Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/61/20231063

Dynamic Optimization of Investment Portfolio Based on Self-Attention Mechanism

- LSTM Model Prediction

Songwei Wu[1,a,*]

[1]*Faculty of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, China*
a. q030018082@mail.uic.edu.cn
**corresponding author*

Abstract: In the era of big data and advanced computational capabilities, financial market participants are continuously searching for innovative strategies to gain a competitive edge. A potential pathway emerges in the domain of deep learning, especially with regard to the LSTM neural architectures, which are renowned for their ability to handle and make predictions based on time series data. This study delves into the utilization of LSTM for predicting stock prices, emphasizing the advantages of dynamic investment portfolios in the rapidly fluctuating market conditions. Utilizing a dynamic window approach for time series data preprocessing, a self-attention mechanism - LSTM model was designed to anticipate the tendency of annual closing prices for five stocks from 2022 to 2023, Utilizing the initial 80% of stock price data as training set and allocating the residual 20% for validation. The performance of the dynamic optimization portfolio model was assessed by dynamically adjusting the weights of the stocks based on the last 20% of the data, and was subsequently compared to actual market cumulative returns. The findings indicate not only that the LSTM model offers a commendable level of accuracy in predicting stock prices, but also that the recursive algorithm for the dynamic optimization portfolio, constrained by maximum returns and minimal standard deviation, consistently outperforms the general market.

Keywords: Long Short Term Memory (LSTM), Dynamic Investment Portfolios, Dynamic Window Approach, Neural Networks, Recursive Algorithm

1. Introduction

In recent years, the landscape of financial investment has undergone a transformative shift, embracing machine learning technologies to bolster the precision and efficacy of portfolio management strategies. Machine Learning has recently shown significant progress in forecasting time series across diverse socio-economic metrics [1]. A focal point of this transition is the deployment of deep learning models, notably the LSTM neural networks, tailored for financial statistics forecasting and consequent portfolio optimization. In recent years, LSTM has become a popular model for data analysis, widely used for predicting financial data. People also combine the most basic LSTM model with other machine learning methods. The ARIMA-WOA-LSTM outperforms single models when it comes to forecasting accuracy, model precision, and stability in predicting pollutants [2]. Metaheuristic strategies, like the Artificial Rabbits Optimization algorithm (ARO), can enhance the precision of stock market forecasts by fine-tuning the hyperparameters of an LSTM model [3]. Compared with the CNN-LSTM model, the BiLSTM-CNN method achieves the most advanced accuracy in the field of facial recognition [4].

While conventional investment strategies predominantly hinge on historical price data and statistical analytics, they often falter in grasping the intricate dynamics and evolving nature of today's financial markets. This limitation underscores the surge in adopting sophisticated machine learning techniques for refined portfolio management.

LSTM stands out as a specialized variant of recurrent neural networks (RNN) with a knack for adeptly handling sequences exhibiting long-term dependencies. Its prowess in deciphering intricate patterns in sequential data positions it as a frontrunner for forecasting market movements [5]. Leveraging LSTM's capabilities facilitates astute short and long-term financial forecasts, laying the groundwork for judicious investment choices.

In this paper, an innovative approach is introduced that employs an LSTM-based price prediction model for dynamic portfolio optimization. The method consists of two phases: Initially, an LSTM model, uniquely tailored with a dynamic window, updates predictions using 80% of the data and incrementally tests the subsequent 20% to forecast asset prices within a specified forward-looking window. Subsequently, these predictions guide the dynamic reallocation of assets within the portfolio. By doing so, the strategy of this research aims to leverage the trajectory of emerging markets, readjust the composition of the portfolio, and potentially surpass returns achieved through traditional static portfolio management techniques.

2. Background knowledge

2.1. Markowitz's Modern Portfolio Theory

2.1.1. Background Introduction

The concept of portfolio theory, introduced by Harry Markowitz in 1952, serves as a cornerstone for contemporary financial practices [6]. The core viewpoint of this theory is that investors should not only focus on expected returns, but also consider risks. Markowitz's concept suggests a strategy for determining investment portfolios that curtails risk for a set projected return or elevates projected return for a designated risk threshold.

2.1.2. Mathematical Forms

Given a combination of n assets, each asset i has an expected return rate r_i and standard deviation σ_i. The weight of the asset is ω_i, which satisfies formula $\sum_{i=1}^{n} \omega_i = 1$. Expected return rate R_p and standard deviation of investment portfolio σ_p can be represented as:

$$R_p = \sum_{i=1}^{n} \omega_i r_i \tag{1}$$

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \omega_j \sigma_{ij} \tag{2}$$

DOI: 10.54254/2754-1169/61/20231063

Among them, σ_{ij} Is the covariance of asset i and asset j.

2.1.3. Optimization problems

Based on the above formula, the portfolio optimization can be formulated as:

$$subject\ to\ R_p \geq R_{target} \tag{3}$$

$$\sum_{i=1}^{n} \omega_i = 1 \tag{4}$$

Here, R_{target} is the expected target rate of return. Portfolio theory provides investors with a quantitative method to balance returns and risks, and to choose the best combination of assets. The two methods used in this article are closely related to this theory, but focus on minimizing risk and maximizing return, respectively. By using constrained optimization methods, the optimal investment portfolio that meets various conditions can be found. These two methods will be explained in detail in Section 5.2.

2.2. Introduction to LSTM Neural Network

2.2.1. Background Introduction

The LSTM architecture, a distinct variant of RNN, was pioneered by Sepp Hochreiter and Jürgen Schmidhuber in the late 1990s. Its design excels in analyzing and predicting noteworthy occurrences spanning vast chronological intervals [7]. Due to its unique structural design, LSTM can effectively learn, store, access, and utilize patterns and dependencies in long-term sequence data, making it particularly effective in processing sequence data containing long-distance dependencies (such as time series, text, etc.) [8].

2.2.2. Structure and Operations

The fundamental architecture of LSTM revolves around its "memory cells" or just "cells". Each cell incorporates three essential gates, and these gates dictate how cells accept, retain, and dispatch data.

Forget Gate: It decides how much information the cell state should keep or discard.

Input Gate: It dictates the extent to which new data should be updated in the cell state.

Output Gate: It decides the segment of the cell state to be outputted. Data and preprocessing

2.3. Data sources and their reasons for selection

For this research, information from Yahoo Finance was utilized to pick the closing stock prices of five firms: Apple (AAPL), Microsoft (MSFT), Google (GOOGL), Amazon (AMZN), and Meta (META), as the research subjects for the past year. Because they represent the forefront and multiple key sub areas of the technology industry, such as hardware, software, the internet, social media, and e-commerce. These companies not only have huge market capitalization and a significant position in the global economy, but their stocks also have high liquidity and extensive analysis coverage, providing rich and reliable data sources for this study.

2.4. Data cleaning and preprocessing steps

Handling missing and outliers: Any missing data caused by transaction interruptions or data acquisition issues needs to be identified and processed. Fortunately, there were no outliers or missing values in the dataset of this study;

2.5. Statistical characteristics and analysis of data

Calculate daily return: For each stock, the daily return in this study is calculated using difference return (See Figure 1).

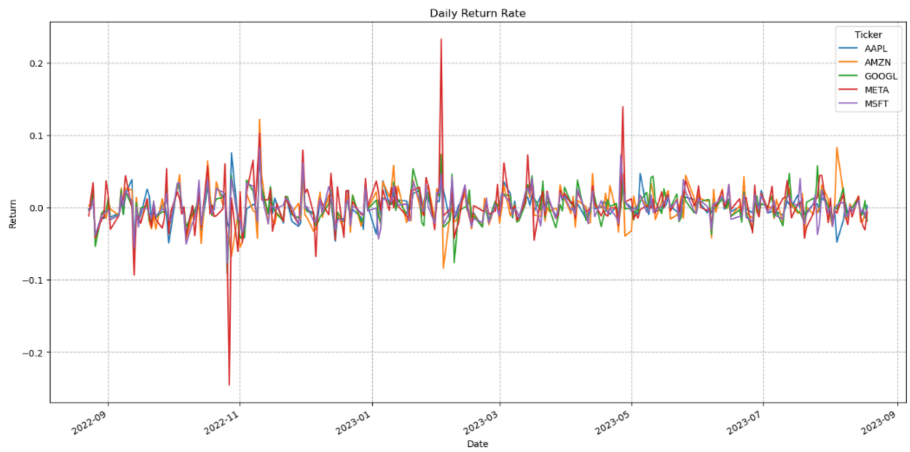

Figure 1: Daily Return Rate

Normalized data: To enhance the LSTM model's efficacy, data normalization is implemented (See Figure 2).

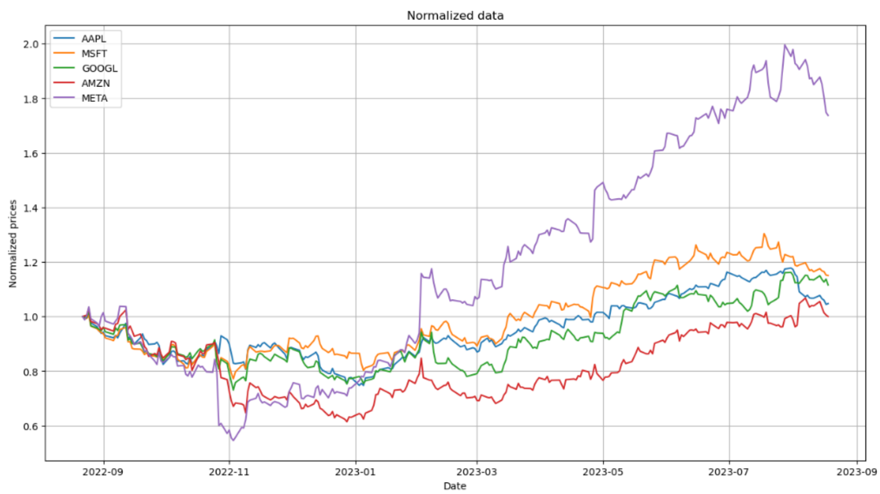

Figure 2: Normalized data

3. Methodology and Results

3.1. LSTM model design and parameter selection

3.1.1. Model architecture

The LSTM model in this study uses a single LSTM layer as the main body, followed by a fully connected output layer. The model architecture has been widely recognized as effective in handling multivariate time series data, especially stock price data.

LSTM layer: The LSTM layer has 50 units and uses the relu activation function. The number of units chosen seeks a compromise between model intricacy and computational speed, while the relu function was chosen as the activation function due to its superior performance in deep learning models.

Dense layer: This is a densely connected layer, with the number of neurons matching the expected count of stocks. It transforms the output from the LSTM layer into predictive results that correspond to the target variable.

3.1.2. Compilation and Training

Optimizer: The 'adam' has been chosen as the optimizer because it is effective for most deep learning tasks in practice and usually converges quickly.

Loss Function: Employ the MSE for the loss function because it quantifies the difference between predicted and true values, especially for regression problems, which are the most commonly used.

Training Cycle and Batch Size: The model trained 100 cycles with a batch size of 32. The selection of these parameters is based on the preliminary experimental results to ensure that the model can effectively converge while maintaining a balance between computational efficiency and model performance.

3.1.3. Optimized model

Based on the preliminary LSTM model, to enhance the model's efficacy and better discern intricate trends within stock price, a self-attention mechanism was incorporated. This mechanism integrating the concept of uncertainty into deep learning models and has shown excellent performance in multiple temporal data processing tasks, particularly within natural language processing [9].

The optimization model combines LSTM and multi head self-attention layers, integrating these two powerful mechanisms into a unified framework. Compared to the previous model, two new layers have been added:

Multi head self-attention layer: This layer uses two heads and a 2D key to process the output of LSTM. The multi head self-attention mechanism can capture different levels of dependency relationships in data, enhancing the model's ability to recognize subtle patterns in the data.

Merge Layer: The outputs of LSTM and multi head self-attention are merged to ensure that the model can fully utilize the information captured by both mechanisms. Detailed results for LSTM predictions are shown in the following Figure 3.

DOI: 10.54254/2754-1169/61/20231063

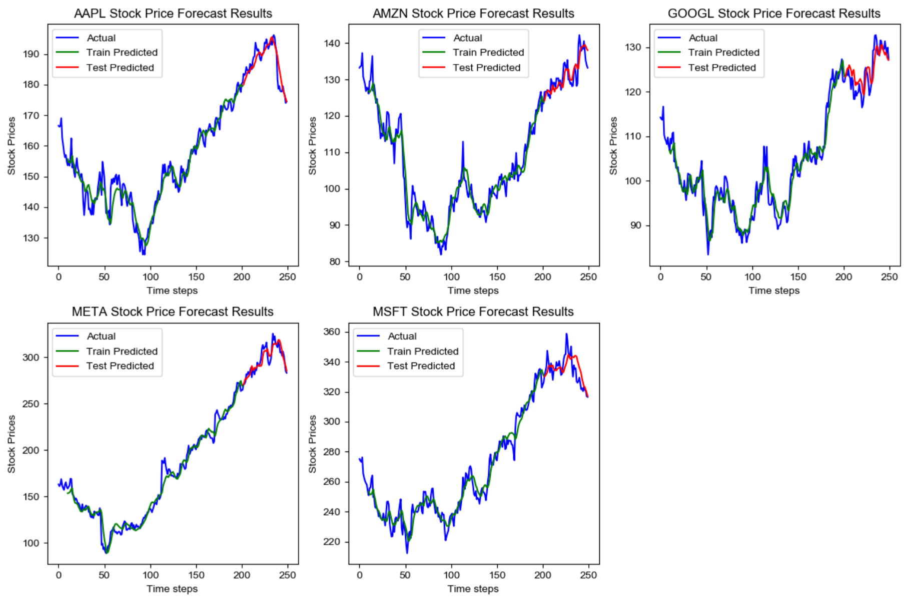

Figure 3: LSTM Forecast results

3.2. Dynamic Portfolio Optimization Strategy

3.2.1. Algorithm Overview

Dynamic portfolio optimization is the process of adjusting the weight of an investment portfolio based on the latest market data within a fixed time window. This strategy is different from the traditional "set and forget" strategy, as it continuously optimizes based on new market information. This study reassesses and adjusts the weight of investment portfolios based on daily stock price changes. A recent study introduced an interactive portfolio optimization technique using SNGDM-PO, considering dynamic trust of experts, risk preferences, consensus feedback mechanisms, and reduced interaction costs, ensuring effective consensus achievement and high group satisfaction within limited interactions [10].

3.2.2. Optimization objectives

The primary aim of this article is to mitigate investment portfolio risks. The risk of an investment portfolio is calculated based on its asset weight and covariance matrix. Use the following formula:

$$\sigma_p = \sqrt{\omega^T \times \sum \times \omega} \tag{5}$$

Among them:
- σ_p is the standard deviation (risk) of the investment portfolio.
- ω is the weight of each asset in the investment portfolio.
- $\sum$ is the covariance matrix.

3.2.3. Constraints

When optimizing, added a constraint condition to ensure that the sum of weights in the investment portfolio is 1. This ensures that all funds are invested and there is no surplus.

3.2.4. Dynamic optimization process

In order to dynamically optimize the investment portfolio, a sliding window method was used. In each step, new market data was added, and the return rate and covariance matrix were recalculated. Next, use the 'minimize' function to find the optimal asset weight, thereby minimizing the risk of the investment portfolio.

3.2.5. Portfolio Weight

Finally, the optimization process of each step obtained the optimal weights and saved them in a data box. This data box provides the optimal weight of the investment portfolio at each time point (See Figure 4).

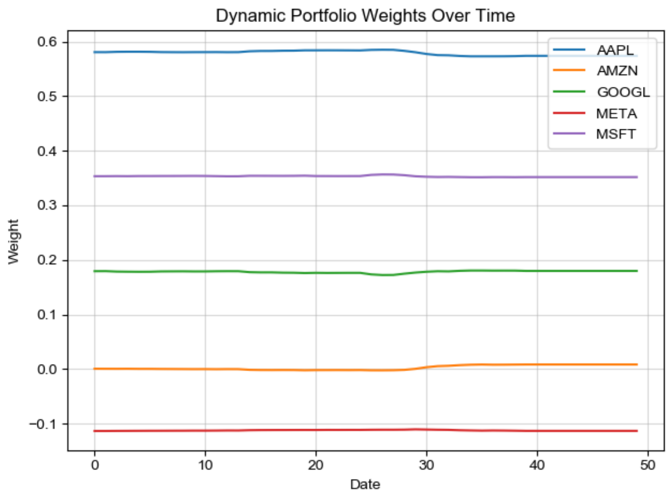

Figure 4: Dynamic Portfolio weights over Time

3.2.6. Comparison of cumulative returns between investment portfolio and S&P 500

To better evaluate the strategy, the daily return of the investment portfolio based on optimal weights was first calculated. This is achieved by multiplying the daily yield by the corresponding weight. Subsequently, the daily cumulative yield was calculated, which was obtained by continuously multiplying the daily yield.

At the same time, the cumulative yield of S&P 500 was also calculated as a reference benchmark. The S&P 500 is a commonly used stock market index that represents the broad performance of the US stock market. Comparing with investment portfolios can provide us with a clear reference point to indicate whether the strategy has exceeded the market average [11].

3.2.7. Visualization

To compare the effect of these two strategies, drew a figure showing the cumulative return of the investment portfolio and S&P 500 over time (See Figure 5).

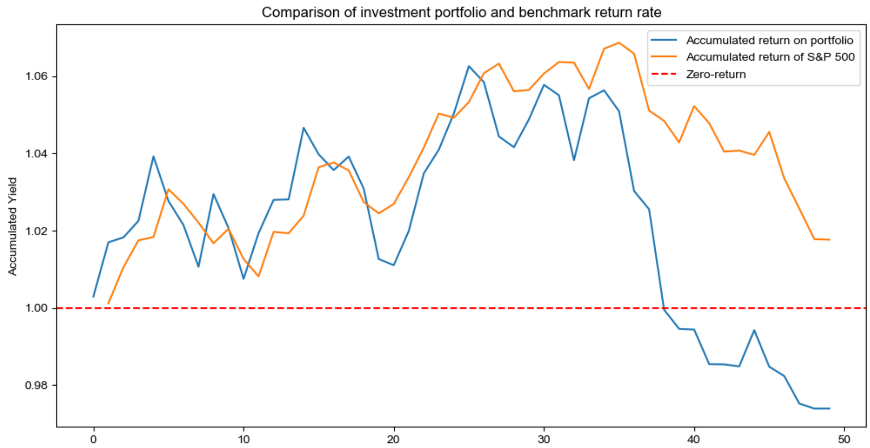

Figure 5: Comparison of investment portfolio and benchmark return rate

From the figure, it can be seen that compared to the S&P 500, the dynamic optimization strategy maintained relatively stable and similar volatility and cumulative returns within the first 20 days but did not show a significant trend higher than market returns. Over time, the market has shown a trend of sustained returns, while the cumulative returns of the investment portfolio have shown a cliff like decline, with continuous losses starting from the 38th day. This is because the limitation of the investment portfolio is to minimize volatility, rather than maximizing portfolio returns. Next, optimize the model and change the restrictions to observe the subsequent results.

3.3. Improved optimized investment portfolio

Additional Constraints:

Weight boundary: Limit the weight of each asset to between -1 and 1. This can be achieved through the bounds parameter in the minimize function.

Leverage limit: Restrict the total weight to 1, ensuring no excessive investment or borrowing.

3.3.1. Objective Function

$$\max R(\omega) = -\omega^T \mu \tag{6}$$

Among them:

- $R(\omega)$ is the expected return.
- ω is the weight vector of each asset.
- μ is the expected return vector of each asset.

3.3.2. Constraints

$$\sum_{i=1}^{N} \omega_i = 1 \tag{7}$$

For each ω_i, there is $-1 \leq \omega_i \leq 1$

The first constraint ensures that the total weight is 1, and the second constraint sets upper and lower bounds for short selling.

DOI: 10.54254/2754-1169/61/20231063

3.3.3. Optimize strategy implementation

Using the above model, the optimal daily weights can be calculated based on historical data. When implementing strategies, first calculate daily returns, and then dynamically adjust asset weights based on these returns.

At each time step, new data is added to the historical dataset and the optimal weight of the investment portfolio is recalculated based on this. In this way, the strategy can be adjusted based on the latest market conditions. Related results are shown in the following Figure 6.

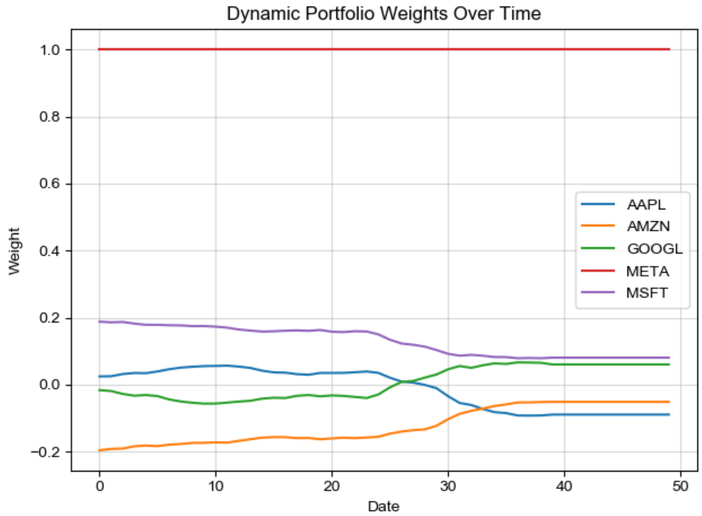

Figure 6: Optimized Dynamic Portfolio Weights Over Time

3.3.4. Comparison Visualization

From Figure 7, it can be seen that the dynamic optimization strategy performs better in certain periods compared to the S&P 500. This demonstrates the importance of dynamic weight adjustment, especially in the face of market fluctuations. In addition, the zero-return line represented by the red dashed line also provides a reference point to help us understand the performance of investment portfolios and markets at any time.

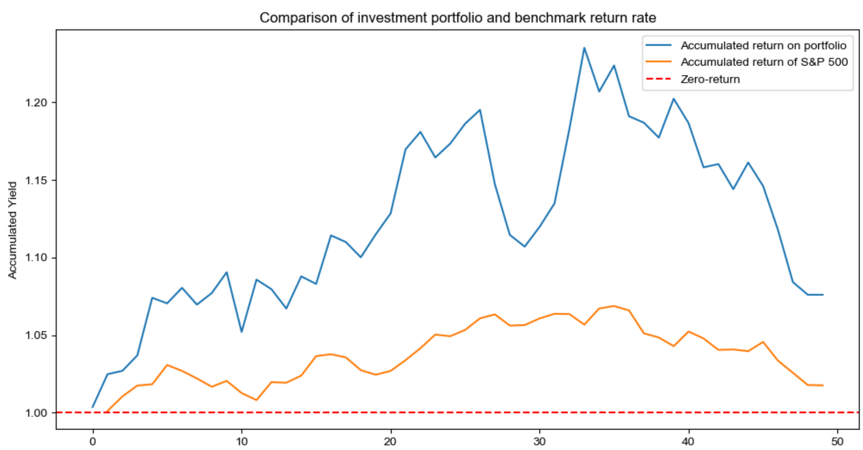

Figure 7: Comparison of cumulative returns after optimization

4. Conclusion and Future outlook

The study explored the use of LSTM-based models for predicting stock prices, and based on these predictions, a dynamically optimized investment portfolio was constructed for investors. By directly comparing with the S&P 500 index, the dynamic portfolio optimization strategy outperformed during certain periods. This not only validates the efficacy of LSTM models in predicting stock prices but also underscores the importance of dynamically adjusting portfolio weights. Of course, this strategy also requires continuous updating and optimization to adapt to the ever-changing market environment.

Although the methods in this study have achieved satisfactory results and provide certain suggestions for financial practitioners, there are still some potential areas for improvement:

Further optimization of the model: Although LSTM performs well in processing time series data, it is still possible to improve prediction accuracy by introducing more advanced models or structures.

More data sets and validation: In order to comprehensively evaluate strategies, validation can be considered across more financial markets and time periods.

Integration of Other Financial Indicators: When constructing an investment portfolio, in addition to expected returns, other financial indicators such as liquidity and market value can also be considered.

Further Research on Risk Management: Future research can further explore how to better balance risk and return in strategies.

References

[1] Grossman, I., Wilson, T., and Temple, J. (2023). Forecasting small area populations with long short-term memory networks. Socio-Economic Planning Sciences, 101658.

[2] Luo, J., and Gong, Y. (2023). Air pollutant prediction based on ARIMA-WOA-LSTM model. Atmospheric Pollution Research, 14(6), 101761.

[3] Gülmez, B. (2023). Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. Expert Systems with Applications, 227, 120346.

[4] Febrian, R., Halim, B. M., Christina, M., Ramdhan, D., and Chowanda, A. (2023). Facial expression recognition using bidirectional LSTM - CNN. Procedia Computer Science, 216, 39-47.

[5] Bhandari, H. N., Rimal, B., Pokhrel, N. R., Rimal, R., Dahal, K. R., and Khatri, R. K.C. (2022). Predicting stock market index using LSTM. Machine Learning with Applications, 9, 100320.

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

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

[8] Wijnberg, K. M., Kilsdonk, R. A. H., and Bomers, A. (2022). Predicting urban flooding due to extreme precipitation using a long short-term memory neural network. Hydrology, 9(6), 105.

[9] Park, J., Lee, K., Park, N., You, S. C., and Ko, J. (2023). Self-Attention LSTM-FCN model for arrhythmia classification and uncertainty assessment. Artificial Intelligence in Medicine, 142, 102570.

[10] Li, D., and Hu, S. (2023). Adaptive consensus reaching process with dynamic weights and minimum adjustments for group interactive portfolio optimization. Computers & Industrial Engineering, 183, 109491.

[11] Pal, R., Datta Chaudhuri, T., and Mukhopadhyay, S. (2021). Portfolio formation and optimization with continuous realignment: A suggested method for choosing the best portfolio of stocks using variable length NSGA-II. Expert Systems with Applications, 186, 115732.