RESEARCH ARTICLE OPEN ACCESS

Diversified portfolio optimization under deep learning prediction and heuristics optimization

Yoochan Noh*, Abhijit Debnath**
**(Korea International School, Korea)*
*** (Department of Basic Science,Asansol Engineering College, Kanyapur, Vivekananda Sarani, Asansol, West Bengal, India. 713305)*

ABSTRACT
In the financial market volatility and uncertainty are the biggest challenges to the proper handling of the portfolio of stocks. Much research has been done on portfolio optimization using a meta-heuristic approach. However, the difficulty still lies in the allocation of stocks in the portfolio to simultaneously maximize the return and minimize risk in diversified stock allocation. In this study, a diversified portfolio with 10 stocks is randomly considered from the semiconductor industry, car industry, pharmaceuticals, retail chain, and service industry. Since the stocks are from different sectors, the covariance analysis is also done. The price prediction and subsequently the return of the portfolio is done with LSTM and Attention mechanism models. After that, the multi-objective optimization is done with mean-variance optimization, genetic algorithm, and particleswarm optimization. The PSO always shows a higher portfolio balance in terms of the Sharpe ratio, but the predictive model of pricing under an attention-based mechanism outperforms the general LSTM model. Comparative analysis and managerial insights for the portfolio managers are also provided based on the results obtained.

***Keywords*–**LSTM, Portfolio optimization, Attention mechanism, Genetic algorithm, Particle Swarm optimization

---

Date of Submission: 09-07-2024 Date of acceptance: 23-07-2024

I. INTRODUCTION

In recent years, there has been a notable increase in interest in the realm of portfolio optimization, primarily attributed to the intricate challenges associated with attaining equilibrium between maximizing returns and minimizing risks. Conventional techniques for portfolio optimization, though effective to some extent, frequently prove inadequate in tackling the dynamic and non-linear characteristics of financial markets. This study aims to bridge this existing disparity by utilizing advanced deep-learning models for price forecasting and incorporating meta-heuristic strategies for optimizing a diversified asset portfolio encompassing three distinct sectors.

The rationale for conducting this research stems from the necessity for more resilient and adaptive methodologies to oversee investment portfolios within a progressively unstable market setting. Conventional strategies like mean-variance optimization frequently make assumptions about unchanging market circumstances and the normal distribution of profits, which do not align with reality. Furthermore, the correlations among assets from diverse sectors introduce layers of complexity that conventional methods may not adequately tackle. [1] has worked on the application of machine learning to accurate prediction of the arithmetic and geometric average options of asset pricing. Also, the future stock value has been predicted using linear and multiple linear regression models [2]. In an advanced study by[3], the asset price prediction was done by setting up an attention-scaled DL algorithm. In stock price movement another significant study was done by[4], where machine learning was amalgamated with belief rules for monitoring and predicting the stock price movements.[5]discussed the application of deep learning for a factor asset pricing model with better return movements. In another study with five-factor models by[6], deep learning models are used for capturing the non-linearity in pricing structures. Deep learning models, renowned for their ability to discern complex patterns and connections from extensive datasets, present a viable approach for precise price forecasting. Augmenting this with meta-heuristic optimization techniques can notably boost the capacity to identify optimal portfolio distributions that can withstand market fluctuations.

The underpinning of portfolio optimization can be historically linked to Harry Markowitz's Modern Portfolio Theory (MPT), which introduced the principle of diversification for risk mitigation.

Throughout the years, numerous advancements and substitutes to MPT have been suggested, encompassing the Capital Asset Pricing Model (CAPM), Arbitrage Pricing Theory (APT), and more intricate quantitative methodologies. Nevertheless, these frameworks frequently rest upon assumptions that may lack practical validity, such as the linearity and stationarity of returns.

Conversely, deep learning models, notably neural networks, have exhibited notable success in capturing non-linear interrelationships and time-evolving trends within financial data. Research has indicated that deep learning surpasses conventional econometric models in forecasting stock prices and market trends. Nonetheless, the integration of these forecasts into a viable portfolio optimization framework poses a persistent challenge.

Meta-heuristic algorithms, such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO), and Simulated Annealing (SA), have garnered traction for resolving intricate optimization quandaries. These algorithms do not necessitate gradient intelligence and are less susceptible to being ensnared in local optima, rendering them well-suited for portfolio optimization given the extensive and multi-modal search space. [7]also worked on the optimization of diversified portfolios with complex features and constraints and different asset classes under meta-heuristic approaches. [8]provided a review of applications of meta-heuristics on portfolio optimization and risk mitigation where different past approaches and current trends were discussed. An advanced study with a stochastic approach to portfolio optimization was done by[9], where the noisy covariance and the average return were taken as the random variables. The objective function of this study was solved with a heuristic approach. Another comprehensive survey was conducted by[10], by which the feature selection in different sectors is discussed with different nature-inspired heuristic approaches. In quantitative investment, portfolio optimization is also done by [11] with the help of the Mayfly algorithm. The objective function of this study was constructed with the cardinality-constrained method. [12]also proposed the model of sequential ensemble under meta-heuristics for analyzing the portfolio efficiently. This study includes the portfolio allocations with maximization of return.

In all the literature portfolio optimization is done with mainly the application of statistical approaches or the meta-heuristics related to the genetic algorithm. If the return is found based on SOTA models, metaheuristics are applied for the optimization. On the other hand, if the ML algorithms are utilized under factor models of the return, some single objective optimizations take place. In our study, the stepwise price prediction is done with deep learning models, and also using the attention mechanism for better prediction using long-term memory. Subsequently, the stock selection is also done with mean-variance optimization. Finally, the meta-heuristics is applied to find the optimized portfolio with minimized risk.

II. DATA ACQUISITION AND PREPROCESSING

The method of web crawling has been considered for the data collection from Yahoo Finance. Data for 10 stocks from sectors of car, IT, Pharmaceuticals, retail, and, FMCG are chosen.

A total of five sectors are selected for making the diversified portfolio. In Python, using tickers from the yfinance the data is downloaded in the range of 15 years until June 30, 2024.

Firstly, historical data for different assets were collected from Yahoo Finance, specifically 10 different assets into a single CSV file. These datasets were loaded into Data Frames using the Python library pandas. To easily deal with date-related data in Python, the "Date" column in each Data Frame was converted to a date time format. After checking, there were no rows with missing data for all files, meaning missing data did not need to be removed. Given that the data spans different periods for each particular asset, the common dates from all of them were identified, and each data frame was filtered so that only the common dates existed. From there, only the "Adj Close" columns from the assets were identified and a new Data Frame was created, only to save those values. In the end, the whole dataset spanned for the range of 15 years. The columns 'Low','High','Open', 'Volume', and 'Close' are eliminated to form the final dataset for further calculations of return and risk.

III. Methodology

This study includes a comprehensive study of price prediction through the deep learning process and then the portfolio optimization is done using the conventional and meta-heuristic approach. The following subsections: 1. Prediction strategy, and 2. Optimization strategy.

In prediction, the Adjusted closing price is first predicted, and the predicted data is stored which finally goes for the optimization for the following objectives:

Expected portfolio return

$$E(ret_p) = \sum_{i=1}^{k} b_i E(ret_i)$$

$b_i$is the weights of asset I and k is the total asset number.

Portfolio Risk

The risk of a portfolio is often represented as the standard deviation of the portfolio return and is calculated considering each asset's risk and their correlations. The mathematical equation for such risk (σ_p) is as follows:

$$\sigma_p = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} b_i b_j \sigma_i \sigma_j \rho_{ij}}$$

Where b_i and b_j are the weights of assets i and j in the portfolio, σ_i and σ_j are the standard deviations of assets i and j, and ρ_{ij} is the correlation coefficient between the returns of assets i and j.

Sharpe Ratio

Sharpe ratio is used to measure the risk-adjusted return of a portfolio. Its equation is as follows:

$$S = \frac{E(R_p) - R_f}{\sigma_p}$$

Where $E(R_p)$ is the expected return of the portfolio, R_f is the risk-free rate, and σ_p is the portfolio risk.

A. Price prediction using LSTM

LSTM networks are a type of RNN, or Recurrent Neural Network, which are very effective in terms of time series prediction due to their ability to capture long-term dependencies. This makes it particularly suitable for asset pricing where a lot of historical data is involved.

These LSTM networks use 3 different gates:
i_t, the input gate, f_t, the forget fate, and o_t, the output gate, which helps manage C_t, the cell state, andh_t, the hidden state. The specific equations used for this are:

- Forget gate : $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$
  This forget gate is used to decide which information from the previous cell state should be removed. The sigmoid function σ determines that extent by outputting a value between 0 and 1.
- Input gate:
  $$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$
  $$C_t = \tanh(\widetilde{W_c \cdot [}h_{t-1}, x_t] + b_c)$$
  This input gate is used to update the cell state with new information. It has two parts: the sigmoidfunction σ to determine which values to update, and the tanh function tanh which creates new candidate values to be added to the state.
- Cell state update:
  $$C_t = f_t C_{t-1} + i_t C_t$$
  The cell state is updated by combining the previous cell state C_{t-1}, adjusted by the forget gatef_t, with the new candidate cell state $\widetilde{C}_t$, adjusted by the input gate i_t.
- Output gate:
  $$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$
  $$h_t = o_t \cdot \tanh(C_t)$$

The illustration of the LSTM model is also depicted in **Figure 1**. ,where the gates along with the cell states are shown.

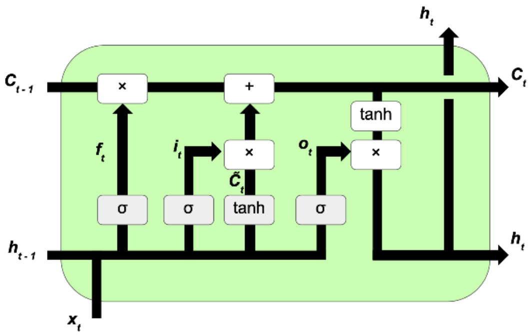

Figure 1. Illustration of LSTM architecture

In the present study, the deep neural network model summary is givenin fig

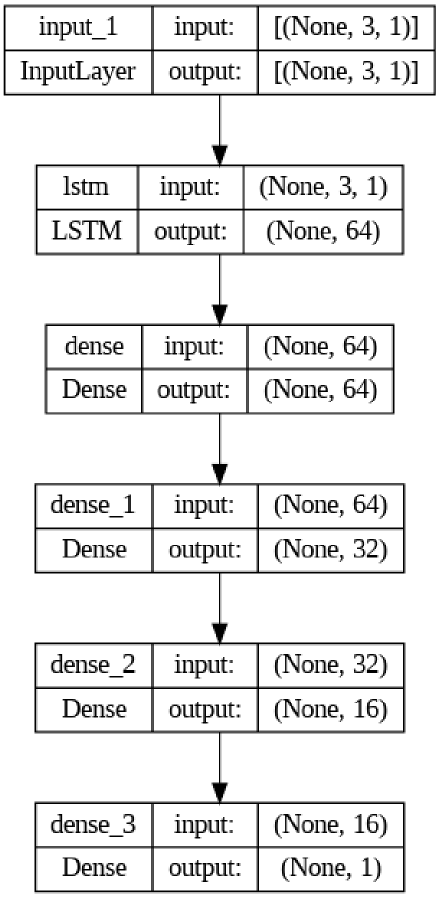

Figure 2. Model summary for LSTM of preset study

B. Price prediction using an attention mechanism

The attention mechanism operates by computing a set of weights known as attention weights for each time step within the input sequence. These weights signify the relevance of individual time steps in influencing the final output. The attention mechanism typically comprises three primary elements: the query, the key, and the value. In the context of Long Short-Term Memory (LSTM) networks, the query represents the current hidden state, the key denotes the hidden states at each time step, and the value corresponds to the same hidden states as the key. By computing the dot product of the query with each key and subsequently normalizing these weights through a SoftMax function, the attention mechanism determines the attention weights. These weights are then utilized to calculate a weighted sum of the values, which in turn contributes to the final prediction.

The importance of the attention mechanism in time series data stems from the varying significance of past observations over time. Through a dynamic focus on pertinent time steps, the LSTM equipped with an attention mechanism enhances its capability to capture temporal dependencies and patterns crucial for precise predictions. This capability facilitates improved handling of long-term dependencies and temporal correlations, thereby enhancing the model's effectiveness in tasks like stock price forecasting, weather prediction, and other sequential data analyses. The architecture of an LSTM integrated with an attention mechanism consists of LSTM layers for processing the input sequence, an attention layer for computing attention weights and generating a context vector, and an output layer for final prediction. The integration of the attention mechanism with LSTM optimizes the model's performance on time series prediction tasks by effectively capturing long-term dependencies and temporal patterns, thereby enhancing its predictive accuracy. The architecture is based on three steps:

1. LSTM layers
2. Attention layers, where the context vector is generated through the attention weights.
3. The output layer gives the final output using the context vectors.

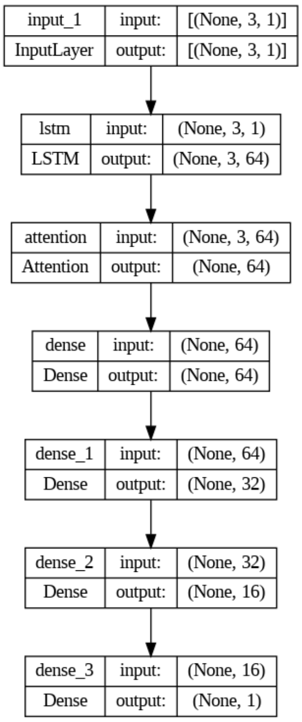

Figure 3. Attention based LSTM model summary

C. Optimization by genetic algorithm (GA)

The Genetic Algorithm (GA) is a computational approach rooted in heuristic search and optimization methods, drawing inspiration from the fundamental concepts of natural selection and genetics. Within the realm of portfolio optimization, GA is designed to enhance returns and mitigate risks through the iterative progression of a population of potential portfolios. Initialization commences with the creation of a random portfolio population, each delineated by an array of asset proportions. Evaluation of portfolio fitness is conducted based on a predefined objective function, typically encompassing anticipated returns and risks such as variance or standard deviation. The selection of portfolios is contingent upon their fitness evaluations, leading to subsequent genetic processes like crossover (integration of segments from two portfolios) and mutation (random adjustments in portfolio weights). This ongoing cycle of selection, crossover, and mutation unfolds across numerous generations, steadily refining the population toward an optimal resolution. By emulating the evolutionary mechanism, GA adeptly navigates the exploration space, managing the equilibrium between exploration and exploitation, ultimately homing in on a portfolio that establishes an optimal equilibrium between maximizing returns and minimizing risks. The pseudocode for the genetic algorithm is given as:

Algorithm for GA

Initialize population of portfolio weights

Define fitness function to evaluate negative Sharpe ratio

Set GA parameters (population size, mutation probability.)

Repeat for a set number of generations or until convergence:

- *Evaluate the fitness of each candidate's solution*
- *Select a subset of solutions based on fitness*
- *Apply crossover to create new solutions*
- *Apply mutation to some new solutions*
- *Replace old population with new solutions*
- *Return the best candidate solution found*

D. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) serves as an optimization technique inspired by the collective behavior observed in birds congregating or fish schooling. In the domain of portfolio optimization, PSO is harnessed to amplify returns and curtail risks by progressively enhancing a swarm of prospective solutions, each mirroring a distinct portfolio. Every particle within the swarm adapts its placement (portfolio proportions) grounded on its individual encounters and those of neighboring particles. Initially, particles are dispersed randomly throughout the exploration space. Throughout each iteration, a particle refines its speed and placement based on the most favorable position it has encountered (referred to as personal best) and the optimum position discovered by the swarm (known as global best). The objective function, encompassing projected returns and risk assessments, steers the particles toward optimal resolutions. PSO's efficacy lies in its rapid convergence facilitated by the communal exchange of information, enabling particles to traverse the exploration space efficiently and collaboratively gravitate towards superior solutions. This dynamic adaptation aids in pinpointing portfolios that achieve an optimal equilibrium between maximizing returns and minimizing risks, positioning PSO as a potent instrument for intricate portfolio optimization predicaments.

Algorithm for PSO

Initialize swarm of portfolio weights and velocities

Define fitness function to evaluate negative Sharpe ratio

Identify initial personal best (pBest) for each particle

Identify initial global best (gBest) among all particles

Repeat for a set number of generations or until convergence:

Evaluate fitness of each particle

Update pBest for each particle

Update gBest for the swarm

Update velocity and position of each particle

Ensure particles' positions remain within valid bounds

Return the best position (portfolio weights) found

Figure 4 Correlation matrix

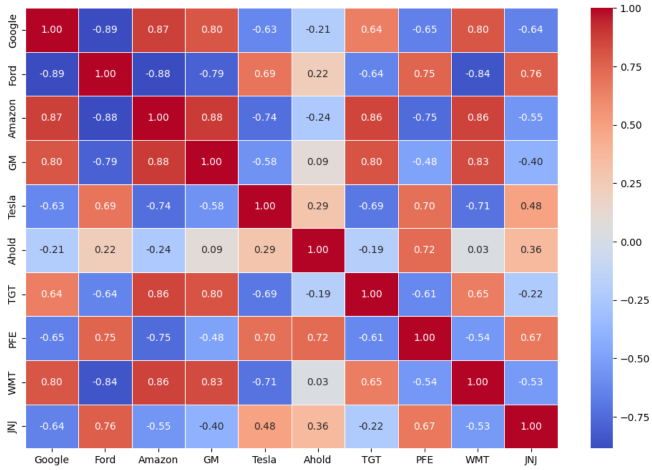

Figure 5. Correlation matrix of the stocks

E. Mean-variance optimization

Mean-Variance Optimization (MVO), originally introduced by Harry Markowitz, serves as a fundamental methodology within contemporary portfolio theory to create an optimal portfolio through the maximization of anticipated return while considering a specific risk level, or the minimization of risk given an anticipated return. MVO operates by taking into account the mean (anticipated return) and variance (risk) of asset returns, requiring the computation of expected returns, variances, and covariances for all assets included in the portfolio. Through the utilization of these parameters, MVO formulates a quadratic optimization conundrum to ascertain the optimal allocation of weights for each asset, effectively managing the balance between return and risk. The primary aim is to pinpoint a portfolio situated on the efficient frontier; wherein no other portfolio offers superior returns at an equivalent risk level. A critical metric within MVO is the Sharpe Ratio, which functions as a pivotal performance gauge, defined as the surplus return of the portfolio over the risk-free rate divided by its standarddeviation. Through the maximization of theSharpe Ratio, investors can pinpoint portfolios that deliver optimal returns in relation to risk exposure. The structured methodology of MVO provides a coherent, mathematically underpinned framework for guiding investment decisions, enabling the methodical evaluation and regulation of return and risk within portfolio development.

IV. RESULTS AND DISCUSSION

This section includes the correlation map among the stocks in the portfolio, the prediction of asset prices through LSTM, and attention-based LSTM. This section includes the analysis of the predicted stock's correlation and also the metrics of price prediction. Finally, the return, risk, and Sharpe ratio analysis through GA, PSO, and MVO algorithms.

A. Correlation matrix among the predicted price of stocks

In this section, the analysis of the correlation of the predicted price of the stocks is illustrated with the correlation matrix **Figure *5***.

The correlation matrix presented yields valuable insights into the interconnections among various stocks within the portfolio. An analysis of the portfolio performance and stock selection reveals certain key considerations:

Elevated Correlations (Approaching 1 or -1):The correlation of 0.87 between Google and Amazon signifies a strong positive relationship, indicating synchronous price movements. The inclusion of both in the portfolio may offer limited diversification advantages due to their likely similar responses to market conditions. Similarly, the correlation of 0.80 between Google and GM also indicates a strong positive relationship, suggesting comparable diversification constraints as seen with Google and Amazon. Conversely, the correlation of -0.88 between Ford and Amazon implies a substantial negative relationship, signifying opposing price movements. Incorporating both in the portfolio could yield significant diversification benefits, potentially reducing overall portfolio risk.

Moderate Correlations (Ranging from 0.5 to 0.75 or -0.5 to -0.75) A correlation of 0.65 between Google and PFE suggests a moderate positive relationship, hinting at some co-movements while still presenting diversification possibilities. On the other hand, the correlation of 0.75 between Ford and PFE showcases a relatively strong positive relationship, which could limit diversification benefits.

Low Correlations (Approaching 0) The correlation of -0.21 between Google and Ahold indicates weak synchronicity in price movements, offering favorable diversification advantages. Similarly, the correlation of -0.19 between Ahold and TGT highlights a low correlation, suggesting potential diversification benefits for a portfolio.

Insights on Stock Selection

Diversification: Optimal diversification of the portfolio involves selecting stocks with low or negative correlations. Pairing Google with Ahold or TGT, for example, can mitigate portfolio risk due to their minimal correlations. High positive correlations, such as those between Google and Amazon or Ford and GM, imply that holding both in substantial proportions could expose the portfolio to heightened systematic risk. It may be more judicious to opt for one from each highly correlated pair.

Risk Management: Integrating stocks with negative correlations, like Ford and Amazon, can aid in risk management. As the value of one stock declines, the other may rise, thereby stabilizing the overall performance of the portfolio. Exercise caution with stocks exhibiting high correlations with multiple counterparts, such as Google, as they have the potential to escalate portfolio volatility.

Maximizing Returns: Stocks with anticipated higher returns but moderate to low correlations with other portfolio elements can enhance the risk-adjusted returns of the portfolio. Evaluating individual expected returns and selecting stocks that strike a balance between potential returns and correlation profiles is advisable.

Balanced Portfolio: A balanced portfolio can be crafted by incorporating stocks from diverse sectors and industries, as they are more likely to display lower correlations. For instance, tech stocks like Google and Amazon, healthcare stocks like PFE, and consumer goods like WMT can offer a well-rounded approach.

In Conclusion, the correlation matrix serves as a fundamental instrument for comprehending the connections among various stocks in the portfolio. To optimize portfolio performance, the focus should be on combining stocks with low or negative correlations to maximize diversification benefits.

B. Price prediction of stocks

This section includes the LSTM and attention-based LSTM for the prediction of asset price. The metrics for the deep learning models for all 10 stocks are given in *Table 1*. The provided table presents the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) for the training, validation, and test datasets of various stocks forecasted utilizing Long Short-Term Memory (LSTM) networks and Attention Mechanisms (ATM). These metrics provide valuable insights into the precision and dependability of the predictive models.

Key Findings:

Tesla:During training, ATM exhibits superior performance, displaying lower RMSE (7.64 vs. 12.56) and MAE (1.35 vs. 1.89). In validation, ATM surpasses LSTM with notably lower RMSE (17.48 vs. 27.97) and MAE (3.32 vs. 4.19). For the test phase, ATM also excels, demonstrating lower RMSE (13.4 vs. 16.41) and MAE (2.99 vs. 3.18).

		Training			Validation			Test		
Stocks		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Tesla	LSTM	12.56	1.89	0.03	27.97	4.19	0.02	16.41	3.18	0.01
	ATM	7.64	1.35	0.02	17.48	3.32	0.16	13.4	2.99	0.01
GM	LSTM	4.83	1.78	0.06	5.41	1.95	0.06	6.88	2.34	0.07
	ATM	3.57	1.47	0.05	12.58	3.42	0.09	12.99	3.46	0.09
FORD	LSTM	0.21	0.33	0.02	0.11	0.28	0.02	0.05	0.18	0.03
	ATM	0.18	0.26	0.02	0.06	0.19	0.02	0.04	0.14	0.02
GOOG	LSTM	0.55	0.59	0.01	1.02	0.81	0.01	110.93	5.61	0.03
	ATM	0.41	0.41	0.01	1.07	0.83	0.01	79.09	4.87	0.03
JNJ	LSTM	6.54	2.03	0.02	51.11	6.49	0.04	200.08	13.73	0.09
	ATM	7.4	2.19	0.02	64.91	7.66	0.05	213.74	14.29	0.09
PFE	LSTM	1.51	0.97	0.03	9.03	2.56	0.07	56.78	7.37	0.26
	ATM	2.06	1.21	0.04	6.02	1.95	0.05	40.21	6.2	0.22
TGT	LSTM	14.46	3.1	0.04	105.82	10.1	0.07	208.75	14.3	0.1
	ATM	7.02	2.15	0.03	27.1	4.86	0.03	93.25	9.47	0.07
AMZN	LSTM	1.79	0.97	0.02	1.75	1.02	0.01	14.82	2.59	0.01
	ATM	0.95	0.63	0.01	1.71	1.04	0.01	4.21	1.37	0.01
WMT	LSTM	0.35	0.46	0.02	1.49	1.18	0.03	38.71	4.93	0.08
	ATM	0.42	0.52	0.02	2.11	1.42	0.03	19.66	3.72	0.06
Ahog	LSTM	0.4	0.52	0.03	1.73	1.21	0.04	4.24	2.04	0.07
	ATM	0.58	0.61	0.04	0.75	0.72	0.03	2.47	1.54	0.06

Table 1 Metrics for the price prediction under LSTM, and Attention-based-LSTM (ATM)

GM:In the training phase, ATM outperforms LSTM, yielding lower RMSE (3.57 vs. 4.83) and MAE (1.47 vs. 1.78). For both validation and test, LSTM performs better, presenting lower RMSE (5.41 vs. 12.58) and MAE (1.95 vs. 3.42). Additionally, LSTM exhibits superior performance in the test dataset.

Ford:Across training, validation, and testing, both models demonstrate comparable performance with minor discrepancies. Generally, ATM showcases marginally improved performance in validation and test datasets.

Google:In the training phase, ATM outperforms LSTM in terms of RMSE and MAE. Regarding validation, both models exhibit similar performance. During the test phase, ATM outshines LSTM, with significantly lower RMSE (79.09 vs. 110.93) and MAE (4.87 vs. 5.61).

JNJ:During training, LSTM slightly outperforms with lower RMSE and MAE. For validation and test, LSTM performs better in validation but similarly in the test set compared to ATM.

PFE:In training, LSTM performs better, achieving lower RMSE (1.51 vs. 2.06) and MAE (0.97 vs. 1.21). For validation and test, ATM excels in validation, while LSTM outperforms in the test dataset.
TGT:Throughout the training, ATM demonstrates enhanced performance with lower RMSE (7.02 vs. 14.46) and MAE (2.15 vs. 3.1). In both validation and test, ATM significantly outperforms LSTM.
AMZN:Across training, validation, and testing, ATM generally exhibits superior performance with notably lower RMSE and MAE.
WMT:In training, LSTM performs better, showing lower RMSE (0.35 vs. 0.42) and MAE (0.46 vs. 0.52). Regarding validation and test, ATM surpasses LSTM with significantly lower RMSE and MAE in the test set.
Ahold:For training and validation, LSTM excels in training while ATM demonstrates slightly superior performance in validation. In the test phase, ATM performs better, showcasing lower RMSE (2.47 vs. 4.24) and MAE (1.54 vs. 2.04).

Summary

- Overall Trend: ATM typically outperforms LSTM in most scenarios, particularly in the validation and test sets, suggesting that attention processes provide more efficient extraction of significant features from the data.
- Stock-specific: Attention-based LSTM performs better in test sets and validation for extremely volatile equities like Tesla and Google. LSTM sometimes performs similarly to more reliable equities like Ford and GM, if not significantly better.

Managerial Insight

- Model Selection: Integrated attention mechanisms (ATM) tend to boost the predictive capabilities of the model, especially in capturing complicated patterns and correlations in the data, whereas long-term support vector machines (LSTM) can offer high baseline performance.

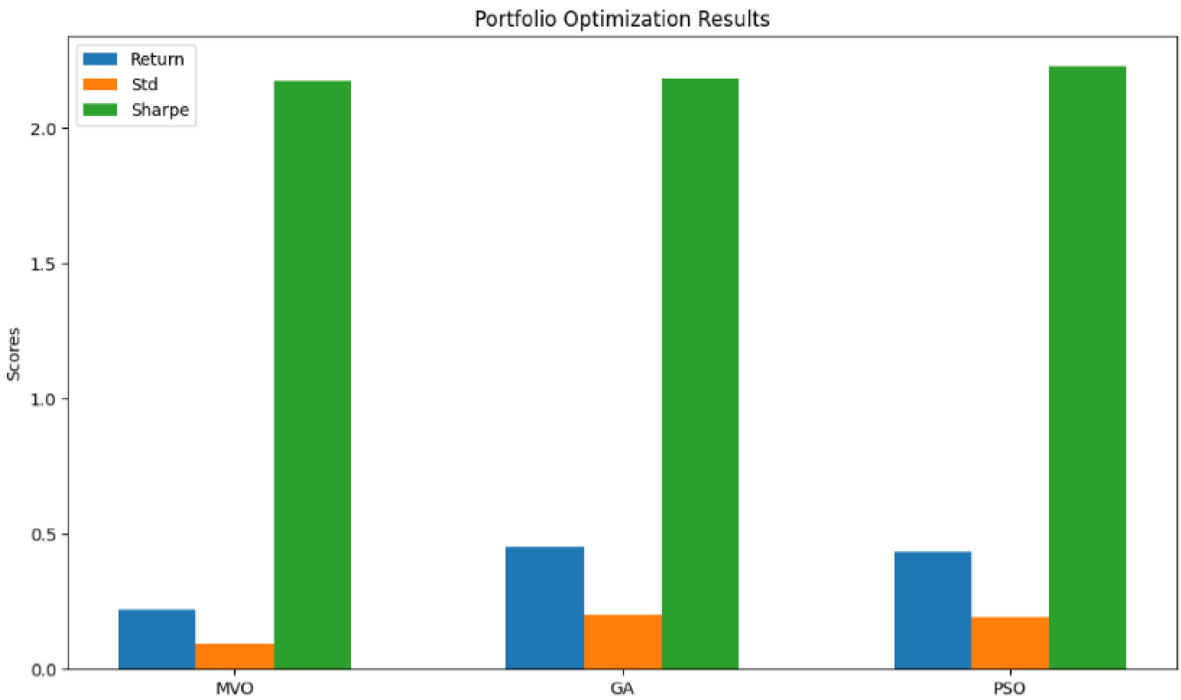

Figure 6 Comparison study of MVO, GA, and PSO return, risk, and Sharpe ratio after LSTM price prediction

- Portfolio management: Making well-informed investment decisions is aided by precise forecasts. Better-performing models for Google and Tesla, for example, can offer information on high-frequency trading tactics.
- Risk management: Reliability in risk assessments is enhanced by reduced prediction errors, which allows for improved hedging tactics and portfolio modifications.
- Resource Allocation: Investing in computing resources and experience to construct ATM

models is worthwhile for improved financial forecasting, as these models often perform better than other models.

C. Optimize, return, risk, and Sharpe ratio

This section introduces the part of the optimization with GA, PSO, and MVO after the price prediction with LSTM and attention-based LSTM or attention mechanism (ATM). The metrics of return, risk, and Sharpe ratio are analyzed with proper insights.

After LSTM price prediction, the return level is almost similar for GA, PSO, and MVO, but there is a slightly higher trend value in GA and PSO than the conventional MVO. The overall risk is found much lower for all three optimization techniques. On the other hand, the higher Sharpe ratio value of PSO indicates that there is a very well-adjusted risk-return balance of the portfolio in the PSO optimization. Even if the PSO is showing a higher Sharpe ratio compared to GA and MVO, all three optimization techniques also show a positive outcome of Sharpe ratio indicating a good construction of the portfolio.

Also, there is a high chance that the portfolio managers may consider the GA, and PSO for portfolio optimization as they are showing marginally better return and higher Sharpe ratio than MVO. **Error! Reference source not found.**shows the comparison study of the three optimization processes after LSTM price prediction.

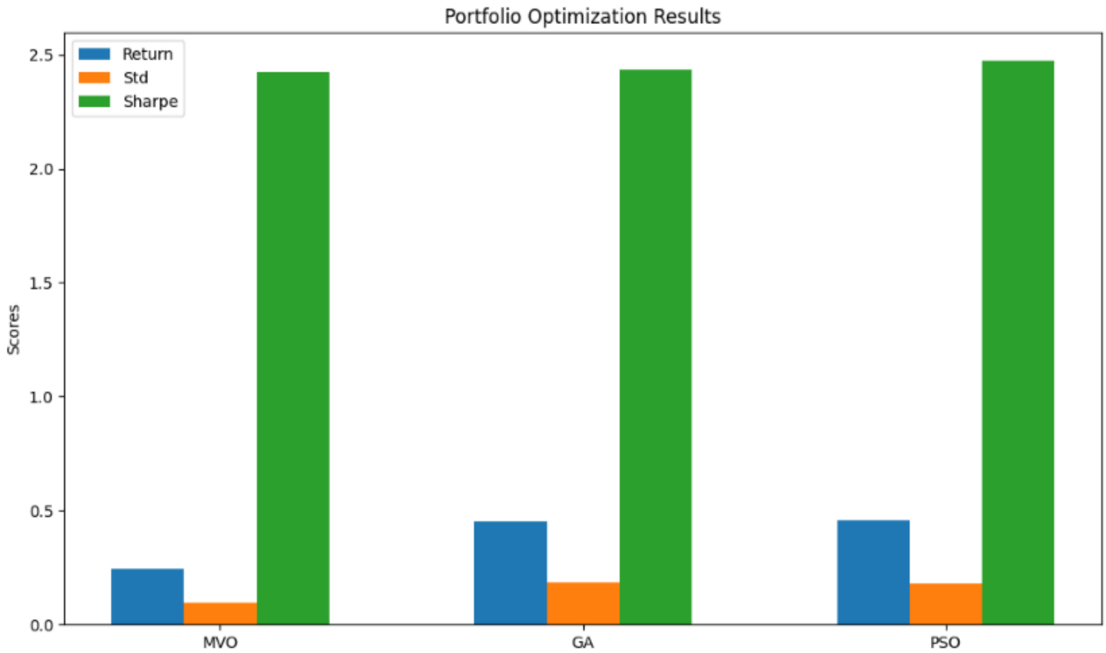

Figure 7. Comparison study of MVO, GA, and PSO return, risk, and Sharpe ratio after attention-based LSTM price prediction

In another optimization experiment after the price prediction using the attention mechanism, GA shows a better return than the PSO and MVO. Similar to the previous model, here also a low standard deviation for all three techniques indicates proper control over risk. Also, regarding the risk-adjusted return, the Sharpe ratio shows a little higher trend for all three optimization techniques. shows the portfolio optimization results for all three techniques with the attention-based price prediction strategy.

In comparison with LSTM, and attention-based LSTM, both GA and PSO have better returns in attention-based pricing strategy. Moreover, MVO has marginally improved return value in attention-based pricing. For risk management, both the deep learning model-based pricing strategies show a controlled behavior. On the other hand, the Sharpe ratio is also higher for the ATM based pricing strategy.

Thus, attention-based pricing leads to better portfolio performance in all senses. Thus, portfolio managers can think of integrating these strategies for better prediction which has a further impact on the later optimization strategies. Thus investing in attention-based predictive modeling will give a well-balanced, low-risk portfolio return for the investors. Table ***2***shows the values of the optimization metrics with GA, PSO, and MVO under LSTM, and attention-based LSTM predictive model.

Table 2 Metrics of portfolio optimization

		Return	Std	Sharpe ratio
LSTM + Attention	MVO	0.24	0.1	2.43
	GA	0.45	0.18	2.43
	PSO	0.46	0.18	2.47
LSTM	MVO	0.22	0.1	2.18
	GA	0.45	0.2	2.19
	PSO	0.44	0.19	2.23

V. Conclusion

This section includes the overview of this study along with the limitations and future scope of the present study.

From this study, the better price prediction and subsequent impact on the optimization is found. Moreover, the heuristics approach like GA and PSO also outperforms the traditional MVO method for optimization technique.

Though the efficacy is found for this price prediction and portfolio optimization through deep learning meta-heuristic approach, the computational complexity may become a barrier for the practitioner to effective implementation of this model during short decision time. Also, the model is run over a limited number of models, so there are many deep learning as well as optimization techniques to run over the models with more complicated scenarios.

In the future scope of this study, some constraints can be added like transaction costs, or liquidity constraints to make the model more realistic. Also to reduce the computational complexity, some advanced versions of GA and other optimization models can be implemented. However, the market movement strategies can also be an important point to be investigated further with the amalgamation of this model with the conditional VAR method.

REFERENCES

[1] L. Gan, H. Wang, Z. Yang, Machine learning solutions to challenges in finance: An application to the pricing of financial products, *Technological Forecasting and Social Change*, 153, 2020, 119928.

[2] R. Soujanya, P.A. Goud, A. Bhandwalkar, G.A. Kumar, Evaluating future stock value asset using machine learning,*Materials Today: Proceedings*, 33, 2020, 4808–4813.

[3] F. Xu, S. Tan, Deep learning with multiple scale attention and direction regularization for asset price prediction, *Expert Systems with Applications*, 186, 2021, 115796.

[4] E. Hossain, M.S. Hossain, P.-O. Zander, K. Andersson, Machine learning with Belief Rule-Based Expert Systems to predict stock price movements, *Expert System Applications*, 206, 2022, 117706.

[5] H. Yao, S. Xia, H. Liu, Six-factor asset pricing and portfolio investment via deep learning: Evidence from Chinese stock market, *Pacific-Basin Finance Journal*, 76, 2022, 101886.

[6] S. Pan, S.C. Long, Y. Wang, Y. Xie, Nonlinear asset pricing in Chinese stock market: A deep learning approach, *International Review of Financial Analysis*, 87, 2023, 102627.

[7] G.A.V. Pai, T. Michel, Metaheuristic multi-objective optimization of constrained futures portfolios for effective risk management, *Swarm and Evolutionary Computation*,19, 2014, 1–14.

[8] J. Doering, R. Kizys, A.A. Juan, A. Fito, O. Polat, Metaheuristics for rich portfolio optimisation and risk management: Current state and future trends, *Operations Research Perspectives*, 6, 2019, 100121.

[9] R. Kizys, J. Doering, A.A. Juan, O. Polat, L. Calvet, J. Panadero, A simheuristic algorithm for the portfolio optimization problem with random returns and noisy covariances, *Computers & Operations Research*, 139, 2022, 105631.

[10] M. Nssibi, G. Manita, O. Korbaa, Advances in nature-inspired metaheuristic optimization for feature selection problem: A comprehensive survey, *Computer Science Review*, 49, 2023, 100559.

[11] X. Zheng, C. Zhang, B. Zhang, A Mayfly algorithm for cardinality constrained portfolio optimization, *Expert System Applications*,230, 2023, 120656.

[12] J.-S. Chou, K.-E. Chen, Optimizing investment portfolios with a sequential ensemble of decision tree-based models and the FBI algorithm for efficient financial analysis, *Applied Soft Computing*,158, 2024, 111550.