Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ31068

Applied Research of Explainable Artificial Intelligence (XAI) in Portfolio Optimization

Hong Li

Jinan University – University of Birmingham Joint Institute, Jinan University, Guangzhou, China
kiomin18@gmail.com

Abstract. Artificial intelligence made great changes in investment management Incorporating machine learning models into portfolio optimization to enhance the portfolio's returns and decrease the risk. However, lacking in explanatory ability of decision making is also challenging to an investor who would like to understand how those automated processes work, to avoid losing too much. To solve this problem, this paper explains how we will use machine learning for portfolio allocation. This research uses SHAP (SHapley Additive exPlanations) to explain the outputs from a machine learning model. The model is trained on the results of a dynamic Markowitz portfolio optimization. CSI 300 market date is applied, the portfolio consists of 10 random stocks having over 1,000 trading days. conduct robustness analysis to see how different rolling windows would affect the explanation of the model. From the results, it is obvious that the framework can identify the drivers of portfolio weights and improve the transparency and understandability of the model.

Keywords: XAI, Portfolio optimization, SHAP, Markowitz model, Machine learning

1. Introduction

After Markowitz introduced the mean-variance model [1], modern portfolio theory has played an important role in much contemporary finance. To be more practical, many papers combine these with machine learning and portfolio optimisation, like M-V and RM-RP. Neural Networks, like RNNs and LSTMs models are used for asset selection as well to reduce the risk Deep neural nets are very good at predicting, but they are often called black boxes and people do not know what they are thinking. It is also unclear how a model would give high weight to certain assets or predict a portfolio to do well. This kind of openness might cause people to misunderstand about AI money stuff, perhaps get even riskier when the model starts acting funny. Also, some studies have found that only a small portion of AI-generated money can beat the market benchmark all the time, and most of which underperform on average.

To make it clearer, there exists more common for XAI or the Explainable Artificial Intelligence. LIME [2], SHAP [3] etc., will be able to explain how the model is making decision by telling you about the different contribution of a feature. It is very useful for making investment decisions which come with a high risk of money. XAI used for portfolio optimization so we can get insight into reasons behind predictions of a model.

DOI: 10.54254/2755-2721/2026.TJ31068

Utilize SHAP values for an explication of the method by which a combination of the machine learning and the portfolio selection works to decide the element from an asset with maximum effect. Even though XAI methods had already been used before studying portfolios, mostly previous works relied on DL which is a model with very high computing time. Thus, they are not good candidates for general application purposes. Therefore, this research aims to formulate a model-ignorant way of trading-off between the prediction accuracy of portfolio selection and the explainable.

2. Literature review

2.1. Traditional portfolio optimization theories

The Finance area had lots of work done on portfolio optimization, and the major idea is reducing risk and bettering returns with proper asset allocation. It is the most influential, and also the most widely used approach, which is the mean - variance (MV) portfolio theory by Markowitz [1], the optimization is done on the expected returns and the risks is being taken into account at the same time and the concept of efficient frontier was introduced too. However, the MV framework is very sensitive inputs; we are assuming normally distributed returns. Biased estimates arise, mostly for the covariance estimation. To tackle the weaknesses, Black and Litterman proposed a model which incorporates investors' subjective views into account [4]. Qian put forward risk parity to handle the equal contribution of each asset to the total portfolio risk [5]. Portfolio optimization after adding variance-based measures, such as mean-VaR models, give us a bit clearer idea of the downside risk too. All these are clear to us, but it does not give an idea about the portfolio, as which asset contributes more to the value of their own portfolio over time. lacks interpretability and thus cannot say how each single of those assets is influencing the risk and returns of the entire portfolio.

2.2. Application of machine learning and deep learning in portfolio optimization

Machine learning is getting more and more applied to portfolio optimization because it can predict well and make models of complicated market actions. Many recent studies use a two-stage design where stock prices are predicted and then used for constructing portfolios. An example is that Zhou et al. [6] show that ensemble learning combined with maximum Sharpe ratio can improve performance. Behera et al. [7] used AdaBoost for price prediction and Mean - VaR (Mean – VaR) for allocation, and Wang et al. [8] further improved it by data denoising and LSTM settings. prove that machine learning can give better portfolio returns and less portfolio risks. However, one downside of most two-stage ways is that they are not easy to understand. They rarely explain the reasons behind making different model decisions (e.g., a model that refuses a loan), which drives people interested in explainable artificial intelligence (XAI).

2.3. Previous research of XAI in portfolio optimization

Though AI-driven models can offer decent predictability and give some support in making decisions about the portfolio, they remain like black boxes which makes it difficult to understand their logic. There are lots of XAI methods made by researchers, among which are: Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP). The LIME algorithm is proposed by Ribeiro et al. [2]. They construct an interpretable and simplified model as the surrogate of a complex, black-box model. SHAP method was proposed by Lundberg and Lee [3]. It uses cooperative game theory and applies Shapley value to calculate the contribution of each feature to a prediction.

DOI: 10.54254/2755-2721/2026.TJ31068

Lots of research have brought XAI into AI portfolio management to make things more transparent and to decrease that number of financial risks. Lina et al [9] and Chorasiya et al [10] used DRL and SHAP/LIME on portfolio decision explanations. They provide useful guidelines as to which explanatory features and outputs to pick, but DRL-based frameworks use too many resources and can take a long time to train, so people may not find them practical. Accessibility is improved by using SHAP for the crypto asset allocation project like Babaei et al. [11], a model agnostic approach which increases the model interpretability at no heavy computational cost.

3. Methodology

In this study, XAI is applied to investigate the stocks that contribute most significantly to portfolio optimization. The primary objective is not only to compare the effectiveness of different methods for asset selection, but also to propose a framework that explains the decisions made by the model. SHAP is employed in this study to interpret the predicted performance of the portfolio. The following sections present the theories and methodologies employed in this study.

3.1. Modern portfolio allocation

The mean-variance (MV) model proposed by Markowitz [1] is employed to construct portfolios. The model assumes that investors are risk averse. That is, given two portfolios with identical expected returns, investors prefer the one with lower risk. The optimization problem is formulated as follows:

$$\max_{w} \ w^T \mu - \frac{\lambda}{2} w^T \Sigma w \tag{1}$$

where w denotes the weight vector, μ represents the expected return vector, λ denotes the risk-aversion coefficient and Σ denotes the covariance matrix of asset returns.

After obtaining the resulting weight, the portfolio's expected return and associated risk can be calculated as follows:

$$E[R_p] = \sum_i w_i R_i \tag{2}$$

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

where σ_i denotes the standard deviation of asset i, ρ_{ij} represents the correlation between assets i and asset j.

However, these parameters alone are insufficient to determine which asset exerts the most significant influence on the expected return and risk of a portfolio during a specific period. Widely used explainable methods typically require feature inputs; however, portfolio optimization (Equation(1)) essentially relies on mathematical models, which are inherently interpretable. To address this question, inspired by the work of Babaei et al. [11], we introduce the concept of the Z-score. The Z-score is a statistical measure that evaluates both return and risk within the context of portfolio performance. The rationale for using Z-scores as features and target variables lies in their ability to integrate both return and risk, which are essential for determining an individual asset's contribution to overall portfolio performance. When these factors are considered separately, it becomes challenging to construct a cohesive formula linking the portfolio to its constituent assets. Its formulation for a portfolio is given as follows:

DOI: 10.54254/2755-2721/2026.TJ31068

$$Z_p = \frac{R_p - E(R_p)}{\sigma_p} \quad (4)$$

And for an asset, its Z-score is given by:

$$Z_i = \frac{R_i - E(R_i)}{\sigma_i} \quad (5)$$

Using the Z-score of portfolios and individual assets, the relationship can be estimated by treating portfolio Z-score Z_p as the dependent variable and the assets' Z-scores Z_i as predictors. To gain enough samples, the dataset is expanded by generating daily portfolio based on the previous 30 days. For example, in the experiment, data for 10 stocks from the CSI 300 index were collected between June 2019 and June 2024. By applying the rolling 30-day strategy, Z_p and Z_i were calculated for each trading day from July 2019 to June 2024.

3.2. Explainable AI with SHAP

In the next step, a machine learning model is trained and tested, with the portfolio's Z-score as the dependent variable and the assets' Z-scores as predictors. The main objective is not only to predict the performance of a one-day portfolio, but also to integrate SHAP values to interpret the marginal contributions of individual assets.

SHAP, derived from the Shapley value in cooperative game theory, is an extension employed for model interpretability. Essentially, it quantifies each feature's marginal contribution to the model output by evaluating the change when the feature is included versus excluded. Larger absolute values indicate stronger positive or negative effects on the model output. Formally, SHAP values are calculated using the formula provided by Lundberg and Lee [3]:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S\cup\{i\}}\left(x_{S\cup\{i\}}\right) - f_S\left(x_S\right) \right] \quad (6)$$

Here, F denotes the set of all features (in this case, the assets' Z-scores), S represents a subset of F, and $f_{S\cup\{i\}}\left(x_{S\cup\{i\}}\right)$ and $f_S\left(x_S\right)$ correspond to the model predictions for a single observation with feature i included and excluded, respectively.

In our study, the prediction function f corresponds to the portfolio's Z-score, while the predictor x_S represents the vector of Z-scores of the securities included in set S.

4. Experiments

This section provides an overview of the experiments which the explainable method SHAP is applied to portfolio optimization. First, the weights of individual assets are calculated to construct daily portfolios. The proposed portfolio allocation strategy is also compared with the $1/N$ portfolio selection benchmark. A machine learning model is subsequently applied to model the relationship between portfolios' and individual Z-scores. Finally, the explainable results are presented, illustrating each asset's contribution to the portfolio over the selected time interval.

4.1. Dataset

In this study, the daily historical price data of the CSI 300 index was employed as the market for testing the proposed method. Ten stocks were randomly selected from the 300 constituent securities

DOI: 10.54254/2755-2721/2026.TJ31068

of the CSI 300 index to construct portfolios, covering the period from June 2019 to June 2024. Consequently, the dataset comprises 1,209 trading days in total. Table 1 presents the tickers of the selected ten securities in the CSI 300 index along with their respective statistical measures. Figure 1 illustrates the monthly closing prices of the selected securities. Figure 2 illustrates the monthly comparative rates of all ten assets after base-period normalization, facilitating intuitive comparison. Base-period normalization is defined as

$$p/p_0 \times 100\% \tag{7}$$

where p denotes the monthly closing price and p_0 represents the closing price in the first month.

Table 1. Selected 10 assets and corresponding statistical parameters

Asset	Mean	Std.	Max	Min
000333	62.54	12.42	108.00	40.09
300059	21.64	7.39	40.57	11.83
300308	60.71	38.53	192.88	25.36
300502	44.98	16.91	93.16	20.02
300750	303.43	166.10	692.00	64.00
600036	39.23	7.80	58.92	26.30
600519	1666.57	318.63	2627.88	840.00
600900	20.87	2.31	27.02	16.01
601318	60.89	17.72	94.62	35.90
601899	9.30	3.58	19.79	3.16

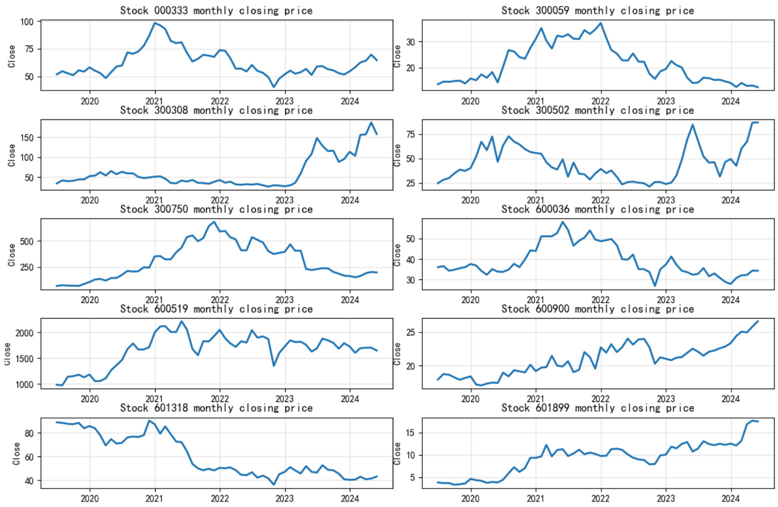

Figure 1. Monthly closing price for each asset during 01-06-2019 to 01-06-2024

DOI: 10.54254/2755-2721/2026.TJ31068

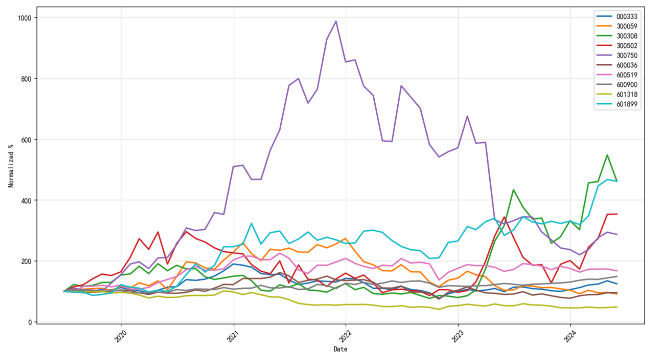

Figure 2. Base-period normalization of 10 stocks selected in the portfolio

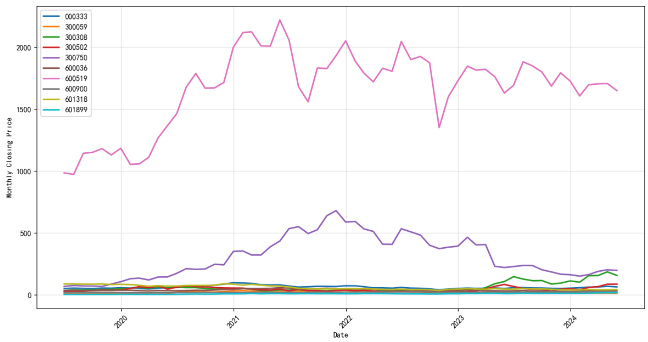

Figure 3. Comparison of 10 stocks' monthly closing prices

From Table 1, it is observed that tickers 300750 and 600519 are the most volatile assets over the period. Figure 2 indicates that ticker 300750 exhibits pronounced fluctuations in its monthly closing prices. Figure 3 shows that ticker 600519 exhibits the highest monthly closing prices, ranging between 1,000 and 2,500.

4.2. Experiment process

As mentioned in Section 3.1, daily portfolios are generated based on the preceding 30 trading days. Specifically, the first portfolio is produced on July 16, 2019, based on the trading data from June 3, 2019, to July 15, 2019. Consequently, after calculation, the portfolios' weights are stored in a 1179×10 matrix, where 1179 represents the number of days available for portfolio construction and 10 represents the number of selected stocks.

In the next step, once the weights assigned to each asset are obtained, the daily portfolio returns and risks are calculated by Equation (2) and (3).

Table 2. Comparison of different indicators between Markowitz model and two benchmarks

	SR	APV	RI (30)	AR	MDD
Equal Weight	1.1789	3.1201	0.0423	0.2237	-0.3205
Risk Parity	0.8646	1.9632	0.0381	0.1846	-0.3361
Markowitz Model	0.6780	1.4668	0.0156	0.1297	-0.2081

DOI: 10.54254/2755-2721/2026.TJ31068

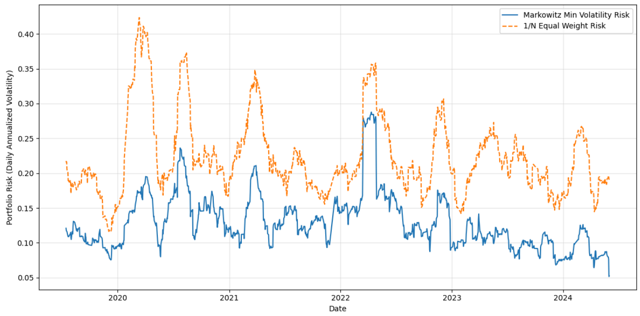

Figure 4. Portfolio risk comparison: Markowitz vs. equal weight

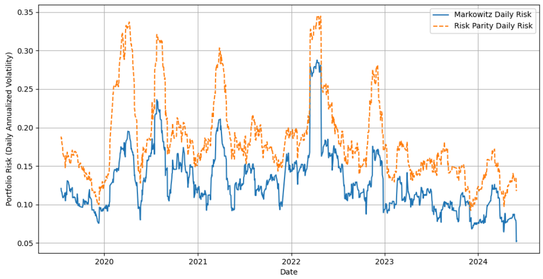

Figure 5. Portfolio risk comparison: Markowitz vs. risk parity

Table 2 compares five performance metrics of the Markowitz model against the equal-weight and risk-parity benchmarks. The five criteria include the Sharpe Ratio (SR), Accumulative Portfolio Value (APV), Rolling Interest over the latest 30 days (RI (30)), Average Risk (AR), and Maximum Drawdown (MDD).

The Sharpe Ratio (SP) quantifies the trade-off between a portfolio's excess return over the risk-free rate and its risk. A higher value of the Sharpe Ratio indicates that, for the same level of risk, the portfolio generates a higher return, or alternatively, for a given return, the portfolio entails less risk. The formula for the Sharpe Ratio is given by Equation(8), where R_p represents the return rate of the portfolio, R_f represents the return rate of risk-free assets (typically set to 0), and σ_p denotes the portfolio's risk.

$$RF = (R_p - R_f)/\sigma_p \tag{8}$$

The Accumulative Portfolio Value (APV) represents the overall return of a portfolio over time, as given by Equation(9). In this context, R_t denotes the return rate at the end of the period, while R_0 represents the initial value, which is set to 1.

$$APV_t = R_t/R_0 \tag{9}$$

The Rolling Interest over the latest 30 days (RI(30)) measures the return rate over the past 30 days, as given by Equation(10). A higher value of RI (30) indicates better recent portfolio performance. Other variations of Rolling Interest, such as RI (20) and RI (50), can be used to measure performance over different time periods.

$$RI\,(30) = R_t/R_{t-30} \tag{10}$$

DOI: 10.54254/2755-2721/2026.TJ31068

The Average Risk (AR) calculates the mean daily risk of a portfolio, as given by

$$AR = \sum_i \sigma_i / \sum_i 1 \tag{11}$$

The Maximum Drawdown (MDD) is a metric that indicates the largest loss over a specified period. A higher value of MDD corresponds to a more significant loss for investors. The formula is given below.

$$MDD = \max_{i=1,...,t} \left(APV_i / \left(\max_{j<i} APV_j \right) - 1 \right) \tag{12}$$

Table 2 shows that although the Markowitz model may yield lower returns, as indicated by the first three metrics, it significantly reduces overall risk, making it a suitable choice for risk-averse investors. Figure 4 presents the distribution of daily volatility for the Markowitz model using both the minimum variance strategy and the $1/_N$ strategy. The figure illustrates that a dynamic robo-advisor strategy, which adjusts portfolio weights over time, results in less risky portfolios compared to a static, fixed-weight approach. Similarly, Figure 5 shows a comparable outcome between the Markowitz model and the risk parity model, with only a slight difference in risk levels. Together, these two figures demonstrate the superiority of the Markowitz model in minimizing portfolio risk.

Next, Z-scores are calculated to identify the asset that significantly contribute to the portfolio's performance. As explained in Section 3.1, Z-scores are calculated for both portfolios and assets that constitute the portfolios. Following Equations(4)and(5), the Z-scores are computed. After computation, a 1179×11 matrix is constructed to store the Z-scores for both portfolios and assets.

Then, a machine learning model is applied to model the relationship between the portfolio's Z-score and the assets' Z-scores. A random forest model is employed, with the dataset split into 70% for training and 30% for testing. The random forest model is selected primarily for practical reasons: Python provides a built-in implementation named "RandomForestRegressor" in the "sklearn" package, and the corresponding "TreeExplainer" in SHAP facilitates model explainability. As the purpose of this study is not to optimize model accuracy for higher returns or lower risks, the goodness-of-fit is not discussed.

Finally, the SHAP explainer is applied to the random forest model to identify the security with the greatest influence on portfolio performance.

4.3. Result analysis

The experiment is conducted following the procedure described above. The following two figures illustrate the mean absolute SHAP values and the SHAP values separated into positive and negative contributions based on the rolling window of 30 trading days.

From Figure 6, it is evident that ticker 600900 exhibits the highest mean absolute SHAP value. When considered together with Table 1, it is observed that ticker 600900 also has the lowest standard error, indicating that the Markowitz portfolio tends to select stocks with lower risk. Although ticker 600519 has the highest overall price among the ten stocks, it exhibits a relatively small SHAP value due to its high standard error. A similar pattern is observed for ticker 300750.

DOI: 10.54254/2755-2721/2026.TJ31068

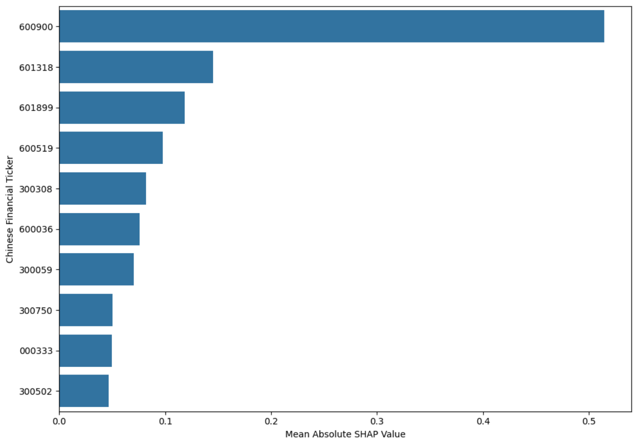

Figure 6. Mean absolute value of tickers

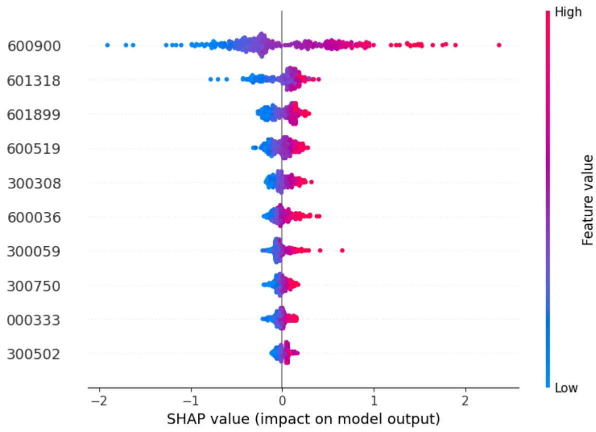

Figure 7. SHAP value of each ticker

Notes: red color represents positive effect and blue color represents negative impact.

From Figure 7, the positive and negative effects of each asset can be directly identified. Ticker 600900 exhibits strong effects in both directions, consistent with its high mean absolute SHAP value. From Figure 1, ticker 601318 followed a downward trend in its price over the period. Correspondingly, its negative SHAP values exceed its positive values, indicating that, to maximize portfolio returns, short-selling could be considered for ticker 601318.

A robustness analysis is also conducted to compare the influence of different rolling windows. To assess the impact of varying the length of the rolling window, experiments using 20-day and 50-day windows are compared, as shown in Figures 8 and 9. Interestingly, for the 20-day, 30-day, and 50-day rolling windows, ticker 600900 consistently exhibits a remarkable SHAP value compared to the others, maintaining its rank as number 1. Conversely, tickers with minimal influence on portfolio performance, such as 300750, 000333, and 300502, despite slight rank variations, remain at the bottom of the list. For the remaining tickers, although their ranks differ across rolling windows, their overall impact on portfolio performance remains relatively insignificant.

DOI: 10.54254/2755-2721/2026.TJ31068

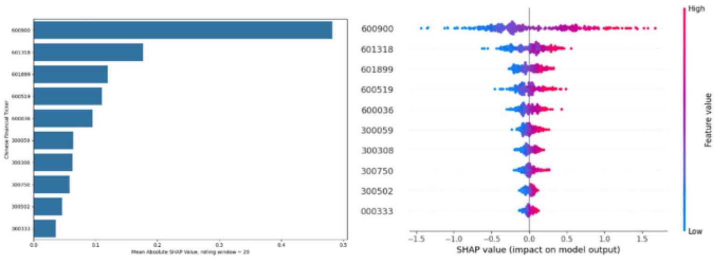

Figure 8. Mean SHAP and two-sided SHAP values of tickers based on 20 rolling trading days

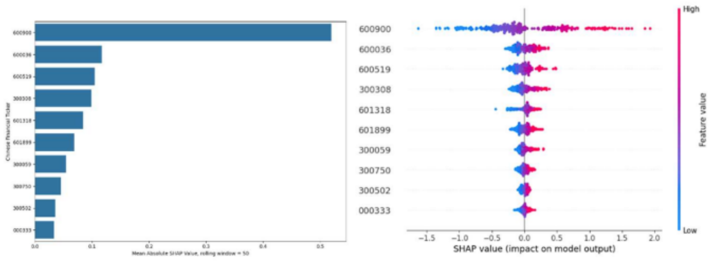

Figure 9. Mean SHAP and two-sided SHAP values of tickers based on 50 rolling trading days

The figures demonstrate the effectiveness of the proposed explainable method. Regardless of the rolling window used, stocks with either the most significant or minimal influence on the portfolio are consistently identifiable. It can be concluded that ticker 600900 has the greatest impact on portfolio performance during the analyzed period.

5. Conclusion

This paper focuses on portfolio selection problem in the Chinese market. The method proposed by Babaei et al. [11] is mainly based on the SHAP (SHapley Additive exPlanations). It is an explainable AI approach to ensure portfolio performance. To figure out how the portfolio is faring, the Z-score is made for the training of the machine learning model.

The results show that the process works, and that the most influential security has been identified. This successful framework also shows adaptability to markets like the S&P 500. Putting XAI into portfolio optimization can help FinTech advising firms with making and watching over AI-driven investment strategies, which can lead to more trust between people and AI systems.

In future work, the presented framework will be extended for different combinations for example risk parity as a portfolio allocation technique with XGBoost as predictive model. To have a better analysis about the robustness and stability of methodology. And due to mathematical limitations of SHAP noted by Kumar et al. [12], alternative methods like p-value correction could be incorporated into the Shapley-value approach to address these biases.

References

[1] Markowitz, H. M., & Todd, G. P. (2000). Mean-variance analysis in portfolio choice and capital markets. John Wiley & Sons.

[2] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https: //doi.org/10.1145/2939672.2939778

DOI: 10.54254/2755-2721/2026.TJ31068

[3] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30. https: //proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
[4] Black, F., & Litterman, R. (1992). Global Portfolio Optimization. Financial Analysts Journal, 48(5), 28–43. http: //www.jstor.org/stable/4479577
[5] Qian, E. (2005). On the Financial Interpretation of Risk Contribution: Risk Budgets Do Add Up. http: //dx.doi.org/10.2139/ssrn.684221
[6] Zhou, Z., Song, Z., Ren, T., & Yu, L. (2023). Two-Stage Portfolio Optimization Integrating Optimal Sharp Ratio Measure and Ensemble Learning. IEEE Access, 11, 1654–1670. https: //doi.org/10.1109/ACCESS.2022.3232281
[7] Behera, J., Pasayat, A. K., Behera, H., & Kumar, P. (2023). Prediction based mean-value-at-risk portfolio optimization using machine learning regression algorithms for multi-national stock markets. Engineering Applications of Artificial Intelligence, 120. https: //doi.org/10.1016/j.engappai.2023.105843
[8] Wang, C.-H., Zeng, Y., & Yuan, J. (2024). Two-stage stock portfolio optimization based on AI-powered price prediction and mean-CVaR models. Expert Systems With Applications, 255(124555). https: //doi.org/10.1016/j.eswa.2024.124555
[9] Lina, J., Banda, S. S., Paib, H.-T., & Rawal, B. S. (2024). EUDRL: Explainable Uncertainty-Based Deep Reinforcement Learning for Portfolio Management. 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 1–7. https: //doi.org/10.1109/ICCCNT61001.2024.10723849
[10] Chorasiya, R., & Kinger, S. (2025). Portfolio Optimization of Indian Stock Market Using Reinforcement Learning and Explainable AI. Technische Sicherheit, 25(6), 814–823. https: //technikwissen.eu/wp-content/uploads/62-TSJ2951.pdf
[11] Babaei, G., Giudici, P., & Raffinetti, E. (2022). Explainable artificial intelligence for crypto asset allocation. Finance Research Letters, 47(102941). https: //doi.org/10.1016/j.frl.2022.102941
[12] Kumar, E., Venkatasubramanian, S., Scheidegger, C., & Friedler, S. A. (2020). Problems with Shapley-value-based explanations as feature importance measures. International Conference on Machine Learning, 5491–5500. https: //proceedings.mlr.press/v119/kumar20e.html