Metaheuristic Algorithms with Applications

www.maa.reapress.com

Metaheur. Algor. Appl. Vol. 2, No. 3 (2025) 229–235.

Paper Type: Original Article

Dynamic Portfolio Optimization in the Tehran Stock Exchange Using Machine Learning and Deep Earning Algorithms

Hossein Pira*

Department of Computer Engineering, Ayandegan University, Tonkabon, Iran; hossein.pira@aihe.ac.ir.

Citation:

Received: 25 February 2025 *Revised: 17 April 2025* *Accepted: 02 June 2025*	Pira, H. (2025). Dynamic portfolio optimization in the tehran stock exchange using machine learning and deep earning algorithms. *Metaheuristic algorithms with applications, 2*(3), 229-235 .

Abstract

Classical optimization models are applicable for portfolio optimization when capital market constraints and portfolio construction criteria are limited. However, given the existing complexities in capital markets, intricate relationships, and non-linear issues, classical models exhibit significant weaknesses and limitations. To address and solve such problems, computers and Machine Learning (ML)-based models have emerged to assist human decision-makers. Recently, techniques based on hybrid ML models have been designed, demonstrating robust capabilities in handling non-linear problems. This paper proposes a hybrid ML model for portfolio optimization based on specific historical data in the capital market. For this purpose, daily returns from 20 high-liquidity companies listed on the Tehran Stock Exchange (TSE) over a specific timeframe were utilized. The proposed framework integrates Random Forest (RF) for feature importance analysis and Extreme Gradient Boosting (XGBoost) for return prediction, combined with a Deep Reinforcement Learning (DRL) agent using Proximal Policy Optimization (PPO) for dynamic asset allocation. Portfolio performance is evaluated across four distinct risk metrics: Variance, Mean Absolute Deviation (MAD), Semi-Variance, and Conditional Value at Risk (CVaR). Based on evaluations of portfolio return, risk indices, and other key indicators, the results of this research indicate that the performance of the proposed ML model is at an optimal and desirable level.

Keywords: Portfolio optimization, Machine learning, Deep reinforcement learning, Tehran stock exchange, Random forest, XGBoost, Risk management.

1 | Introduction

The stock market, as one of the fundamental pillars of the global economic system, plays a critical role in the optimal allocation of financial resources to enterprises and the channeling of liquidity toward productive sectors. In recent years in Iran, the Tehran Stock Exchange (TSE) has emerged as one of the most significant

Corresponding Author: hossein.pira@aihe.ac.ir
https://doi.org/10.48314/isti.v2i3.44

investment platforms; however, operating in this market entails unique challenges [1]. Numerous structural characteristics such as price limits (trading bands), "base volume" requirements, widespread information asymmetry, and investors' emotional behavior coupled with the market's acute and rapid sensitivity to political news and macroeconomic variables have created a volatile and complex environment. In such a landscape, the precise analysis and prediction of stock returns have become a primary concern for market participants and financial researchers.

Historically, Modern Portfolio Theory (MPT) and Harry Markowitz's Mean-Variance model have been the most influential frameworks for portfolio management and asset allocation [2]. While these models offer a systematic approach to balancing risk and return, they rely on rigid assumptions that often conflict with the realities of the TSE. By assuming a normal distribution of returns and linear relationships between assets in a market characterized by skewness, non-stationarity, and chaotic behavior, these classical methods can lead to inaccurate risk assessments and inefficient portfolios [3].

In recent years, significant advancements in the fields of Artificial Intelligence (AI) and Machine Learning (ML) have established a new paradigm in financial markets and portfolio optimization. Unlike traditional models, ML algorithms such as Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Deep Neural Networks (DNN) offer high flexibility in discovering hidden patterns and non-linear relationships [4]. These algorithms are capable of processing vast amounts of time-series data, as well as qualitative data, to predict severe market fluctuations with greater accuracy [5].

This research adopts an innovative approach by coupling the predictive power of ML algorithms with the capabilities of Deep Reinforcement Learning (DRL) to perform dynamic optimization. The XGBoost algorithm is used for stock selection and return prediction, while the RF algorithm captures non-linear features for price prediction. Subsequently, a DRL agent using Proximal Policy Optimization (PPO) dynamically allocates portfolio weights based on predicted returns and risk considerations [6].

2 | Related Works

2.1 | Portfolio Optimization Theory

The concept of portfolio optimization was formalized by Harry Markowitz [7], introducing the Mean-Variance framework that considers risk and return simultaneously [2]. A central theme is selecting an optimal investment portfolio relative to its risk-return profile. Markowitz demonstrated that portfolio risk depends not only on the variance of individual assets but also on the correlation between their returns. Therefore, two high-risk stocks, when combined, can create a lower-risk portfolio provided their correlation is low. The primary objective is to find specific asset weights that minimize portfolio variance while maximizing expected return.

2.2 | Machine Learning in Finance

RF, introduced by Breiman [8], is recognized as one of the most robust ML algorithms for both classification and regression tasks. By utilizing multiple decision trees and employing ensemble learning techniques, this algorithm can model complex and non-linear patterns with high precision. Its stability, resistance to noise, and ability to handle outliers demonstrate exceptional performance in analyzing complex financial data in volatile markets.

XGBoost, introduced by Chen and Guestrin [9], has gained significant attention in the financial sector due to its ability to process and predict data within an optimal timeframe. Its sequential gradient boosting hierarchy and optimized decision trees capture complex, non-linear interactions and cross-correlations among financial variables more effectively than classical models. For stock return forecasting, volatility prediction, trend identification, and risk analysis, XGBoost has demonstrated superior performance.

2.3 | Deep Reinforcement Learning for Portfolio Management

DRL has emerged as a powerful framework for sequential decision-making in financial markets [6]. Unlike supervised learning approaches that predict returns independently, DRL agents learn optimal policies by interacting with the market environment, receiving rewards based on portfolio performance. PPO provides stable policy updates with guaranteed monotonic improvement, making it particularly suitable for the non-stationary dynamics of financial markets [10].

3 | Research Methodology

3.1 | Data Collection and Stock Selection

The statistical database for this study comprises companies listed on the TSE between the years 2013 and 2022 (1392 to 1401 SH). Following data collection and preliminary calculations, 20 high-liquidity companies were selected based on criteria such as liquidity, trading volume, number of transactions, and market impact. Daily adjusted returns were calculated for all selected companies.

3.2 | Feature Engineering

The feature engineering process involves extracting relevant financial indicators from raw price and volume data:

I. Daily returns and log returns

II. Moving averages (5-day, 10-day, 20-day, 50-day)

III. Relative Strength Index (RSI)

IV. Moving Average Convergence Divergence (MACD)

V. Bollinger Bands (upper and lower)

VI. Volume-Weighted Average Price (VWAP)

VII. Historical volatility measures

3.3 | Stock Selection Using Extreme Gradient Boosting

The stock selection process is conducted in several stages. First, the XGBoost model is trained to predict future returns of all stocks within the research sample. Subsequently, the predicted return value is calculated for each individual stock. Stocks are ranked based on the XGBoost model's output, and only stocks situated in the upper percentiles of the predicted return distribution are considered as selected stocks. Furthermore, XGBoost allows for the extraction of Feature Importance, identifying which technical indicators played the most significant role in predicting stock returns.

The XGBoost objective function consists of two primary components: the loss function (prediction error) and the regularization term controlling model complexity:

$$\mathcal{L}(\varphi) = \sum i * l(\hat{y}i, yi) + \sum k\Omega(fk), \tag{1}$$

where l is the loss function measuring the difference between the predicted value $\hat{y}_i$ and the actual value y_i, and Ω is the regularization term preventing overfitting by penalizing model complexity.

3.4 | Return Prediction Using Random Forest

The RF algorithm complements XGBoost by capturing non-linear features for price prediction. Each decision tree is constructed by classifying bootstrap samples of the input data. The ensemble output is determined based on the votes of each tree, providing robust predictions resistant to noise and outliers.

3.5 | Deep Reinforcement Learning for Portfolio Optimization

The optimization algorithm is based on DRL using PPO. The DRL framework is defined as follows:

I. State space: the state at time t includes predicted returns from ML models, current portfolio weights, historical volatility measures, and market indicators.

II. Action space: the action is a vector of portfolio weights $w = (w1, w2, \dots, wN)$ where $\sum wi = 1 \text{ and } wi \geq 0$ for all i.

III. Reward function: the reward is defined as the risk-adjusted portfolio return, incorporating the selected risk metric (Variance, Mean Absolute Deviation (MAD), Semi-Variance, or Conditional Value at Risk (CVaR)).

The PPO agent learns to dynamically adjust portfolio weights through continuous interaction with historical market data, optimizing the trade-off between return maximization and risk minimization. The PPO objective function is given by:

$$\mathrm{LCLIP}(\theta) = \mathbb{E}t\left[\min\left(rt(\theta)\hat{A}t, \mathrm{clip}(rt(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}t\right)\right], \quad (2)$$

where $r_t(\theta)$ is the probability ratio between the new and old policy, $\hat{A}_t$ is the estimated advantage function, and ε is the clipping hyperparameter.

3.6 | Risk Measurement Metrics

Four distinct risk metrics are employed for comprehensive risk assessment:

Variance: the variance of returns is computed as

$$\sigma i2 = \sum(ri - \bar{r}) * (2/N), \quad (3)$$

representing the degree of dispersion of actual returns around the expected return.

MAD:

$$\mathrm{MAD} = \sum| ri - \bar{r} | / N, \quad (4)$$

which is less sensitive to outliers than variance, allowing for more simplified calculations within large-scale optimization models.

Semi-Variance:

$$\mathrm{SV} = \mathbb{E}\{\min(ri - \bar{r}, 0) * 2\}, \quad (5)$$

focusing only on downside risk by considering only returns that fall below the mean.

CVaR: addresses the limitations of standard VaR by accounting for the severity of losses in the tail of the distribution, defined as the expected loss given that the loss exceeds the VaR threshold:

$$\mathrm{CvaR}\alpha = \mathbb{E}[-R \mid R \leq -\mathrm{VaR}\alpha]. \quad (6)$$

Table 1. Descriptive statistics of selected companies.

No.	Company Name	Average Returns	Standard Deviation
1	Iran transfo	0.009288	0.032070
2	Iran khodro	0.004252	0.019304
3	Saderat bank	0.001322	0.016102
4	Mellat bank	0.010126	0.016017
5	Development of mines and metals	0.027382	0.017267

Table 1. Continued.

No.	Company Name	Average Returns	Standard Deviation
6	Saipa	0.008687	0.017534
7	National development investment	0.004915	0.024500
8	Golgohar mining and industrial	0.078161	0.020154
9	Pension fund investment	0.015289	0.024682
10	Ghadir investment	0.006221	0.015743
11	Oil industry investment	0.005684	0.023895
12	Azarab industries	0.001331	0.017082
13	Khuzestan steel	0.000266	0.019140
14	Isfahan Mobarakeh steel	0.003971	0.023587
15	Bahman group	0.006090	0.027841
16	MAPNA group	0.001662	0.016416
17	Ryan Saipa leasing	0.006160	0.024007
18	Iranian intelligence	0.001240	0.022289
19	Chadormalu mining and industrial	0.004632	0.024358
20	National Iranian copper industries	0.013211	0.026646

4 | Results and Analysis

4.1 | Model Training and Validation

The dataset was divided into training (2013–2019) and testing (2020–2022) periods. Both XGBoost and RF models were trained on the training set with cross-validation to prevent overfitting. The DRL agent was trained using the PPO algorithm with the following hyperparameters: learning rate = 0.0003, batch size = 64, epochs = 500, discount factor $\gamma = 0.99$, clip ratio $\varepsilon = 0.2$.

4.2 | Prediction Performance

The combined ML pipeline (XGBoost + RF) achieved strong prediction accuracy. The error rate of the algorithm was found to be close to zero, indicating accurate learning and prediction capabilities. Feature importance analysis revealed that volume-related features, momentum indicators, and historical volatility were the most discriminative predictors.

4.3 | Portfolio Performance

Table 2. Portfolio performance across risk metrics.

Risk Metric	Expected Return	Risk Index	Sharpe Ratio
Mean-Variance	0.001	0.010	0.10
Mean-Semi-Variance	0.001	0.010	0.10
Mean-MAD	0.001	0.010	0.10
Mean-CVaR	0.001	0.030	0.03

The portfolio optimization results demonstrate that the expected rate of return across all four risk models achieves a value of 0.001 with risk indices of 0.01 for Variance, Semi-Variance, and MAD models, and 0.03 for CVaR. The CVaR model shows a higher risk index due to its conservative accounting of tail risks, providing a more comprehensive view of extreme loss scenarios.

4.4 | Comparative Analysis

Table 3. Comparison with benchmark strategies.

Strategy	Annual Return (%)	Annual Volatility (%)	Sharpe Ratio	Max Drawdown (%)
Equal weight portfolio	12.5	18.3	0.68	−22.4
Mean-Variance (Markowitz)	15.2	16.1	0.94	−18.7
RF only	17.8	15.4	1.16	−16.2
XGBoost only	18.5	14.9	1.24	−15.5
Proposed hybrid (RF+XGBoost+DRL)	22.3	13.2	1.69	−11.8

The proposed hybrid model outperforms all benchmark strategies across every metric. The combination of ML-based stock selection with DRL-based dynamic allocation achieves the highest annual return (22.3%) with the lowest volatility (13.2%) and maximum drawdown (−11.8%). The Sharpe ratio of 1.69 indicates a significantly superior risk-adjusted return profile compared to the traditional Mean-Variance approach (0.94) and individual ML models (1.16 and 1.24 for RF and XGBoost, respectively).

5 | Conclusion

In this research, portfolio optimization was studied using the RF algorithm and the XGBoost algorithm combined with DRL. According to the results of evaluating the portfolio return, risk index, and other desired indicators across all criteria, the results obtained in this model were satisfactory. The model demonstrates an acceptable ability to form an optimal portfolio of stocks.

By combining these models and algorithms, in addition to linear features, the portfolio can be optimized and selected based on nonlinear features such as political news and macroeconomic variables. However, due to the abundance of features governing the market and factors affecting the determination of companies' portfolios, decision-making in this case is a very complex issue and the optimization process is completely unstructured.

Although classical models have been able to meet the information needs of investors to some extent, they are useful only when the factors affecting the capital market are limited. In reality, many factors affecting the capital market are nonlinear and complex, requiring advanced computational approaches. Using ML models, data can be collected in very short time periods and hundreds of stocks can be monitored in real time, providing the best investment insights.

The key contributions of this research are: 1) integration of XGBoost and RF for robust return prediction in an emerging market context, 2) application of PPO-based DRL for dynamic portfolio rebalancing, 3) comprehensive risk evaluation across four distinct metrics, and 4) demonstration of superior performance compared to traditional optimization approaches in the TSE environment.

Future work should focus on: incorporating transaction costs and market impact models; extending the framework to multi-market portfolio optimization; exploring alternative DRL architectures such as Soft Actor-Critic (SAC) and TD3; and investigating the integration of sentiment analysis from news and social media as additional features.

References

[1] Zare, H. (2021). Political uncertainty and stock market fluctuations in Iran: With consideration of international political changes. *Journal of development and capital*, *6*(1), 123–145. https://doi.org/10.22103/jdc.2021.15440.1091

[2] Fabozzi, F. J., Markowitz, H. M., & Gupta, F. (2008). Portfolio selection. In *Handbook of finance* (pp. 3–13). Wiley. https://doi.org/10.1002/9780470404324.HOF002001

[3] Fahimi, A., & Shahbandarzadeh, H. (2021). Developing the Markowitz portfolio optimization model concerning investor non-financial considerations and supporting domestic products. *Industrial management journal*, *13*(1), 53-79. https://doi.org/10.22059/imj.2021.316204.1007811

[4] Ghasempour, A., Khodamipour, A., Zainali, H., & Taebi Naghandari, A. (2025). Designing an optimal model for predicting stock liquidity using the random forest. *Journal of accounting knowledge, 16*(2), 1-26. (**In Persian**). https://doi.org/10.22103/jak.2025.24973.4156

[5] Ghods, M., Amiry, M., Peymani Forooshani, M., & Dehghan Dehnavi, M. (2025). Prediction of Tehran stock exchange total index using natural gradient boosting and SHAP values. *Financial management perspective*, *15*(2), 34–53. https://doi.org/10.48308/jfmp.2025.239954.1498

[6] Jiang, Z., Xu, D., & Liang, J. (2017). *A deep reinforcement learning framework for the financial portfolio management problem.* https://doi.org/10.48550/arXiv.1706.10059

[7] Markowitz, H. (1952). Modern portfolio theory. *Journal of finance*, *7*(11), 77–91. https://www.raymondjames.com/-/media/rj/advisor-sites/sites/s/t/sturgeonellett/files/february-2023.pdf

[8] Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[9] Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). Association for Computing Machinery (ACM). https://dl.acm.org/doi/abs/10.1145/2939672.2939785

[10] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). *Proximal policy optimization algorithms.* https://doi.org/10.48550/arXiv.1707.06347