DOI: 10.54254/2754-1169/90/20242006

Integrating LSTM Networks with Mean-Variance Optimization for Enhanced Portfolio Construction: An Empirical Study on UK Stock Market

Shanshan Liu[1,a,*]

[1]*Faculty of social sciences, University of Southampton, Southampton, The United Kingdom*
a. sl1y22@southamptonalumni.ac.uk
**corresponding author*

Abstract: Portfolio optimization has long been a central theme in finance. With ongoing advancements in machine learning, there is a significant opportunity to integrate predictive methods into portfolio optimization. This paper proposes leveraging Long Short-Term Memory (LSTM) networks alongside the established Mean-Variance (MV) optimization framework to construct optimal portfolios. These portfolios aim to help financial investors effectively manage and mitigate risk while maximizing returns. The study meticulously screened the leading stocks of 12 prominent UK companies listed on the London Stock Exchange (LSE), known for their influence and visibility. Initially, the study applies the LSTM networks to predict stock price volatility and integrates these predictions into the MV model to allocate portfolio weights effectively. To underscore the superiority of the proposed approach, the study compares cumulative returns from portfolios optimized for maximum and minimum variance Sharpe ratios with real data against the FTSE100 index over the same period. The results demonstrate that the portfolio optimized with the maximum Sharpe ratio significantly outperforms both the conservative minimum variance portfolio and the FTSE100 index. This empirical evidence underscores the practical value of integrating advanced machine learning techniques with established financial theory, enabling more innovative and efficient investment strategies to meet the evolving needs of financial investors.

Keywords: Portfolio Management, LSTM Networks, Mean-Variance optimization, Stock price forecasting.

1. Introduction

In contemporary society, investment portfolios are crucial for enhancing financial stability and promoting economic growth. Economic phenomena often have a significant tendency to move more or less in coherence in economies, resulting in times of comparatively high or low overall economic activity and strong positive connections between economic entities. However, the inherent diversification of investment portfolios can reduce such risks and enhance the stability and potential growth of assets. In today's unpredictable markets, where economic uncertainty and geopolitical events may have a substantial influence on financial stability, this diversification technique is becoming more and more necessary.

The portfolio theory proposed by Harry Markowitz in 1952 provides positive explanations and normative guidelines for designing strategies and constructing portfolios [1]. Since then, more researchers and scholars have relied on this theory to validate and innovate. Širůček and Křen selected U.S. stocks as a sample and successfully used the Markowitz portfolio theory to determine the most suitable method for constructing an optimal portfolio [2]. This paper empirically verifies that optimal portfolios compiled according to the MPT theory provide investors with the optimal rate of return for a given risk. Kolm et al. show that portfolios focus on the optimal allocation of assets as a means of maximizing return and minimizing risk, which requires investors to accurately grasp stock price forecasts to provide key inputs for portfolio optimization and to make the asset allocation decision more informed [3]. However, it is difficult to find models for financial market price prediction because financial asset prices are chaotic, dynamic, and nonlinear [4]. Dixon et al. have recently revealed the benefits of machine learning as a quantitative tool for portfolio optimization and its use in the definition of modeling and decision framing in the financial industry [5]. To create the ideal portfolios, the MV model uses historical data, which means it can only provide the ideal portfolios within the parameters of the data inputs. According to Henrique et al., by reviewing a large body of literature, it was found that researchers have been applying machine learning to make predictions including future returns and volatility for time series data [6]. Consequently, the LSTM networks were chosen as the modeling tool for stock price prediction in this study.

This research aims to develop a methodology for constructing investment portfolios. The methodology utilizes the LSTM networks for stock price prediction and combines mean-variance optimization methods in conjunction with Monte Carlo simulation to form an optimal portfolio. This approach aims to achieve investment weights that maximize the Sharpe ratio and minimize volatility. To evaluate the effectiveness of the model, the cumulative return of the portfolio is calculated and compared with the cumulative return of the FTSE100 index at the same time frame. The conclusion that the portfolio constructed in this study outperforms the index fund serves as an effective reference for investors.

The data for this study was sourced from Yahoo Finance, focusing on 12 leading companies listed on LSE across six main sectors: energy, financials, consumer goods, healthcare, technology, and telecommunications. The research process follows a rigorous and objective methodology, ensuring that all important indicators are visualized, and the conclusions drawn are both scientific and of reference value.

The article is organized as follows in its entirety. A thorough introduction and description of data gathering are provided in Part 2. Part 3 introduces the experimental process and briefly describes the various research methods and evaluation indicators used in the research process. In Part 4, the results of the experiment are reported in detail and the data from the experimental process are visualized. Finally, Part 5 discusses the main findings, theoretical realizations, and future works.

2. Data

The stock data for this study comes from Yahoo Finance, obtained using the Python package yfinance (https://finance.yahoo.com/). Yahoo Finance's data is sourced from public exchanges and financial markets and is widely used for automated trading, technical analysis, financial modeling, and market research. These data sources are often authoritative, so the daily stock prices obtained are reliable. The data covers the stock performance of UK companies listed on the LSE for the period March 1, 2022, to June 7, 2024. In addition, daily movements of the FTSE 100 index from March 1, 2024, to June 7, 2024, were also obtained. The FTSE 100 index is a crucial indicator for global investors to gauge the movement of European stock markets. It covers 80% of the market capitalization of the London Stock Exchange, making it a highly representative metric.

DOI: 10.54254/2754-1169/90/20242006

2.1. Data Screening

The dataset covers 12 selected companies in six key industries: energy, financials, consumer products, healthcare, technology, and telecommunications. Table 1 provides a detailed presentation. Diversification across industries reduces the risk of a single industry downturn and is more resilient to economic fluctuations. For example, if one sector performs poorly, investments in other sectors can help offset the losses, thereby protecting the stability of the overall portfolio. This protects the portfolio from extreme damage by managing exposure to individual assets, diversifying the portfolio's income streams, and enhancing its defense against market turbulence. Therefore, this study balances returns and provides stability by selecting a combination of cyclical and defensive sectors. A portfolio that encompasses multiple sectors of society provides a comprehensive understanding of the economy and helps in making better investment decisions.

Companies within an industry are selected based on their size, market capitalization, stock returns over the last year, and their visibility in their respective industries to ensure a comprehensive analysis of industry performance. The timeframe for data selection captures the complete market cycle and includes any significant economic events during this period that may have affected stock performance. The study enriches the empirical experience in this field and aims to provide valuable insights into sector resilience and portfolio strategies for the UK stock market by analyzing the share prices of these companies over a specific period.

Table 1: 12 Stocks Selected for Portfolio Optimization

Industry	Stock Symbol	Company
Energy	SHEL.L	Shell
Energy	CNA.L	Centrica
Finance	HSBA.L	HSBC Holdings
Finance	LLOY.L	Lloyds Banking Group
Consumer Goods	ULVR.L	Unilever
Consumer Goods	ABF.L	Associated British Foods
Health Care	AZN.L	AstraZeneca
Health Care	HIK.L	Hikma Pharmaceuticals
Technology	DARK.L	Darktrace
Technology	SGE.L	The Sage Group
Telecommunications	BT-A. L	BT Group
Telecommunications	VOD.L	Vodafone Group Public Limited Company

2.2. Data Application

In this paper, the study chooses ‘Adj Close’ as the daily price V_t^i and computes the daily returns of the i-th particular asset R_{it} on day t as follows:

$$R_{it} = \frac{V_t^i}{V_{t-1}^i} - 1 \tag{1}$$

"Adj Close" takes into account corporate behavior such as dividends, stock splits, and other events that may affect the stock price, thus ensuring that the historical price series reflects a continuous investment scenario. Choosing adjusted prices for stock portfolio calculations provides a more accurate, consistent, and comprehensive picture of stock performance [7].

DOI: 10.54254/2754-1169/90/20242006

3. Methods

This study employs a series of key steps, as illustrated in Figure 1, for stock price forecasting and portfolio optimization. Initially, stock price data ranging from March 1, 2022, to March 1, 2024, are selected and collected based on the criteria outlined in the Data section. Subsequently, the price fluctuations of the selected stocks are forecasted using LSTM networks. Once the predictions are complete, the optimal asset allocation weights are ascertained by integrating the mean-variance optimization (MV) technique with Monte Carlo simulation. Furthermore, the portfolio's cumulative return and the fluctuation in the FTSE 100 index's return are analyzed using actual market share price data from March 1, 2024, to June 7, 2024. Ultimately, the effectiveness of the refined portfolio's performance is evaluated through a comparison of its earnings against the benchmark returns of the FTSE 100.

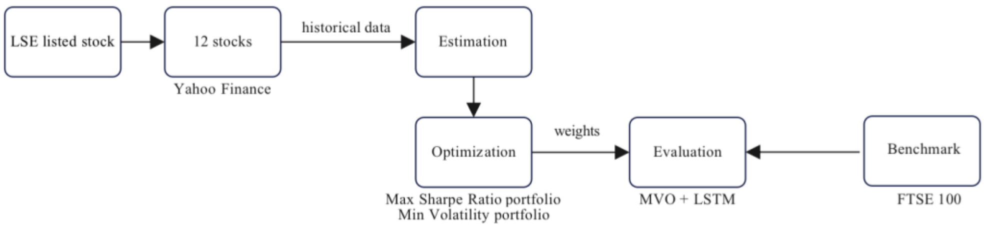

Figure 1: Flowchart of the Study

3.1. Modern Portfolio Theory

Markowitz's Mean-Variance (MV) framework is a foundational investment framework focused on selecting and constructing portfolios to maximize expected returns while minimizing investment risk. According to theory, diversification helps investors to achieve more stable and potentially higher returns than they would by investing in a single asset. This is accomplished by spreading their investments across a range of assets, thereby reducing the portfolio's overall risk [1].

3.2. Mean-Variance (MV) Model

The Mean-Variance (MV) model is a quantitative tool used in Modern Portfolio Theory (MPT) applications. It involves selecting an asset portfolio through mathematical optimization. In the realm of asset allocation, the primary objective is to guide investors toward achieving an optimal equilibrium between potential risks and anticipated rewards [1]. This approach satisfies both maximizing returns and minimizing risk. Theoretical research primarily centers on these fundamental elements: expected return, which denotes the average return across the portfolio's investments, and risk, measured by variance, indicating the uncertainty and volatility of these returns [1]. The allocation factor w_i refers to the share of the i-th asset within the entire portfolio, while the expected earnings for this asset are denoted by r_i.

$$Expected\ Return\ of\ Portfolio\colon \mathrm{E}(R_P) = W^T R = \sum_i w_i r_i \tag{2}$$

Variance serves as a metric for assessing the volatility or instability of investment returns. When applied to multi-asset portfolios, variance calculation extends the method used for two assets [1]. Unlike expected returns which measure average returns, variance focuses on the fluctuation of returns. This allows investors to evaluate the risk level associated with combining different assets and to optimize the risk-return balance through strategic asset allocation. Denote σ_i as the standard deviation of the i-th asset and ρ_{ij} to signify the degree of association between the returns of the i-th and j-th assets, as indicated by their correlation coefficient.

$$Variance: \sigma_p^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{3}$$

The efficient frontier serves as a vital mechanism for illustrating the relationship between risk, characterized by volatility, and the potential returns of different investment portfolios. It enables investors to graphically assess various portfolios and identify the one that maximizes potential earnings for a defined level of risk acceptance. A portfolio lying on this frontier is deemed optimal when it delivers the highest anticipated returns for a specific appetite for risk.

The maximum Sharpe ratio portfolio, or tangency portfolio, is a significant position on the efficient frontier. It shows the point on the Capital Market Line (CML) where the tangent of the efficient frontier meets the line. The market portfolio's Sharpe ratio establishes the CML's slope, which is a line that originates from the risk-free interest rate (Rf). This portfolio represents the ideal balance between risk and return in the context of the market.

$$Sharpe\ Ratio = \frac{R_p - R_f}{\sigma_p} \tag{4}$$

3.3. Monte Carlo Simulation

This research uses Monte Carlo simulation to generate random samples from a probability distribution of asset returns to mimic a range of probable portfolio returns. When building a portfolio, Monte Carlo simulation is frequently used to evaluate risk, allocate assets optimally, and forecast future portfolio performance in the face of uncertainty. It helps to improve knowledge of risk and uncertainty in prediction [8]. The method of realizing a portfolio is explained in detail in Figure 2.

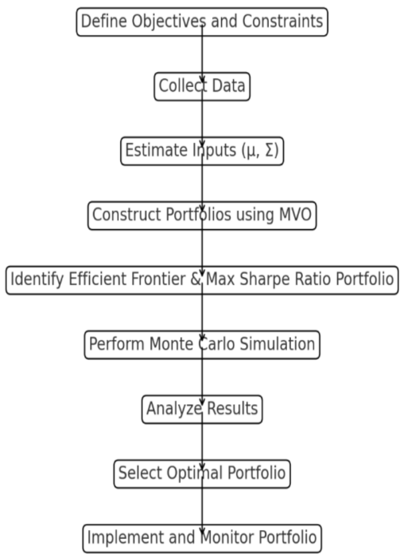

Figure 2: Flowchart for Portfolio Realization

3.4. Long Short-Term Memory

Time series forecasting is the process of projecting future values by taking pertinent information out of historical data. The learning of long-range dependencies presents in time series, a problem that most algorithms find challenging, is efficiently addressed by the LSTM network, a particular deep learning technique, and a kind of recurrent neural network (RNN) [9]. RNNs are robust artificial neural networks with internal memory for maintaining inputs. As a result, they work especially well at handling issues with time series and other sequential data [9]. An important component of LSTM networks is the cell state. It is present throughout the entire sequence, and the gradient is maintained by its linear interactions. Through meticulously controlled structures known as gates, information may be added to or deleted from the cell state. Three different kinds of gates are available in LSTMs to regulate the information flow:

The cell state in an LSTM is selectively pruned by the forget gate, which identifies and removes outdated information. The input gate plays a role in the cell's state by selecting relevant values from the incoming data to update the existing state. The output gate then curates a segment of the cell state to be conveyed as part of the LSTM's output, thereby shaping the subsequent hidden state.

The "gate" structure and cell transitions, which employ outputs around 0 or near a sigmoid function of 1, form the core of the LSTM. This structure facilitates the gradient's smooth passage through the LSTM and lowers the likelihood of gradients vanishing or exploding [10]. Figure 3 illustrates a detailed view of how the LSTM cell works:

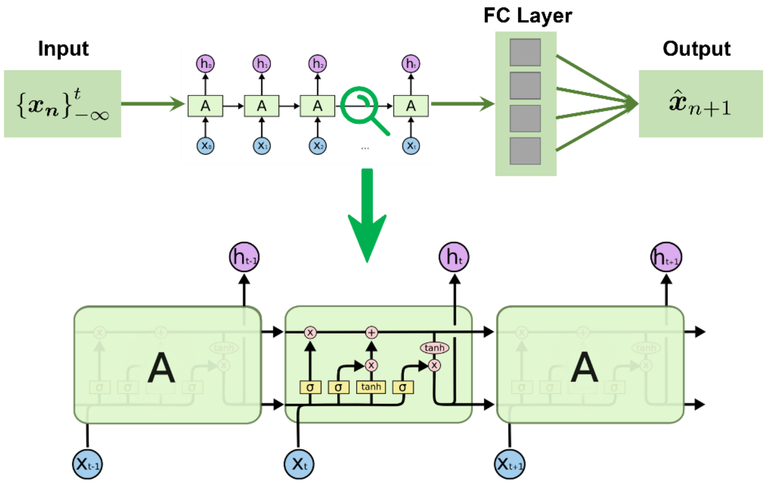

Figure 3: Long and short-term memory network LSTM model

3.5. Standard of Measurement

The discrepancy between the expected and actual values is measured in this research using the mean absolute error (MAE) and root mean square error (MSE). A smaller MSE and MAE indicate a better fit between the model and the data [11]. During the training process of a machine learning model, MSE and MAE can be used as loss functions to guide the model learning and minimize the prediction error. Additionally, they are used in the hyperparameter optimization process to find the best model configuration.

4. Result

4.1. Forecasting

This research presents a complete neural network architecture-based LSTM stock price prediction method. By applying the LSTM model to acquire a forecasting dataset of twelve well-chosen stock prices, the usefulness of the model in time series forecasting is illustrated. The steps involved in setting up a model include data scaling, model training, and iterative forecasting. The metrics used to assess the model's performance are MAE and MSE. The study's conclusions demonstrate the model's ability to accurately forecast changes in stock values.

The model was constructed by first scaling historical price data using MinMaxScaler to normalize values within a range of 0 to 1. The data used from March 1, 2022, to March 1, 2024, is utilized to make the forecast. Eighty percent of the historical data constituted the training dataset, with the study

DOI: 10.54254/2754-1169/90/20242006

utilizing 90 days of past data to predict the 91st day, iteratively incorporating each day's latest prediction. This sliding window approach served as input and output for training the model, enabling continuous learning and enhancement of predictive accuracy as more data became available. Limiting the model to a fixed-size window helps mitigate overfitting to older, potentially outdated data, thereby enhancing robustness against changes in the underlying data distribution.

In order to avoid overfitting, the study also built an LSTM model with three LSTM layers and dropout regularization. During model development, an Adam optimizer and mean squared error loss function were employed. To mitigate overfitting, training incorporated early stopping callbacks. The training regimen encompassed 100 epochs, utilizing a batch size of 32 and a 33% validation split. The training and validation losses of the model were displayed to show its performance. The training and validation losses are shown in Figure 4 to evaluate the model's generalization and convergence. Finally, to assess the accuracy of the predicted data, Table 2 calculated the MSE and MAE, which indicated that the MSE was controlled to 1% and the MAE to 10%.

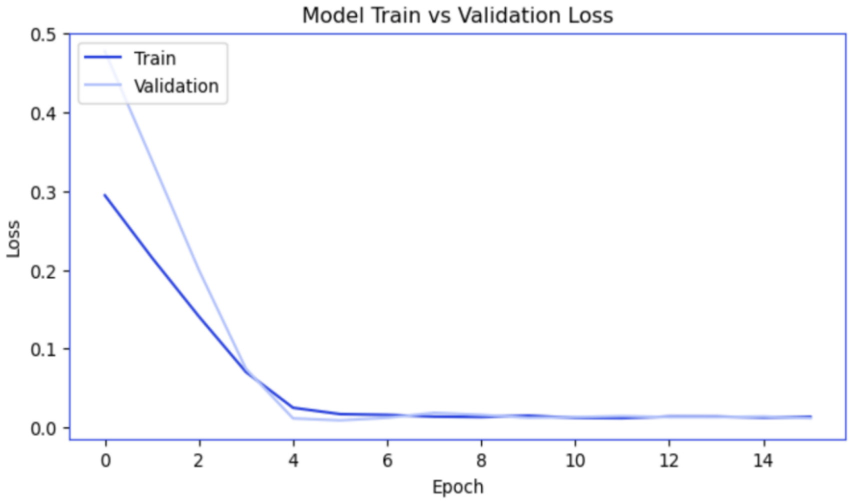

Figure 4: Model train loss and validation loss

Table 2: MSE and MAE

	SHEL.L	CNA.L	HSBA.L	LLOY.L	ULVR.L	ABF.L
MSE	0.0103	0.0080	0.0052	0.0353	0.0257	0.0030
MAE	0.0856	0.0771	0.0590	0.0558	0.0417	0.0477
	AZN.L	HIK.L	DARK.L	SGE.L	BT-A.L	VOD.L
MSE	0.0101	0.0136	0.0015	0.0294	0.0119	0.0013
MAE	0.0758	0.0897	0.0321	0.0633	0.0943	0.0296

4.2. Investment Portfolio

This study uses a Monte Carlo model to assign 106 different portfolio weights to the results. Figure 5 illustrates the anticipated returns and volatilities for portfolios exhibiting the highest and lowest Sharpe ratios.

DOI: 10.54254/2754-1169/90/20242006

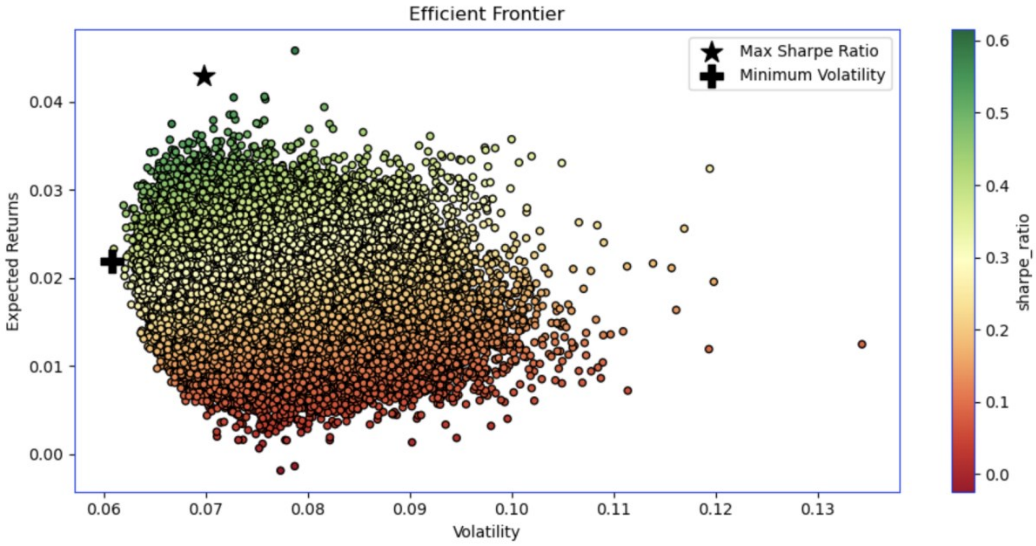

Figure 5: Portfolio weight distribution

The weights assigned to each stock for portfolios that correspond to the optimal Sharpe ratio and the lowest possible variance, as detailed in Table 3.

Table 3: Weights

Stocks	Maximum Sharpe Ratio portfolio	Minimum Variance portfolio
SHEL.L	7.02%	10.39%
CNA.L	2.04%	11.94%
HSBA.L	8.71%	4.69%
LLOY.L	28.81%	1.12%
ULVR.L	3.93%	11.36%
ABF.L	3.95%	10.29%
AZN.L	0.11%	18.71%
HIK.L	10.46%	6.38%
DARK.L	24.22%	3.77%
SGE.L	4.14%	19.65%
BT-A.L	2.53%	0.26%
VOD.L	4.08%	1.45%

4.3. Cumulative Return Rate

In order to assess the validity of the weights derived from the study, the study combines the above weights and utilizes real data to compare the cumulative return changes with the FTSE100 index. The analysis begins by obtaining the prices of 12 selected stocks for the period March 1, 2024, to June 7, 2024. Daily returns were then calculated for each stock and weighted according to the values derived from the study. Subsequently, a cumulative return was calculated based on the daily returns. As of June 7, 2024, the cumulative return of the portfolio optimized for the Sharpe ratio peaked at 24.8%, while the portfolio tailored for minimal variance registered a cumulative return of 10.4%. In stark contrast, the FTSE100 Index rendered a cumulative return of 7.32% over the identical timeframe. These comparative outcomes are graphically represented in Figure 6.

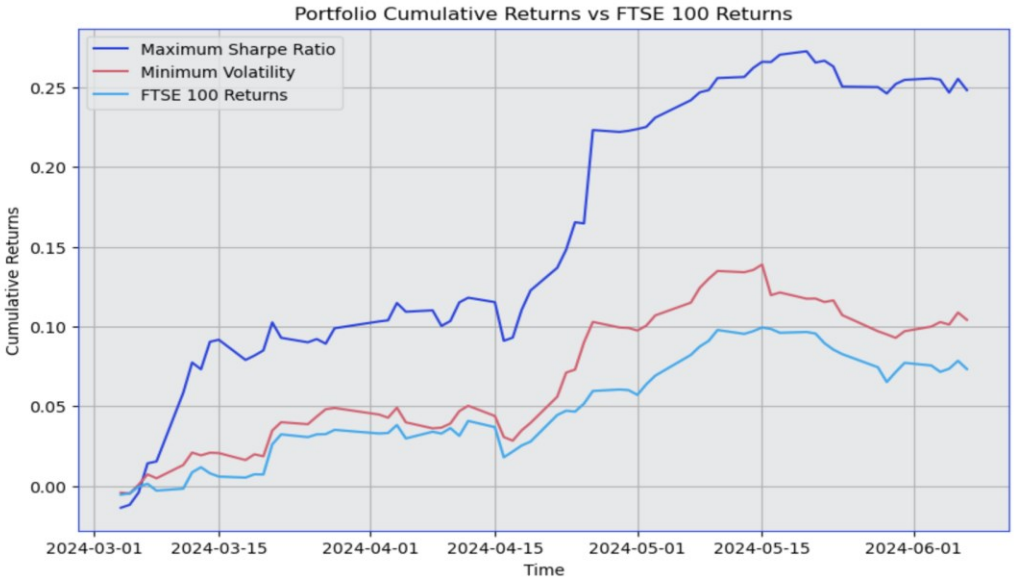

Figure 6: Portfolio Cumulative Returns vs FTSE 100 Returns

The results show that both the Maximum Sharpe Ratio Portfolio and the Minimum Variance Portfolio constructed from the data forecasted in the time series approach of this study significantly outperform the FTSE100 Index over the same period. The significant difference in returns proves the validity of the proposed weights. Therefore, it can be used as a reference for other investors in their allocation practices.

5. Conclusion

To sum up, this study demonstrates that the Mean-Variance (MV) optimization framework and the LSTM network work well together to create optimal portfolios for financial investors. The cumulative returns highlight the potential for improved portfolio performance when stock price volatility is predicted using the LSTM network's predictive capability and then incorporated into the MV model. Empirical results based on data from top UK firms listed on the London Stock Exchange (LSE) between March 1, 2024, and June 7, 2024, show that portfolios optimized for the maximum Sharpe ratio significantly outperform the FTSE100 Index and conservative minimum variance portfolios. This outperformance demonstrates the advantages of integrating cutting-edge machine learning methods with conventional financial theory to offer a more adaptive and flexible approach to portfolio management. The study's conclusions offer valuable insights to financial investors who aim to maximize profits and minimize risk in a constantly shifting market. To further advance the field of portfolio optimization, future studies may explore the incorporation of additional machine learning models and optimization strategies.

References

[1] Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), 77-91.

[2] Širůček, M., & Křen, L. (2017). Application of Markowitz Portfolio Theory by Building Optimal Portfolio on the US Stock Market. Advances in Finance, Accounting, and Economics, 24-42.

[3] Kolm, P. N., Tütüncü, R., & Fabozzi, F. J. (2014). 60 Years of Portfolio Optimization: Practical challenges and current trends. European Journal of Operational Research, 234(2), 356-371.

[4] Paiva, F. D., Cardoso, R. T. N., Hanaoka, G. P., & Duarte, W. M. (2019). Decision-making for financial trading: A fusion approach of machine learning and portfolio selection. Expert Systems with Applications, 115, 635-655.

[5] Dixon, M. F., Halperin, I., & Bilokon, P. (2020). Machine Learning in Finance. Springer International Publishing.

[6] *Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications, 124, 226-251.*
[7] *Jensen, M. C. (1969). Risk, The Pricing of Capital Assets, and The Evaluation of Investment Portfolios. The Journal of Business, 42(2), 167-247.*
[8] *Li, Q., & Zhong, W. (2022). The Application of Monte Carlo Simulation and Mean-Variance in Portfolio Selection. BCP Business & Management, 26, 1216-1221.*
[9] *Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., & Zhang, H. (2019). Deep Learning with Long Short-Term Memory for Time Series Prediction. IEEE Communications Magazine, 57(6), 114-119.*
[10] *F. Karim, S. Majumdar, H. Darabi, and S. Harford (2019), Multivariate LSTM-FCNs for time series classification. Neural Netw, 116, 237-245.*
[11] *Wen, X., & Li, W. (2023). Time Series Prediction Based on LSTM-Attention-LSTM Model. IEEE Access, 11, 48322-48331.*