Portfolio Optimization Strategies: New Approaches Based on Machine Learning Forecasting

Xuanlin Lyu *

School of Economics and Business, Chongqing University, Chongqing, 400044, China

* Corresponding Author Email: 20215590@stu.cqu.edu.cn

Abstract. This study provides an in-depth discussion and comprehensive review of the latest applications of machine learning techniques in the field of portfolio optimization. The article begins with an overview of traditional portfolio optimization theory and its limitations, and then focuses on how machine learning predictive models, which have flourished in recent years, can provide new perspectives and tools for solving the problems of non-linearity, dynamics and uncertainty in investment decision-making. This paper provides a detailed overview of the application practices of various machine learning algorithms (e.g., deep learning, reinforcement learning, integrated learning, etc.) in the areas of asset return prediction, risk assessment, and optimal weight allocation, and analyses their advantages and challenges compared to traditional methods. The analysis of relevant research cases reveals the significant effect of machine learning predictions in increasing expected portfolio returns, reducing risk exposure, and achieving effective diversification. The study also explores possible future trends and potential research directions for machine learning in portfolio optimization, highlighting the importance of combining domain knowledge with big data-driven intelligent investment decisions. This review aims to provide financial scholars and practitioners with a new way of thinking about portfolio optimization, and to promote the combination of theoretical research and practical operation in the fields of financial engineering and investment management, so as to achieve more accurate and efficient investment decisions.

Keywords: Portfolio Optimization; Machine Learning; Investment Forecasting; Deep Learning.

1. Introduction

In the modern financial market, portfolio optimization is an important topic in the field of financial engineering and risk management, which aims to maximize returns **while** effectively controlling risks through rational allocation of various assets. Traditional portfolio optimization theories, such as Markowitz's Modern Portfolio Theory (MPT), provide investors with an effective asset allocation framework, but are often subject to a number of limitations in practical application, such as the assumption of market efficiency, the accurate prediction of the distribution of returns, and the inability to dynamically adapt to market environmental changes in the market environment, etc.

With the advent of the big data era and the rapid development of machine learning technology, it brings new opportunities and challenges for portfolio optimization strategies. Machine learning algorithms are capable of automatically extracting features, discovering patterns, and predicting future trends from massive amounts of historical data, which makes it possible for portfolio management strategies to break through the limitations of traditional models and achieve more accurate risk-return trade-offs and dynamic adjustments. In addition, the plethora of machine learning models presents a dilemma for portfolio managers seeking to leverage these techniques: which model should be used? Should forecasts or portfolios obtained using different machine learning models be combined, and if so, how?

In recent years, cutting-edge machine learning techniques such as deep learning and reinforcement learning have been widely used in portfolio optimization research, e.g., Zhang et al. use deep learning for yield prediction and portfolio model construction [1]. Zhuang et al. propose a new model for dynamic portfolio optimization using reinforcement learning [2]. There are also integrated deep reinforcement learning portfolio models proposed by Jie Long et al. [3]. Machine learning techniques are able to deal with non-linear relationships, high-dimensional data and uncertainty, leading to more granular and intelligent management of portfolios. However, although the application of machine learning to portfolio optimization shows great potential, how to effectively combine these complex algorithms with the characteristics of financial markets to design robust and practical investment decision-making solutions is still an important topic that requires in-depth research.

Therefore, this paper first introduces the commonly used machine learning methods in portfolio, discusses the main problems encountered in the current machine learning portfolio optimization methods, points out the differences between the application scenarios of different algorithms, analyses the principles, advantages and disadvantages of different algorithms, and provides an outlook on the application of machine learning in the portfolio optimization problem.

2. Machine Learning Methods Based on Feature Processing

2.1. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a commonly used data degradation and feature extraction method with a wide range of applications in the field of machine learning. PCA transforms the original high-dimensional data into a new set of variables that are orthogonal to each other in each dimension, called principal components, by linear transformation. These principal components are ranked in order of variance size, with the first principal component reflecting the information about the largest variation within the dataset, and subsequent principal components reflecting, in turn, the information about the largest remaining, independent variation. PCA achieves data compression and simplification while maximizing the preservation of the original information in the dataset. The advantages of this approach are clear. It allows the data to be downscaled, removes noise and makes multi-dimensional data easy to understand and visualize and highlights key information. At the same time, the drawbacks are obvious PCA is based on linear relationships for dimensionality reduction and may not perform well for highly non-linear data. The results of PCA are also susceptible to outliers and require pre-cleaning or pre-processing of the data.

Sun proposed a weighted principal component analysis (WPCA) algorithm that takes into account both expected return and variance-covariance matrix estimation error [4]. The algorithm overcomes the problem of the limitations of the factors extracted by classical principal component analysis in explaining returns by introducing a weighting term for the estimation error of the first-order moments of returns in the objective function. And the estimates of the parameters constructed from the statistical factors extracted by the WPCA algorithm are brought into the mean-variance (MV) strategy to obtain the WPCA-MV strategy for portfolio optimization in the A-share market. The results show that the WPCA-MV strategy achieves superior sample performance in terms of mean return, standard deviation, Sharpe ratio, and cumulative return metrics compared to the common portfolio strategies MV, GMV, EW, BS, and TZ, and it also holds true on a 75-factor dataset of U.S. stocks.

2.2. Boost Algorithm

Boost algorithm is an important branch of machine learning, specifically, it belongs to one of the methods of Ensemble Learning (EL). Integration learning is a technique to gain greater predictive power by building and combining multiple learners. The Boosting algorithm is a concrete implementation of this idea, and its core concept is to gradually train a series of weak classifiers (or regressors) in an iterative manner and combine them in a specific way to form a strong classifier (or regressor), so as to improve the overall prediction performance. By integrating multiple weak learners, the Boosting algorithm reduces the sensitivity of individual models to data noise or outliers and improves model robustness. Also, its built-in regularization mechanism helps prevent overfitting and ensures that the model still performs well on unlearned data. When dealing with high-frequency trading data, boosting models may focus too much on short-term patterns in the market's microstructure, resulting in models that underperform in the long term or fail quickly when market conditions change.

Qian et al. in their study used the Adaboost algorithm in Boost algorithm for portfolio optimization, using the whole stock sample data of the A-share for nearly 25 years and the US-share for 40 years, to prove theoretically and empirically that the machine learning techniques can help to improve the optimization of portfolio strategies [5].

2.3. Random Forest

Random Forest is an integrated learning method that consists of multiple decision trees and uses a randomized approach to generate and combine these decision trees for classification or regression tasks. It was proposed by Breiman in 2001 to improve the generalization and overfitting resistance of decision trees by introducing randomness. The Random Forest algorithm has a wide range of applications in portfolio management, especially in stock selection, risk assessment, and market forecasting. It has the advantage of being able to quantify the importance of each feature for predicting the target variable. For example, in the stock selection process, this helps to identify the financial indicators, market indicators or other relevant factors that have the greatest impact on stock returns, thus guiding investors to prioritie these key factors to screen for the most promising stocks. Also Random Forest is good at dealing with high dimensional data and datasets that contain complex non-linear relationships. Its disadvantages are also clear; the performance of random forests, like other machine learning models, is highly dependent on the quality and representativeness of the training data. If the market environment changes significantly, factors and relationships that were valid in the past may no longer be applicable, causing the model to fail, and investors will need to retrain the model and adjust the factor weights on a regular basis.

Zhou in his study used common factors to construct a random forest model to predict the returns of the constituent stocks of the CSI 500 index and construct portfolios [6]. The results indicated that random forest can fit and predict the relative returns of individual stocks better.

2.4. Support Vector Machine (SVM)

Support Vector Machine (SVM) belong to the supervised learning algorithms of machine learning algorithms and are particularly suitable for applications in classification and regression tasks. Until the advent of deep learning algorithms, SVM was considered the most successful and best performing algorithm in machine learning. The basic idea of Support Vector Machines is to establish an optimal decision boundary, i.e. a hyperplane, in order to maximize the distance between different classes of samples, the fundamental intention is to ensure that the model maintains excellent judgement in the face of unknown data, i.e. good generalization performance. In the field of portfolio management, SVM are widely used in a variety of areas such as asset classification, risk measurement, and investment strategy development and adjustment. The application of SVM in portfolio management has shown some significant advantages as well as limitations. It has the advantage of being able to efficiently deal with complex nonlinear relationships in financial data through kernel functions, which is crucial for capturing the wide range of nonlinear phenomena (e.g., asset price fluctuations, market dynamics, interactions between economic indicators, etc.) in financial markets. And it is stable in financial data that contains outliers or is volatile. The disadvantage is that its performance is highly dependent on the choice of kernel function, and there is a certain threshold for its use. There is also the fact that in portfolio management, certain events (e.g., extreme risk events) may occur much less frequently than normal, resulting in a category imbalance in the dataset. SVM may be overly sensitive to most classes of data in their default settings, neglecting the ability to predict rare but important events (e.g., financial crises) and requiring additional techniques to improve.

Silva et al. in their study proposed a portfolio optimization method based on pre-selecting stocks using SVM models [7]. The specific method is used to select assets prior to portfolio optimization and then optimize them by maximizing the Sharpe ratio. The classification method mentioned in the study achieved 61% accuracy.

3. Deep Learning Methods for End-to-End Processing

3.1. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN) are deep learning models specifically designed to process input data with a lattice structure (e.g., images, time series, or text sequences). CNN initially achieved remarkable success in the field of computer vision, and have since been widely used in speech recognition, natural language processing, and financial data analysis. Fu and Wang in their study pointed out that the advantage of CNN is that the desired features can be extracted automatically by convolution and it also shares convolution kernel which can be used to process high dimensional data [8]. The disadvantage is that a lot of useful information is lost in the sampling process. In practice, it is not common to apply CNN directly to portfolio optimization, as portfolio optimization usually involves statistics, mathematical optimization methods and traditional machine learning algorithms. However, CNN can extract valuable information from massive amounts of unstructured data, which can be further integrated into the investment decision-making process, e.g. for enhancing factor analysis or constructing new quantitative strategies.

Ashrafzadeh et al. proposed a return prediction model based on Particle Swarm Optimization Algorithm (PSO) with Convolutional Neural Networks (CNN) for stock pre-selection in portfolio optimization [9, 10]. Its model does not differ significantly from traditional methods in terms of forecasting accuracy, and in the portfolio optimization step, the model shows superior financial performance compared to other benchmark models.

3.2. Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) is a special type of neural network architecture designed for processing sequential data that captures the time dependence of the data. RNN introduce a cyclic structure in the network so that the output of the current moment depends not only on the current input, but is also related to its hidden state of the previous moment, so that past sequential information can be remembered. Recurrent neural networks, with their ability to capture time-series dependencies, have a promising application in portfolio management, especially when dealing with financial data with time-ordered correlations, which can help investors better grasp market dynamics and trends. RNN have a unique architecture that can flexibly cope with sequence inputs of any length, and at the same time take into account the backward and forward connections between the information at each time point when processing sequence data. The circular propagation mechanism allows the model to reuse the same parameters when processing the entire sequence, reducing parameter complexity. However, RNN have limitations in dealing with long-term memory and facing the gradient vanishing problem, and to successfully apply RNN, it is necessary to overcome the training challenges mentioned above and incorporate domain expertise for effective modelling.

3.3. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) neural networks are a special kind of Recurrent Neural Networks (RNN) that are particularly suitable for processing time-series data, and have attracted attention for their ability to capture long-term dependencies. In portfolio management, LSTM networks are applied to predict the price movements of financial assets such as stocks, futures, foreign exchange, etc., to provide investors with a reference of future price movements and assist in buying and selling decisions and risk control. By analyzing historical price series and related market data, LSTM can also identify patterns and trends in asset price fluctuations, helping to quantify and predict portfolio risk levels such as volatility, maximum retracement, etc. LSTM can be combined with textual data (e.g., news, social media posts, research reports) to capture changes in the sentiment of market participants, providing portfolio management with insights about market expectations and potential turning points. LSTM networks are also widely used in sequence data analysis tasks such as automatic speech recognition, language translation, handwritten character recognition, and time-series data prediction by virtue of their ability to remember both short-term and long-term values.

Zhang et al. proposed a two-stage mean-CVaR portfolio model based on LSTM [11]. The principle is to use LSTM to predict stock returns and select suitable stocks in the first stage, and apply the mean-CVaR model to determine the investment ratio in the second stage. The empirical results show that the use of the LSTM-CVaR model achieves higher average returns, return-to-risk ratios, cumulative returns and Sharpe ratios than the traditional portfolio model.

4. Conclusion

Portfolio optimization strategies using machine learning techniques have already made significant gains in the decision-making process in financial engineering and risk management. However, this emerging field has only scratched the surface of its enormous potential. The following section outlines several promising future research and development directions that promise to further revolutionize investment practices.

As deep and reinforcement learning algorithms continue to advance, increasingly complex models will emerge. These models are capable of adapting to market dynamics in real time. Such models can facilitate an adaptive, dynamic process of optimizing investment decisions. Portfolios can iteratively refine their strategy choices through continuous interaction with changing market conditions. This will enable investors to respond quickly and effectively to market changes, seizing fleeting opportunities while mitigating the risks associated with rapid change.

Integration of multiple machine learning models (e.g., integration learning and transfer learning) is significant in combining the strengths of various approaches to enhance the robustness and accuracy of investment decisions. By merging predictions from multiple models, integration can mitigate the weaknesses of individual models, reduce sensitivity to noise, and provide more reliable predictions. In addition, hybrid intelligent models that combine machine learning with traditional financial theory can find a balance between theoretical rigor and practical applicability, driving a more comprehensive and contextually grounded approach to portfolio management.

Incorporate non-traditional factors into models. Many contemporary machine learning models have begun to incorporate non-financial information, sentiment indicators and social media data as unconventional influences. Future research should develop a deeper understanding of how these variables shape portfolio optimization, thereby enriching and refining the dimensionality of investment decisions. This may involve exploring novel data sources, developing advanced feature extraction techniques, and refining algorithms that effectively capture the complex interactions between traditional financial indicators and emerging, alternative data signals.

Due to regulatory requirements and the need for trust in investment decisions, there is an increasing focus on improving the explainability and transparency of machine learning-based investment strategies. The adoption of Explainable Artificial Intelligence (XAI) techniques is critical to elucidating the decision logic behind these models, helping to enhance stakeholder understanding and acceptance. Transparent models not only increase investor confidence, but also promote regulatory compliance and facilitate knowledge transfer within investment teams.

Given the increasing globalization and interconnectedness of financial markets, machine learning-driven portfolio optimization methods are expected to be applied beyond single-market or single-asset-class applications. Future developments may see these strategies extended to cover global, multi-market, multi-asset class portfolio optimization scenarios, enabling investors to navigate complex international investment environments with greater precision and flexibility.

In summary, the field of machine learning-driven portfolio optimization stands at the forefront of a profound transformation that will be shaped by a combination of sustained technological innovation, methodological sophistication, and the diversification of application scenarios. Future research and development will unfold along the aforementioned paths, including but not limited to the deep integration of deep reinforcement learning, the application of multi-dimensional integration methods, the effective integration of non-traditional factors, the significant improvement of model interpretability and transparency, and the wide range of cross-market and cross-asset class applications. These advances will not only further enhance the accuracy and efficiency of investment decisions, but also create more advanced, diverse and market-responsive solutions for investors seeking optimal investment strategies in the complex and ever-changing global financial environment.

References

[1] Zhang, N., Yan, S., & Fan, D. (2022). Deep learning-based yield prediction and portfolio modeling. Statistics and Decision Making, 38(23), 48-51.

[2] Zhuang, W., Chen, C., & Qiu, G. (2022). A new deep reinforcement learning model for dynamic portfolio optimization. Journal of University of Science and Technology of China, 52(11), 12-25.

[3] Long, J., Xie, L., & Xu, H. (2024). Integrated deep reinforcement learning portfolio model. Computer Applications, 44(1), 300-310.

[4] Sun, H. (2023). Portfolio optimization based on weighted principal component analysis. System Science and Mathematics, 43(7), 1888-1903.

[5] Qian, L., Wei, J., Zhao, H., & Ni, X. (2022). Portfolio optimization based on AdaBoost. System Science and Mathematics, 42(2), 271-286.

[6] Zhou, L. (2021). A study on multi-factor investment in stocks based on random forest model. Finance Theory and Practice, (7), 97-103.

[7] Silva, N. F., de Andrade, L. P., da Silva, W. S., de Melo, M. K., & Tonelli, A. O. (2024). Portfolio optimization based on the pre-selection of stocks by the Support Vector Machine model. Finance Research Letters, 61.

[8] Fu, Y., & Wang, M. (2022). A research overview of the application of deep learning in the financial field. Software Engineering, 25(3), 1-4.

[9] Ashrafzadeh, M., Taheri, H. M., Gharehgozlou, M., & Zolfani, S. H. (2023). Clustering-based return prediction model for stock pre-selection in portfolio optimization using PSO-CNN plus MVF. Journal of King Saud University-Computer and Information Sciences, 35.

[10] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In International Conference on Neural Networks, (4), 1942-1948.

[11] Zhang, P., Yang, Y., Li, J., & Zeng, Y. (2023). A Two-stage Mean-CVaR Investment Strategy Based on LSTM. Journal of South China Normal University (Natural Science Edition), 55(5), 93-102.