Kurdish Studies
Apr 2023
Volume: 11, No: 2, pp. 5854-5862
ISSN: 2051-4883 (Print) | ISSN 2051-4891 (Online)
www.KurdishStudies.net

Received: May 2023 Accepted: June 2023
DOI: https://doi.org/10.58262/ks.v11i2.425

New Stocks Prediction Portfolio Optimization on Vn30: Machine Learning Method

Anh Phong, Nguyen[1], Truc Linh, Nguyen[2]

Abstract

This research applies automatic portfolio selection based on machine learning algorithms combined with the Markowitz portfolio optimization method. The experiment was conducted for stocks on the VN30 index, the search results selected stocks that showed optimized value between risk and rate of return. This result can be applied in online investment consulting or building an app for investment consulting based on actual real data on any Stock exchanges.

Keywords: *Portfolio optimization, VN30, Stock Prediction, Machine Learning*

JEL Classifications: C61; E22; E47

1.Introduction

Vietnam's stock market has managed to remain robust and dynamic, despite the adverse effects of the Covid19 pandemic. The market has been experiencing consistent trading sessions, and the number of new accounts has been maintained at a rate of over 100,000 per month in 2021, indicating an enthusiastic and lively phase of development. It is common for new investors who are entering the market to lack selectivity in their investments, lack experience, and be influenced by crowd psychology, resulting in investments based on personal feelings and preferences rather than sound analysis, informed decisions, proper orientation, and research.

Studies and literature on this topic have indicated that technical analysis can serve as a valuable tool for analyzing stock market transactions. Several recent studies have also produced evidence of trades based on technical analysis (Du et al, 2018; Yang, 2021). Test the performance of trading models using synthetic machine learning algorithms, based on inputs from technical analysis. From the ability to learn and process powerful big data in the financial field of machine learning, combined with the theoretical basis of technical analysis, the fluctuations in the stock market are always considered complex and stem from many different causes, the seemingly impossible forecast problem will become clearly possible. Theoretically, this combination is expected to improve the role of a reference channel through analyzing and making the most of historical data. Some highlights of the application of machine learning in the field of stock market trend forecasting include the ability to "learn" from past and unlimited data, contrary to the limitations of human thinking, eliminating the problem of distributional

[1] Assosiate Prof, PhD (Economics), Lecturer, University of Economics and Law, Ho Chi Minh City, Vietnam, (Corresponding Author) Email: phongna@uel.edu.vn, Orcid Id: 0000-0003-2253-2051
[2] Graduate (Fintech), Vietnam National University, Ho Chi Minh City, Vietnam, Email: linhnt19414c@st.uel.edu.vn

psychology. Machine learning also looks at even the smallest price changes, updates and compares existing data with past data, aggregates and assists in making effective investment decisions.

Recognizing the advantages of using machine learning and technical analysis methods and recognizing the importance of having a tool that can assist investors in making informed investment decisions, the individual responsible for this research has formulated a research idea that involves creating algorithms and conducting research to achieve this goal.

2.Literature Reviews

Markowitz (1952) introduced the concept of portfolio optimization through the use of statistical analysis. Markowitz's Modern Portfolio Theory (MPT) provided a framework for investors to quantify the risks and returns associated with investing in a portfolio of assets. MPT assumes that investors are rational and risk-averse, and that they seek to maximize their expected returns while minimizing their risk exposure. Despite its contributions to the field of finance, Markowitz's paper has some limitations. One of the main limitations is that MPT assumes that asset returns are normally distributed, which may not always be the case in practice. Additionally, MPT does not account for the potential impact of extreme events or black swan events, which can have a significant impact on portfolio returns.

The use of deep learning LSTM models and the Markowitz Mean-Variance portfolio optimization method for portfolio optimization has been explored in several recent papers. For example, Zhang et al (2020) utilized a deep learning LSTM model to predict future stock prices and optimize portfolio allocation using the Markowitz Mean-Variance method. Similarly, Sen (2021) and Chaweewanchon et al (2022) also employed a machine learning and deep learning LSTM model for stock price prediction and optimized portfolio allocation. These papers share several commonalities in their approaches to portfolio optimization. They all utilize advanced machine learning and deep learning techniques to analyze complex financial data and predict future market trends. By combining these techniques with the Markowitz Mean-Variance method, they are able to optimize portfolio allocation and manage risk effectively. However, there are also limitations to these approaches, such as the need for high-quality data, the risk of overfitting, and the need for continuous refinement to adapt to changing market conditions. Overall, these papers demonstrate the potential for utilizing machine learning and deep learning techniques in portfolio optimization and suggest promising avenues for future research.

In the field of portfolio optimization, the use of efficient frontier with machine learning and deep learning techniques has been studied in various papers. Chen et al (2021) proposed the Efficient Frontier theory, which is a fundamental concept in modern portfolio theory. The paper discussed the optimization of the portfolio by constructing the efficient frontier and selecting the optimal portfolio from it. Similarly, Sharma and Salian (2020) demonstrated the use of efficient frontier in Python to optimize the portfolio using various machine learning techniques. Moreover, Raut and Singh (2021) presented an approach to optimize the portfolio using efficient frontier and machine learning techniques, specifically using Python. The research analyzed the performance of the portfolio optimization models using different methods, including principal component analysis, linear regression, and random forest. In contrast, Thakur and Singh (2021) used modern portfolio theory and Python to optimize the portfolio by minimizing the risk and maximizing the returns.

Although these papers differ in their techniques, they all share a similar goal of optimizing the portfolio by minimizing risk and maximizing returns. The use of efficient frontier and machine learning techniques allows for a more accurate and effective portfolio optimization strategy by analyzing complex financial data. However, the limitations of these techniques include the need for high-quality data to ensure accuracy and the risk of overfitting, which can lead to incorrect investment decisions. The continuous updating and refinement of the deep learning models and efficient frontier approach are also necessary to adapt to the dynamic financial market.

Optimizing investment portfolios in specific stock markets or equivalent funds with the objective of maximizing profits and/or minimizing risks is a common research topic in finance. The advantages of this approach include better risk management and the potential for higher returns. For instance, Syamni and Almuntasir (2013) using the single index model can provide a more optimal portfolio allocation strategy for the Indonesian stock market. Meanwhile, Kim and Park (2019) demonstrate how combining the efficient frontier and asset allocation strategies can lead to higher returns in the Korean stock market. However, a limitation of this approach is that it relies heavily on historical data and assumptions about future market conditions, which can introduce errors in the analysis. Furthermore, Nguyen and Truong (2017) point out that the lack of transparency in the Hanoi Stock Exchange can pose challenges for investors seeking to optimize their portfolios. Therefore, while this approach has its benefits, investors should exercise caution and be aware of its limitations when using it to optimize their investment portfolios.

3.Methodology

3.1.Data Setting

This research utilizes daily price data (including opening price, closing price, highest price, lowest price, and trading volume) of stocks traded on stock exchanges in Vietnam from the time these stocks began trading to the present day through the API of the vnstock library. Among the stocks in the Vietnam stock market, VN30 stocks are the ones with the largest market capitalization and highest liquidity, specifically accounting for 80% of the total market capitalization value and 60% of the total trading value in the market, going through many market fluctuations, making historical price data of these stocks valuable in research, analysis, or even predicting future price movements. Moreover, VN30 stocks are always closely monitored and carefully considered by the Ho Chi Minh City Stock Exchange (VN30), with annual stock adjustments. The VN30 stocks are mostly held by leading industry companies, leading in terms of revenue, profit, and brand, with ineffective companies or those with reduced liquidity being completely removed from the VN30. Therefore, these stocks have a relatively safe degree, with few issues such as: vague information, fraud, stock price manipulation, and strong influence by other stocks. Thus, the author chose VN30 stocks as a premise for running experiments and creating a foundation for building a web platform.

3.2.Model Training

In the field of portfolio optimization, the use of efficient frontier with machine learning and deep learning techniques has been studied in various papers. Sen et al (2021) proposed the Efficient Frontier theory, which is a fundamental concept in modern portfolio theory. Sharma and Salian (2020) demonstrated the use of efficient frontier in Python to optimize the portfolio using various machine learning techniques. Moreover, Thakur and Singh (2021) used modern portfolio theory and Python to optimize the portfolio by minimizing the risk and maximizing the returns.

In this research, we apply Efficient Frontier method is to set points that are investment portfolios with the highest expected return for a given level of risk or the lowest risk for a certain level of return. Portfolios that are below the Efficient Frontier are not optimal because they do not provide sufficient risk-reward. The portfolios located to the right of the Efficient Frontier are near optimal because they have a higher level of risk for a defined rate of return. We also use the auto generate algorithm in the Efficient Frontier library in Python to generate a desired number of investment portfolios to gain a better understanding of the optimal investment portfolio.

4.Results and Discussion

The process of obtaining data is just the first step in building a stock investment portfolio. However, having a small data table of one stock alone is not sufficient to proceed with optimizing the weight of investment in a large portfolio consisting of multiple shares. In order to achieve this, there are several ideas and implemented them in the following manner: Step 1: Combine 30 stocks into a big data table (only take the closing price); Step 2: Calculate the average profit rate by day; Step 3: Choose the top 5 stocks with the highest average daily return. For completeness, Annual Volatility is a more common measure that can be calculated simply by multiplying the daily volatility by the square root of the number of trading days in a year (252 days)

Daily volatility is the average difference between profits on a given day and average profits over a period of time. Mathematically, it is just the standard deviation of daily profit. Volatility is one of the measures of risk in which highly volatile investments can carry greater risk. In which, 5 stocks have the following results:

BCM	0.029989
BID	0.024883
ACB	0.021176
FPT	0.017772
VIB	0.022332

Figure 1. Daily Volatility (Source: Author's Calculations).

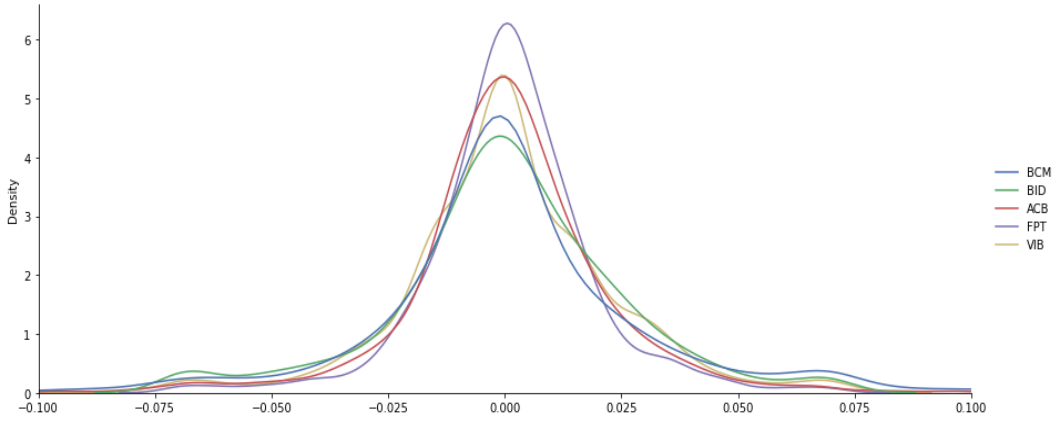

Figure 1. Density Chart of Top 5 Stocks.
Source: Author's Calculations.

Figure 2 can be observed that there is low volatility in FPT, whereas BID has the highest volatility. By comparing the density charts of the daily returns as shown in Figure 3, it is noticeable that the curve of FPT is narrower with a higher peak, indicating a lower level of volatility. On the other hand, BID has a wider curve which implies a higher standard deviation and thus, higher volatility. Also according to this volatility ranked in the order of FPT to VIB to ACB to BCM and finally BID.

Date	BCM	BID	ACB	FPT	VIB
2018-02-22	-0.008368	-0.016062	-0.006681	-0.016809	-0.013158
2018-02-23	-0.092827	0.008162	0.011261	0.035944	-0.013333
2018-02-26	-0.055814	0.026939	0.013452	-0.001650	0.000000
2018-02-27	0.009852	0.001326	0.024261	-0.009958	0.013514
2018-02-28	0.004878	0.022296	0.012959	0.000000	0.066667
...	...	...	...	...	...
2023-02-17	-0.001176	0.017778	0.000000	0.000000	0.028708
2023-02-20	0.012956	0.030568	0.040323	0.014706	0.023256
2023-02-21	-0.005814	-0.020127	-0.011628	0.000000	-0.009091
2023-02-22	-0.011696	-0.027027	-0.025490	-0.012077	-0.036697
2023-02-23	0.001183	0.012222	0.022133	-0.004890	0.023810

Figure 3: Daily Profitability of Top5 Stocks.
Source: Author's Calculations.

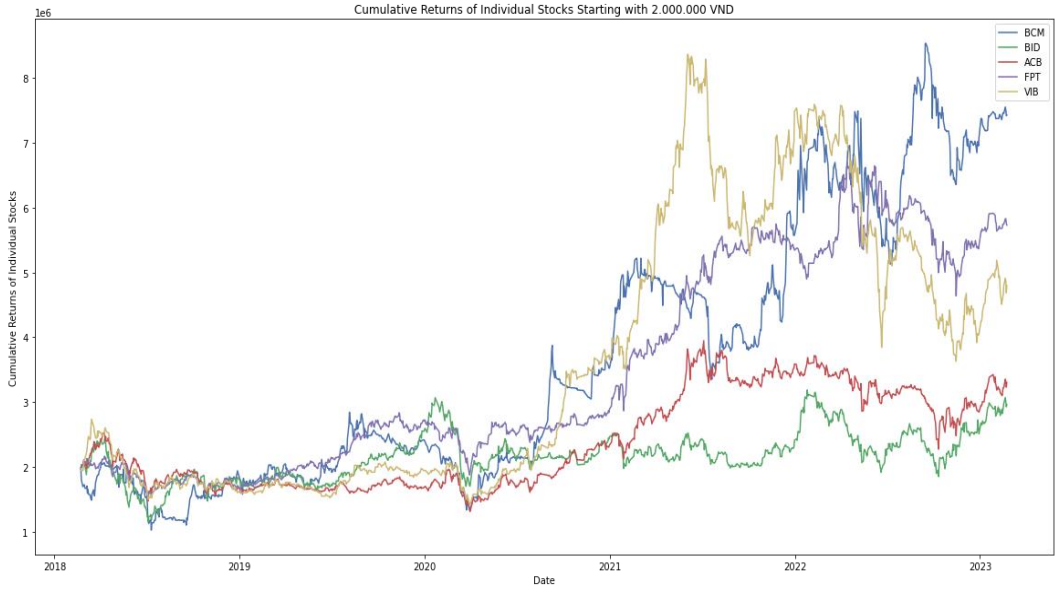

Figure 4. Chart of Profit Growth If the Top 2,000,000 VND Per Share
Source: Author's calculations

Based on Figure 4, it is evident that investing 2,000,000 VND in any of the 5 chosen stocks will result in a profit of around 15,000,000 VND to 18,000,000 VND, almost every time in the current market. Investing all the money in a single stock is a highly risky approach that no one would dare to take. This is because of the common saying in business, "Don't put all your eggs in one basket." Hence, it would be quite absurd to invest a substantial amount of money like 2,000,000,000 VND or 20,000,000,000 VND in just one stock.

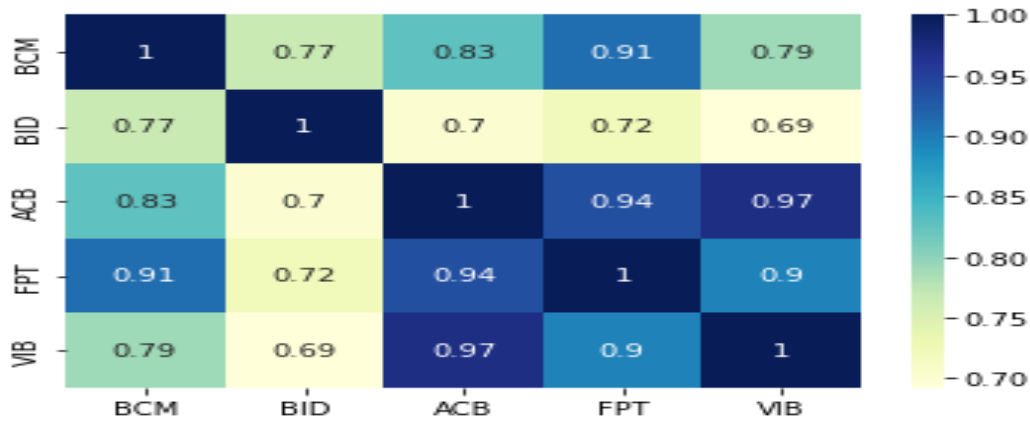

Figure 5. Chart Showing the Correlation Between 5 Stocks.
Source: Author's Calculations.

From Figure 5, it is evident that there is a strong positive correlation among all the stocks as indicated by the coefficients of 0.65. This can be explained when this is also the stage when Vietnam's economy enters the development stage in all aspects.

An Efficient Frontier is the set of optimal portfolios that yield the highest expected return for a given risk level (volatility) or the lowest risk (volatility) for a given expected return level. It is represented by a line on the Return vs Volatility chart. Portfolios with a maximum Sharpe ratio are above the Efficient Frontier. In order to provide visual representation, 1000 portfolios of selected stocks with random weights were created and their returns and volatility were plotted by running the code. The Efficient Frontier and the maximum Sharpe ratio portfolio are also plotted on the chart.

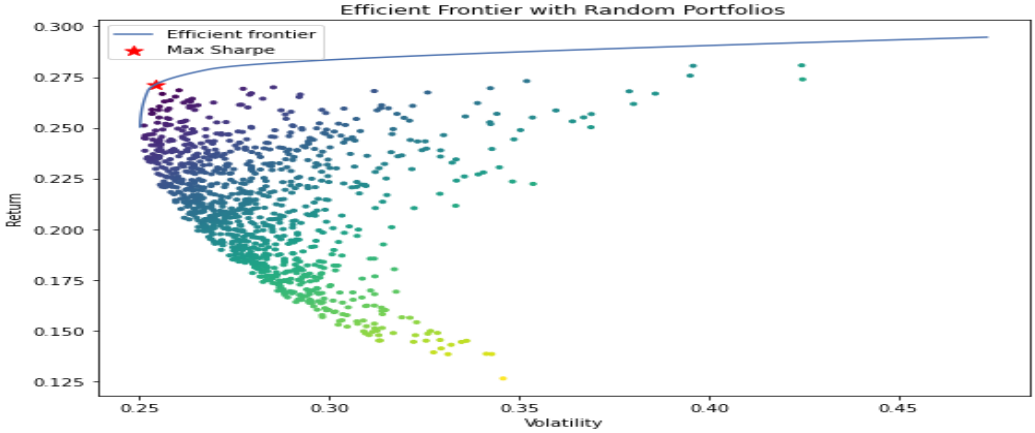

Figure 6. The Efficient Frontier and 1000 Portfolios.
Source: Author's Calculations

From the above graph, it can be observed that the efficient frontier line coincides with the boundary line. This implies that it is not possible to have a portfolio on the boundary line since a portfolio with the same volatility as the one on the boundary line but with a higher expected return would be available. The boundary line in the chart above acts as an effective boundary line, meaning that the portfolio cannot be located on the line itself. If it were, the portfolio would have the same volatility as the one located on the boundary line, but with a higher expected return. Similarly, the portfolio cannot be located to the left of the boundary line since this would result in portfolios with the same expected return as the one located on the boundary line, but with lower volatility. The results show that building a portfolio based on the Markowitz portfolio combined with machine learning algorithms gives better results than other methods (Wang et al, 2020). And the significance of stock preselection is similar to the conclusions by Wang et al (2020), Ta et al (2020), and Chen et al (2021).

	weights
BCM	0.19525
BID	0.00000
ACB	0.00000
FPT	0.64120
VIB	0.16355

Figure 7. Investment Weighting in the Portfolio.
Source: Author's Calculations

Here, it is clear that investment weighting is strictly studied on the side of FPT shares, so it should be noted that certain weights tend to be dominated by high-performance stocks (high profits with relatively stable volatility), while some weights tend to be zero. The maximum Sharpe portfolio optimization method has a weakness where the portfolio may not have enough diversification across various stocks or industries according to our preference. Also, it sounds cliché, past performance is not an indication of future results. Expected returns and volatility are calculated based on historical fluctuations in stock prices, but it's important to note that future returns and volatility may vary significantly.

```
Expected annual return: 27.13%
Annual volatility: 25.43%
Sharpe ratio: 0.99
```

Figure 8. Expected Indicators of the Portfolio
Source: Author's Calculations

The expected annual growth of this optimal portfolio is 29.83% with an annual volatility of 25.46%

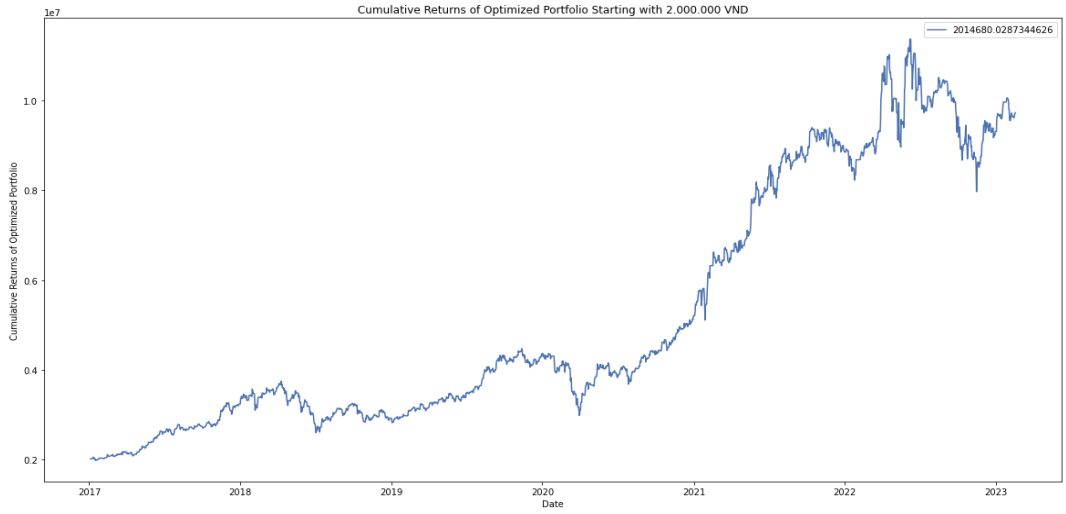

Figure 9. Amount Growth Chart If Investing by Portfolio
Source: Author's Calculations

Based on Figure 9, investing in the portfolio presented in the research article would yield a profit of approximately 18,146,800 VND if the initial investment is 2,000,000 VND. This indicates that the profit is substantial and can be further increased with a larger initial investment. In general, if compared to the results of guest studies, this can also be called a research success.

5.Conclusion and Recommendation

The aim of the research is to create a model that maximizes investment portfolios based on the theory of the boundary margin in the Vietnamese stock market. The model will provide investment suggestions to help investors decide on buying or selling options, thus increasing investment efficiency. To build the model, the authors used Python support packages, theories of the boundary margin, sharp ratio, risk-free rate, and incorporated them into the efficient frontier model. From this comprehensive model, the research yielded the following conclusions: First, it can be concluded that the stock prices of the stocks tend to increase from the past to the present based on the performance of the Vietnamese stock market in the short term, and investors should pay attention to the influence of past prices on the present. Second, the model for maximizing investment portfolios with the technical variables used in this research is suitable for novice investors who are new to securities and do not have the ability to analyze a lot. Finally, the predictive model can be seen as a tool to help investors refer to a variety of investment options that are selected depending on their individual characteristics and risk tolerance, and from there determine how to invest and earn profits for themselves.

Acknowledgements and Funding: This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number DS2022-34-03/HĐ-KHCN

Contribution: The authors contributed equally to this work.

Data Availability Statement: The dataset is available from the authors upon request.

References

Chaweewanchon, A. and Chaysiri, R. (2022) Markowitz Mean-Variance Portfolio Optimization with Predictive Stock Selection Using Machine Learning", IJFS, 10(3), pp. 1-19.

Chen, Wei, Haoyu Zhang, Mukesh Kumar Mehlawat, and Lifen Jia (2021). Mean–variance portfolio optimization using machine learning-based stock price prediction. Applied Soft Computing 100: 106943

Du, K., Han, L., & Wu, T. (2018) "Markowitz Mean-Variance Portfolio Optimization with Predictive Stock Selection Using Machine Learning"

Jong-Min Kim and Se-Hak Park (2019) "Portfolio Optimization by Using Efficient Frontier and Asset Allocation Strategies"

Markowitz, Harry. 1952. Portfolio selection*. The Journal of Finance 7: 77–91

Nguyen V.A and Truong T.N Chi (2017) "Stock Portfolio Optimization: A Case of the Hanoi Stock Exchange"

Quynh Anh Nguyen and Cong Thanh Nguyen (2017) "Stock portfolio optimization using Markowitz model and particle swarm optimization"

Sen, Jaydip & Dutta, Abhishek & Mehtab, Sidra. (2021). Stock Portfolio Optimization Using a Deep Learning LSTM Model - Presentation. 10.13140/RG.2.2.33046.47685/1.

Sharma and Shreyas Salian (2020) "Efficient Frontier: A Modern Portfolio Theory based approach using Python"

Syamni, Ghazali and Muntasir, Al (2013) Optimal portfolio analysis using single index Model in the lq-45 index in indonesia stock Exchange. In: International Conference, 29-30 Agustus 2013, UNISBANK, Semarang

Ta, Van-Dai, Chuan-Ming Liu, and Direselign A. Tadesse (2020). Portfolio optimization-based stock prediction using long-short term memory network in quantitative trading. Applied Sciences 10: 437

Thakur and Srishti Singh (2021) "Portfolio Optimization using Modern Portfolio Theory by Python"

Zhang, Zihao & Zohren, Stefan & Roberts, Stephen. (2020). Deep Learning for Portfolio Optimization. The Journal of Financial Data Science. 2. 8-20. 10.3905/jfds.2020.1.042.

Wang, Wuyu, Weizi Li, Ning Zhang, and Kecheng Liu (2020). Portfolio formation with preselection using deep learning from long-term financial data. Expert Systems with Applications 143: 113042

Yang, J. (2021) "Efficient Frontier Portfolio Optimization using Python"