing belief in the effectiveness of linear regression models for predicting stock prices, despite challenges such as overfitting and data quality issues.

Company fundamentals are perceived as the most important factor for predicting stock prices using linear regression, highlighting the significance of fundamental analysis in stock market analysis. Despite challenges, linear regression models are seen as outperforming other prediction methods, and many respondents express a willingness to incorporate them into their future investment strategies. Overall, the findings underscore the importance of education, data quality management, and continuous improvement in leveraging linear regression models effectively for stock market analysis. By addressing these factors, stakeholders can enhance decision-making processes and drive more informed investment strategies in the dynamic and complex world of the stock market.

RECOMMENDATIONS

There are several key recommendations to bolster the effectiveness and adoption of linear regression models in stock market analysis. Firstly, focusing on education and training initiatives can significantly enhance understanding and confidence levels among individuals regarding the application of linear regression techniques. By offering tailored workshops, online courses, and educational resources, participants can gain the necessary knowledge and skills to leverage linear regression effectively in predicting stock prices.

Secondly, addressing data quality issues is crucial to improving the reliability and accuracy of linear regression models. Investing in robust data collection, cleaning, and validation processes can help mitigate concerns related to data quality, thereby enhancing the credibility of the models' predictions. Moreover, emphasizing the importance of data integrity and transparency can foster trust and confidence among users in the reliability of the insights generated by linear regression models.

Lastly, fostering a culture of continuous improvement and collaborative learning can drive innovation and experimentation in the use of linear regression models for stock market analysis. Providing platforms for peer learning, knowledge sharing, and collaborative problem-solving can enable participants to exchange insights, experiences, and best practices, leading to the refinement and advancement of predictive techniques. By embracing a collaborative and iterative approach to learning and innovation, stakeholders can unlock the full potential of linear regression models in informing decision-making processes and driving investment strategies in the dynamic landscape of the stock market.

REFRENCES

- Seethalakshmi, Ramaswamy. (2018). Analysis of stock market predictor variables using linear regression. International Journal of Pure and Applied Mathematics. 119. 369-377.

LINKS:

- https://www.researchgate.net/publication/326253896_Analysis_of_stock_market_predictor_variables_using_linear_regression
- https://www.kaggle.com/datasets/akshayr009/adani-ports-stock-price
- https://www.kaggle.com/datasets/anishasinghania/reliance
- Box–Jenkins[1] , White [2,3,4] , Henry [5] , Chiang et al.[6] , Kim and Han [7] , Romahi and Shen [8] , [Lee, 2004][9] , Öğüt and Aktaş [2009][10] , Minand Jeong [2009][11]