Proceedings of the 2023 International Conference on Management Research and Economic Development
DOI: 10.54254/2754-1169/26/20230569

Enhancing Portfolio Performances through LSTM and Covariance Shrinkage

Guangqi Li [1, a, *]

[1] *SWUFE-UD Institute of Data Science, Southwestern University of Finance and Economics, Chengdu 611130, China*
a. lgq@udel.edu
**corresponding author*

Abstract. Portfolio optimization is a crucial aspect of finance, requiring advanced analytical tools and modeling techniques. This paper proposes a new method for portfolio optimization that combines Long Short-Term Memory (LSTM) forecasting with Covariance Shrinkage and Mean-Variance Optimization (MVO) to construct diversified portfolios that maximize risk-adjusted returns. The study utilizes an LSTM-based model to predict stock prices, evaluating its performance using the RMSE metric. The calculated RMSE of 0.0849 indicates accurate and robust predictions. The portfolio constructed shows different weights each day for different assets based on the minimum variance and maximum Sharpe ratio portfolios. As of January 3rd, 2023, the assets with the largest proportion in the Maximum Sharpe Ratio portfolio and in the Minimum Volatility portfolio, are respectively BA, accounting for 27.64% of the portfolio and PG, accounting for 32.66% of the portfolio. This paper compares the performance of the proposed method and benchmark methods by applying 30 daily portfolio weights to real returns. The portfolio constructed by the proposed method has higher cumulative return with a higher Sharpe ratio and lower maximum drawdown, indicating a better ability to diversify risks and create returns. The proposed method offers a new perspective on portfolio optimization, which can potentially benefit investors and asset managers.

Keywords: LSTM, covariance shrinkage, MVO

1. Introduction

Portfolio optimization is a crucial task in finance, aiming to construct a portfolio of assets that maximizes returns while minimizing risk. With the ever-growing complexity of financial markets and the abundance of investment options, constructing an optimal portfolio has become a challenging task that requires sophisticated modeling techniques and advanced analytical tools.

The importance of portfolio construction cannot be overstated, as it has a direct impact on the financial performance of investors and asset managers. In recent years, there have been numerous notable events in the financial industry that underscored the importance of diversification and risk management in portfolio construction [1]. For example, the GameStop short squeeze in early 2021 and the pandemic-induced market crash highlighted the vulnerability of concentrated portfolios and the importance of risk diversification.

DOI: 10.54254/2754-1169/26/20230569

There is a vast body of research on portfolio optimization, and various methods have been developed over the years. One of the popular aspects of these approaches is that forecasting methods are linked to portfolio construction. Many studies have shown that incorporating forecasting methods in portfolio optimization can improve portfolio performance.

For instance, Sahamkhadam, Stephan, and Östermark analyzed the effectiveness of portfolio optimization approaches incorporating a model combining ARMA, Garch EVT and copula [2]. Their findings demonstrated that the portfolio which relies on the proposed model, surpasses the benchmark portfolio that is grounded in historical returns. In addition, Chen, Zhang, Mehlawat, Jia proposed a portfolio optimization method based on improved extreme Gradient Boosting (XGBoost) and found that the approach outperformed traditional methods in terms of portfolio returns and risks [3].

LSTM is a popular forecasting method that has gained increasingly attention these years because of its capacity to identify long-term relationships within time-series data [4]. However, existing research combining LSTM with portfolio construction is limited and not enough to draw definitive conclusions. Hence, this paper proposes a portfolio optimization method that combines LSTM forecasting with MVO and covariance shrinkage to construct a diversified portfolio that maximizes risk-adjusted returns.

The proposed method in this paper combines Long Short-Term Memory (LSTM) forecasting and Covariance Shrinkage method with Mean-Variance Optimization (MVO) to construct portfolios. The forecasting process uses historical stock prices to generate predictions, which are then combined with covariance matrices generated using the Covariance Shrinkage method. These predictions and matrices are then used in the MVO model to optimize portfolio weights. Monte Carlo Simulation (MCS) is employed to generate the optimal weights for each day in test interval. To evaluate the performance of the portfolios generated, a back test is conducted using real returns and compared against the S&P 500 index and the returns under equal-weight allocation (EQ). The results show that the portfolio generated by the proposed LSTM + Covariance Shrinkage + MVO method achieved a much higher cumulative return of 14.2%, with a higher Sharpe ratio and lower maximum drawdown, indicating a better ability to diversify risks and create returns.

In summary, this paper highlights the importance of portfolio construction, and proposes a novel approach that combines LSTM forecasting and Covariance Shrinkage method with MVO and covariance shrinkage. The proposed approach provides a new perspective on portfolio optimization and offers potential benefits for investors and asset managers.

2. Data

The data utilized in this paper was sourced from Yahoo Finance [5]. Ten stocks were selected based on their market capitalization and diversification across different sectors. The tickers of these stocks are 'TSLA', 'DIS', 'GS', 'PFE', 'BA', 'AAPL', 'PG', 'PLD', 'SHEL', 'UNP'. The data covers a time range from January 1st, 2018, to February 15th, 2023, and consists of 1289 pieces of daily closing stock prices for each stock. The data is then divided into train set and test set. The train set are composed of 1189 pieces of data from January 1st, 2018, to September 22nd, 2022. The train set contains the remaining 100 pieces of data is subdivided into test set for input and test set for result validation. The test set for input contains 70 pieces of data from September 23rd, 2022, to December 31st, 2022. The test set for result validation contains 30 pieces of data from January 1st, 2022, to February 15th, 2023. Their relationship are illustrated in Table 1.

Table 1: Data split.

	Train set	The test set for input	Test set for result validation
Amount	1189	70	30
Data	Jan 1st, 2018, to Sep 22nd, 2022	Sep 23rd, 2022, to Dec 31st, 2023	Jan 1st, 2023, to Feb 15th, 2023

On the basis of preserving the original stock prices data to evaluate the performance of stock price estimation, they are further transformed into daily simple returns to verify the performance of portfolio optimization. The transformation is based on the formula:

$$R_t = \frac{P_t - P_t}{P_{t-1}} \quad (1)$$

where P_t denotes the closing price at time t. To provide an overview of the dataset, this paper conducts a descriptive statistical analysis on the daily returns of the ten assets. The evaluation encompassed metrics like mean, standard deviation, minimum, and maximum. The results are illustrated in Table 2.

Table 2: Descriptive statistics of daily returns for the ten assets.

	AAPL	AMZN	BA	DIS	GS	PFE	PG	PLD	SHEL	UNP
Mean(%)	0.125	0.066	0.031	0.020	0.058	0.047	0.054	0.081	0.039	0.056
std	0.021	0.023	0.032	0.021	0.021	0.017	0.014	0.019	0.023	0.018
Max(%)	11.981	13.536	24.319	14.412	17.580	10.855	12.009	11.810	19.680	13.004
Min(%)	-12.865	-14.050	-23.848	-13.163	-12.705	-7.735	-8.737	-17.271	-17.172	-13.034

In addition, cumulative returns of ten stocks are calculated and plotted in Fig. 1 according to the formular:

$$R_t(k) = (R_t + 1) * (R_{t-1} + 1) * \dots * (R_{t-k+1} + 1) - 1 \quad (2)$$

Where the $R_t(k)$ represents the k-period simple return from time $t - k$ to t.

Figure 1: Cumulative returns of ten stocks from January 1st, 2018, to February 15th, 2023.

3. Methodology

3.1. Overall Framework

Basic framework of this project mainly consists of three parts, respectively estimation, optimization, validation & evaluation process, as is shown in Fig. 2.

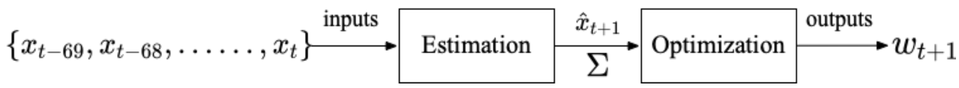

Figure 2: Flowchart of overall framework.

Estimation Side: The portfolio construction is based on estimation results, and thus, it is crucial to pay close attention to this process. The forecasting method adopted in this paper is based on Long Short-Term Memory (LSTM). Based on historical stock closing price $\{x_0, x_1, \ldots\ldots, x_t\}$, this paper selects 100 pieces of stock closing price $\{x_{t-69}, x_{t-68}, \ldots\ldots, x_t\}$ as training data to be input into LSTM to obtain $\hat{x}_{t+1}$, the predictive value of x_{t+1}. The prediction results and covariance matrix Σ generated by Covariance Shrinkage method using $\{x_{t-69}, x_{t-68}, \ldots\ldots, x_t, \hat{x}_{t+1}\}$ are then provided for portfolio optimization in the next step.

Optimization Side: Drawing on the predicted values and covariance matrices, this paper constructs a Mean-Variance Optimization model (MVO). The optimal portfolio weights w_{t+1} can be obtained using Monte Carlo Simulation (MCS). Subsequently, this paper calculates portfolio returns and evaluate its real performance in day t+1, which is precisely what this paper aims to accomplish in the following step.

Validation and Evaluation: In this part, this paper conducts a back test on the final portfolio to evaluate the real returns under the weights generated by the model and real daily returns. The real returns are compared with S&P 500 and the returns under Equal-weights allocation (EQ) over the same time frame to determine the model's performance.

3.2. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a kind of recurrent neural network (RNN) engineered to tackle the vanishing gradient issue, a prevalent challenge in RNNs that complicates the learning of long-

DOI: 10.54254/2754-1169/26/20230569

term dependencies [6]. The primary advancement of LSTM lies in the implementation of memory cells and gates, enabling the network to selectively retain or discard information over time [7]. The equations of LSTM can be defined as follows:

$$\begin{gathered} f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \tilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t = f_t * C_{t-1} + i_t * \tilde{C_t} \\ o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \\ h_t = o_t * \tanh(C_t) \end{gathered} \tag{3}$$

Where x_t is input vector; i_t, o_t, f_t are the input, output and forget gates, respectively; $\tilde{C_t}$ represents the candidate cell state; C_t and h_t are the new cell and hidden states, respectively; W_f, W_i, Wc, W_i are weight matrices; b_i, b_C, b_o, $b_f$are bias vectors; and σ is the sigmoid activation function. The structure of a single LSTM cell and repeating LSTM module are shown in Fig. 3 and Fig. 4.

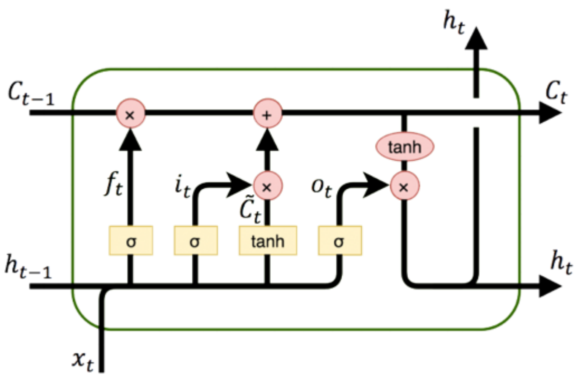

Figure 3: Structure of the LSTM cell.

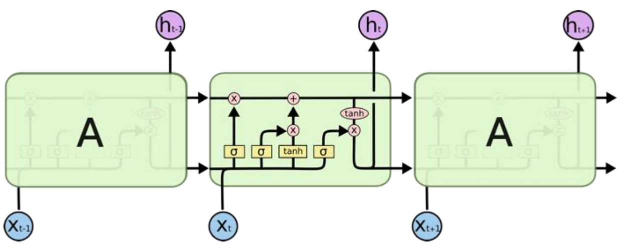

Figure 4: The repeating module in an LSTM.

The four essential components of an LSTM cell work in concert to manage the flow of information. First, the forget gate determines which parts of the previous cell state should be preserved or discarded. Next, the input gate identifies new data to be integrated into the cell state. Concurrently, the candidate cell state generates potential information for inclusion. Lastly, the output gate selects the portions of the updated cell state to be released as the hidden state.

In this paper, author use TensorFlow to construct a neural network with 2 LSTM layers and 2 fully connected layers. The LSTM layers allow the network to learn long-term dependencies in the time-series data, while the fully connected layers help to extract relevant features for the prediction task.

DOI: 10.54254/2754-1169/26/20230569

3.3. Mean-Variance Optimization

Mean-Variance Optimization (MVO) is a widely used mathematical framework in finance for constructing optimal portfolios of assets. MVO aims to find the weights of assets in a portfolio that maximize its expected return while minimizing its variance or risk.

The MVO framework, initially presented by economist Harry Markowitz in 1952, has evolved into a fundamental aspect of contemporary portfolio theory [8]. MVO's critical observation is that an asset's risk and return should be assessed in relation to its impact on the portfolio's overall risk and return [9]. The equations of MVO can be defined as follows:

$$\mathrm{E}(R_p) = \sum_i w_i \mathrm{E}(R_i)$$
$$\sigma_p^2 = \sum_i \sum_j w_i w_j \sigma_i \sigma_j \rho_{ij} \tag{4}$$

Where R_p represents the portfolio return; R_i denotes the return of asset i; w_i signifies the proportion of each constituent asset i in the assembled portfolio; σ_i and σ_j are the standard deviations of the periodic returns of assets i and j, respectively; ρ_{ij} is the correlation coefficient measure the relationship between the returns on assets i and j.

The efficient frontier in MVO symbolizes the collection of portfolios delivering the highest expected return at a given level of risk or the lowest risk for a specified expected return. Portfolios situated beneath the efficient frontier are deemed inferior, as they fail to yield sufficient returns relative to their associated risk. To identify the efficient frontier, the following equation is minimized:

$$\min \mathrm{f} = w^T \Sigma w - q \times R^T w \tag{5}$$

Where w is vector representing the proportion of each asset in the portfolio; Σ is a measurement for co-variance among the select asset; q is a parameter which means risk tolerance of investors; R is a vector of expected returns.

MVO provides a formal and quantitative way for investors to diversify their portfolios and optimize their risk-return tradeoff. By balancing the expected return and variance of a portfolio, MVO helps investors construct portfolios that meet their individual investment objectives and risk tolerance.

3.4. Covariance Shrinkage

In the field of portfolio optimization, the estimation of the covariance matrix is a critical step as it measures the degree of association between asset returns. However, the traditional method of estimating covariance from historical data is limited by small sample size, noise, and non-stationarity of the underlying process [10]. Covariance shrinkage, also known as regularization or regularization estimators, is a popular technique that improves covariance matrix estimation by reducing noise and estimation bias. One of the most widely used covariance shrinkage methods is the Ledoit-Wolf shrinkage, which can be expressed as follows:

$$\hat{\Sigma}_{LW} = \delta \hat{\Sigma}_o + (1 - \delta) \hat{\Sigma}_S \tag{6}$$

Where $\hat{\Sigma}_{LW}$ is constructed by a parameter δ that controls the degree of shrinkage of the original matrix; $\hat{\Sigma}_o$ is the target matrix proportional to the identity matrix; and $\hat{\Sigma}_S$ is the sample covariance matrix. The optimal value of δ is obtained through cross-validation, which balances the bias and variance of the estimation.

DOI: 10.54254/2754-1169/26/20230569

4. Result

In the following section, the results of the LSTM models' stock prediction performance during the entire testing period are presented. Two different portfolios were then constructed using the MVO model combined with the LSTM-based estimation and the EQ strategy. The real return rates of these portfolios were applied to compare the overall performance of the LSTM+MVO and EQ models during the entire testing period's Estimation Result

4.1. Stock Prediction Evaluation

In this section, the LSTM model is constructed utilizing the Keras library, incorporating two LSTM layers and a pair of fully connected layers. The model is then trained using the training set that contains 1189 pieces of data and the mean squared error loss function. This paper iteratively tunes the number of epochs from 1 to 30, to select the optimal parameter for the program. Ultimately, the model is employed to generate predictions for the test data. The outcomes are assessed using the Root Mean Square Error (RMSE) metric and graphed alongside the actual prices. RMSE is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{T} (\hat{y}_t - y_t)^2}{T}}. \tag{7}$$

Where $\hat{y}_t$ represents the predicted value of time t; y_t represents the observed value of time t; T represents the total number of predictions that are made. This paper pick AAPL to demonstrate. The historical adjusted closing price data are acquired from yahoo finance, as is shown in Fig. 5.

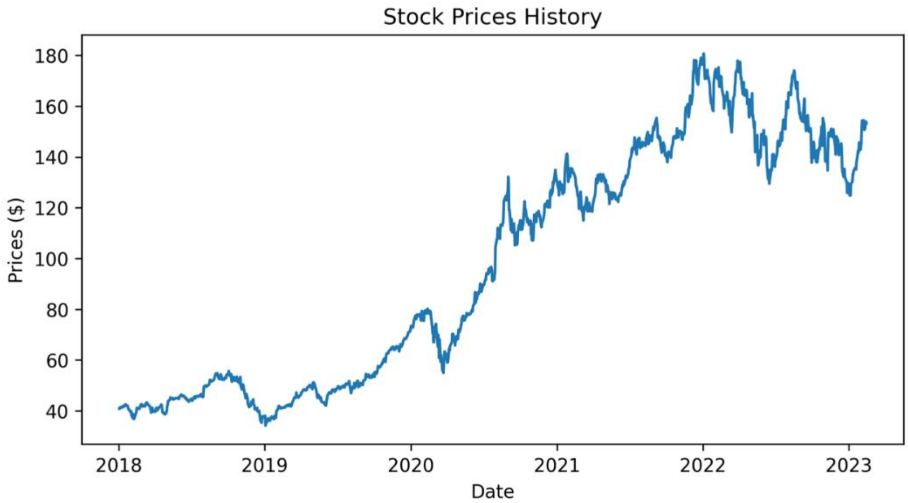

Figure 5: History adjusted closing price of AAPL.

Number of epochs from 1 to 30 against the loss are plotted as is shown in Fig. 6. This paper discover that after 10 rounds of fitting, the loss is basically stable. So, this paper chooses 10 epochs for further prediction.

DOI: 10.54254/2754-1169/26/20230569

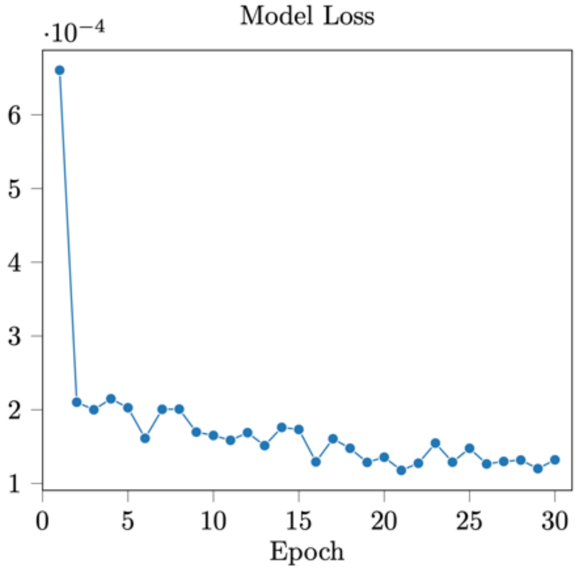

Figure 6: Loss under different number of epochs.

After training, RMSE is employed to assess the performance of predictions generated by LSTM model. The RMSE of the result is 0.0849, which shows LSTM is able to prevent vanishing gradient problems. The predicted stock prices are plotted along with the real stock prices as is shown in Fig. 7. LSTM proves to be quite accurate and robust since the prediction line and validation line overlap almost everywhere.

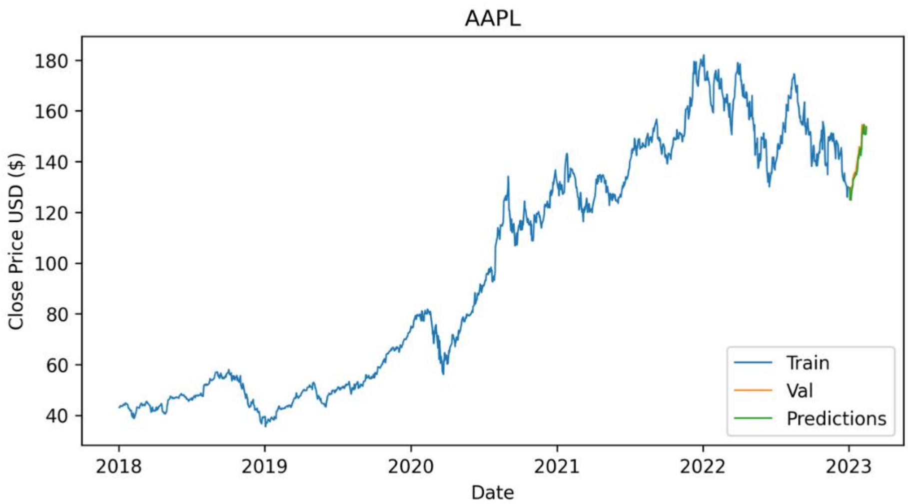

Figure 7: Visualized prediction performance of LSTM model.

4.2. Portfolio Performance Evaluation

In this panel, this paper uses LSTM model to forecast the share price of each stock on next trading day based on a sliding window. The MVO model with Covariance Shrinkage is then adopted to construct a portfolio for the specific day, based on the historical adjusted closing price of each stock for the past 69 trading days and the forecast for the trading day after. The portfolio weights are updated daily as the magnitude of the forecasts and the range of historical data are changing daily and so are the expected returns and covariance matrices.

This paper picks portfolio constructed on December 31st. 2022 for the investment on January 3rd. 2023 to demonstrate. Based on the LSTM, this paper forecast the adjusted closing price of each stock on 3 January 2023 and calculate the daily forecast return for each stock on that date, as shown in Table 3.

Table 3: Forecast daily return for January 3rd. 2023 (%).

	AAPL	AMZN	BA	DIS	GS	PFE	PG	PLD	SHEL	UNP
Value	3.7405	2.5723	2.4056	0.8271	0.0390	0.0066	0.0355	1.7033	12.2422	0.2463

Then based on 69 daily historical returns of each stock from September 23rd. 2022 to December 31st, 2022, and forecast daily return for January 3rd. 2023, covariance matrix Σ is estimated by Covariance Shrinkage method as is shown in Fig.8.

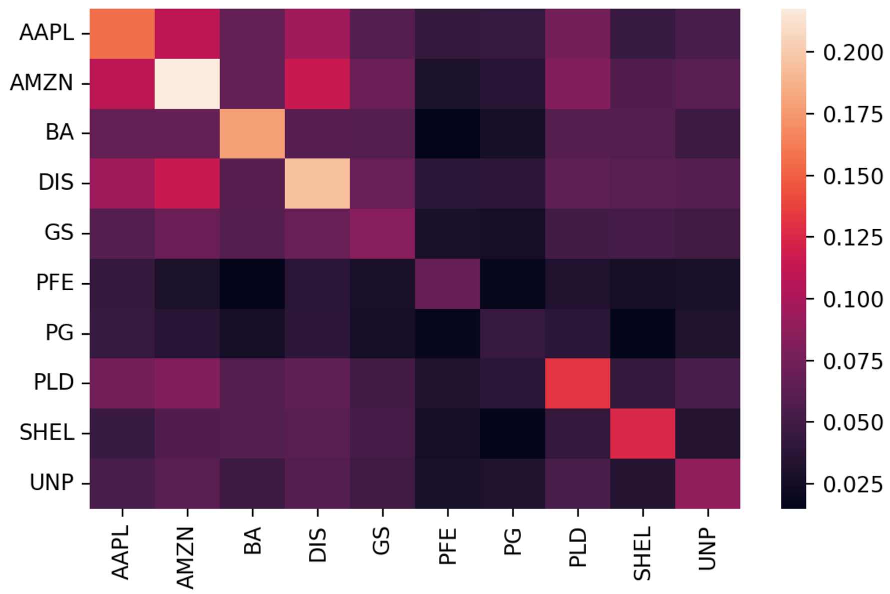

Figure 8: Covariance matrix.

To perform mean-variance optimization, this paper conduct Monte Carlo Simulation for 100,000 times. The 69 historical daily returns of each stock from September 23rd. 2022 to December 31st, 2022, and forecast daily return for January 3rd. 2023 are used to generate expected return in MVO. The covariance matrix used in MVO is covariance matrix Σ estimated by Covariance Shrinkage method. The results are shown in the scatter plot in Fig. 9.

DOI: 10.54254/2754-1169/26/20230569

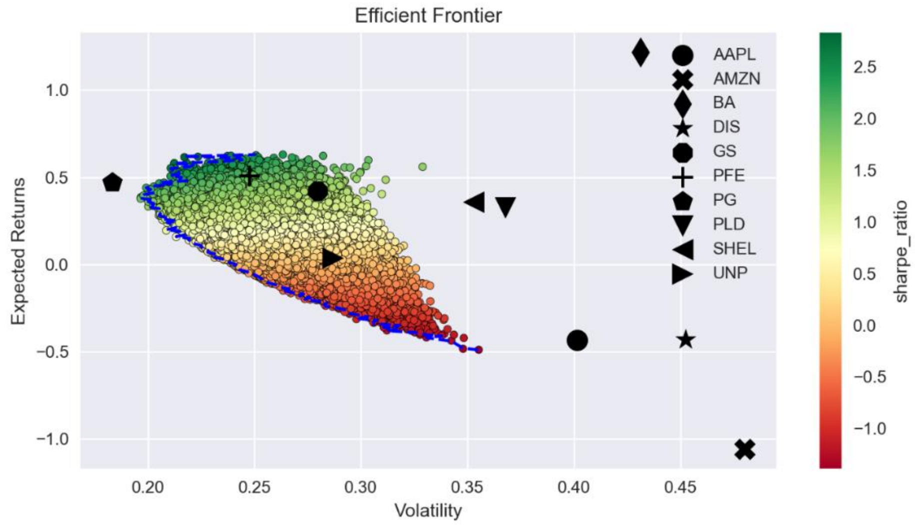

Figure 9: Efficient frontier plot with single stock return and volatility.

The efficient frontier is indicated by the upper part of the blue edge. The black markers in the graph indicate where the portfolio would be if it consisted of only one stock. The sub-optimal portfolio lies below the efficient frontier and provides a relatively low return for a given level of risk. The two target portfolios, respective the minimum volatility portfolio and the maximum sharp ration portfolio are calculated and demonstrated in Fig. 10.

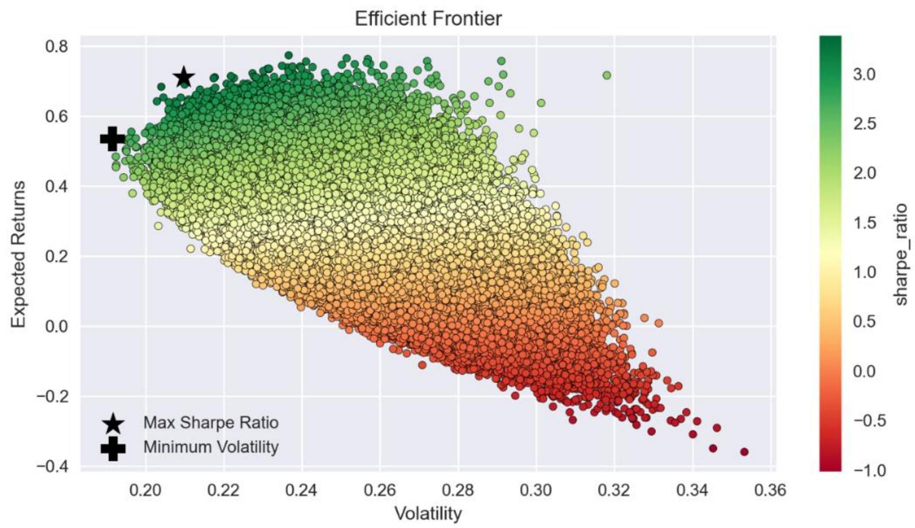

Figure 10: Efficient Frontier Plot with minimum volatility and the maximum sharp ration portfolios.

The weight of each stock under the two portfolios for the investment on January 3rd. 2023 are shown in Table 4.

Table 4: Weight of each stock in the two optimal portfolios for January 3rd. 2023 (continue).

	Max Sharpe Ratio	Min Volatility
AAPL	0.45	4.09
AMZN	1.27	0.17

DOI: 10.54254/2754-1169/26/20230569

Table 4: (continued).

BA	27.64	4.30
DIS	2.06	2.02
GS	0.10	0.76
PFE	26.24	27.89
PG	27.41	32.66
PLD	3.59	4.53
SHEL	7.43	9.41
UNP	3.79	14.16

This paper construct 30 distinct daily portfolios from January 3rd, 2023, to February 15th, 2023, in the same way as shown above. By applying these portfolio weights to real returns generated by adjust closing price in the test set, the daily returns of these portfolios and the cumulative return within 30 trading day can be calculated, since these portfolios are constructed under the same strategy. This paper takes portfolio that achieves maximum sharp ratio to demonstrate. To evaluate the performance of the portfolio's construction strategy, S&P 500 index and the real return under equal weight (EQ) portfolio are brought in. The result is shown in Table 5 below:

Table 5: Performance under different portfolios from Jan 3 to Feb 15, 2023.

	S&P500	EQ	LSTM+Covariance Shrinkage+MVO
Cumulative Return	7.7%	12.3%	14.2%
Volatility	16.3%	16.9%	16.7%
Sharp Ration	3.91	5.85	6.77
Max Drawdown	-2.5%	-2.1%	-1.5%

The benchmark S&P 500 shows a 7.7% cumulative return, which means the general market is in a good condition. The equal weighted portfolio provides a 12.3% cumulative return during this period. And our portfolio weight generated by LSTM + Covariance Shrinkage + MVO has a better performance. At the same level of volatility, it can achieve much higher returns, 14.2% cumulative return to be specific, with a higher sharp ratio and lower max drawdown. LSTM + Covariance Shrinkage + MVO method can generate a portfolio with a higher capability to diversify risks and create returns.

5. Conclusion

In conclusion, this paper has presented a new approach to portfolio optimization that combines LSTM forecasting with Covariance Shrinkage and MVO to maximize risk-adjusted returns. By incorporating historical stock prices, the LSTM model generates predictions that are combined with covariance matrices generated using the Covariance Shrinkage method. These predictions and matrices are then used to determine optimal portfolio weights in the MVO model. To assess the effectiveness of this approach, a back test was conducted using real returns and compared against the S&P 500 index and the returns under equal-weight allocation (EQ). The results demonstrate that the proposed method outperforms traditional methods, achieving a much higher cumulative return with a higher Sharpe ratio and lower maximum drawdown. The proposed method outperformed existing portfolio optimization methods and provided valuable insights into portfolio construction. The study demonstrated the importance of diversification and risk management in portfolio optimization and showed how advanced modeling techniques can improve portfolio performance.

The proposed method still has limitations with LSTM's vanishing gradients and the omission of macroeconomic factors. Future research could explore alternative models and incorporate macroeconomic indicators for improved accuracy. Additionally, incorporating constraints like transaction costs and leverage could enhance the model's sophistication.

References

[1] *Kalymon, B. A.: Estimation risk in the portfolio selection model. Journal of Financial and Quantitative Analysis 6(1), 559-582 (1971).*

[2] *Sahamkhadam, M., Stephan, A., Östermark, R.: Portfolio optimization based on GARCH-EVT-Copula forecasting models. International Journal of Forecasting 34(3), 497-506 (2018).*

[3] *Chen, W., Zhang, H., Mehlawat, M. K., Jia, L.: Mean–variance portfolio optimization using machine learning-based stock price prediction. Applied Soft Computing 100, 106943 (2021).*

[4] *Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation 9(8), 1735–1780 (1997).*

[5] *Yahoo finance, https://finance.yahoo.com/,last accessed 2023/4/1.*

[6] *LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436–444 (2015).*

[7] *Gers, F. A., Schmidhuber, J., Cummins, F.: Learning to forget: Continual prediction with LSTM. Neural Computation 12(10), 2451–2471 (2000).*

[8] *Markowitz, H.: Portfolio Selection. The Journal of Finance 7(1) 77–91(1952).*

[9] *Elton, E. J., Gruber, M. J.: Investments and portfolio performance. World Scientific (2011).*

[10] *Ledoit, O., Wolf, M.: A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis 88(2), 365–411 (2004).*