Eng Int Syst (2025) 4: 365–375

Engineering
Intelligent Systems

Optimizing Quantitative Investment Strategies and Asset Allocation with Machine Learning

Zhanyong Wu, Xiaohang Ma and Yanxue Li*

School of Economics and Management, Handan University, Handan 056005, China

In the decision-making process of financial markets, machine learning techniques, especially support vector machines, have shown their potential to improve the efficiency and accuracy of investment strategies. This study explores the application of machine learning for the optimization of quantitative investment strategies and asset allocation by constructing and optimizing an SVM-based multi-factor stock selection model and asset allocation system. This study verifies the actual performance of the proposed SVM model in the financial market and its ability in terms of risk control and maximization of returns. The results show that SVM provides a higher rate of return and lower risk than traditional investment methods, which confirms its application value in modern financial strategies. This study provides a new perspective on, and technical support for, the field of quantitative investment and presents theoretical support for the further development and application of machine learning technology in the financial market.

Keywords: machine learning; support vector machine; finance; investment strategy; asset allocation

1. INTRODUCTION

Over the past few decades, the complexity and volume of data in financial markets have continued to increase, and traditional investment strategies and asset management approaches have faced increasing challenges. To adapt to this change, quantitative investment strategies have emerged, using mathematical models and big data technology to make investment decisions, aiming to improve investment efficiency and returns on investment (ROI). With the development of artificial intelligence and machine learning technologies, the application of these methods in the financial field is also increasing, particularly for the analysis of the stock market, asset allocation, and risk management. Machine learning is able to handle complex large-scale data, and to discover nonlinear relationships between data by learning historical data, which is particularly important for market effects that are difficult to capture with traditional statistical methods. This paper explores the application effect and potential of machine learning technology and support vector machine algorithm in quantitative investment strategy and asset allocation. By constructing an SVM-based multi-factor stock selection model and asset allocation strategy, and analyzing their performance in the actual financial market, the effectiveness of machine learning in improving the quality of investment decisions, optimizing the efficiency of asset allocation and enhancing the ability of risk management is verified. The study also evaluates the stability and adaptability of the strategy in different market environments through systematic model back testing, in order to provide scientific decision-making tools for financial institutions and individual investors.

In China and abroad, the application of machine learning, especially support vector machines (SVMs) in quantitative investment and asset allocation has become a research hotspot. Zhang et al. (2020) explained policy-based asset allocation facilitated by machine learning and emphasized the importance of interpretative machine learning in financial

*Corresponding author's Email: yanxue_li@hotmail.com

decision-making [1]. Sahu and Kumar (2024) developed a portfolio rebalancing model using support vector machines to optimize asset allocation and proved the effectiveness of SVM in processing large-scale financial data [2]. Lee et al. (2019) explored global stock market investment strategies by using machine learning technology through financial network indicators, showing the application potential of advanced machine learning technology in the global capital market [3]. Bradrania and Neghab (2022) discussed the use of neural networks for state-dependent asset allocation, providing a new dynamic asset allocation method for financial markets [4].

Kaczmarek et al. (2022) studied the target volatility strategy based on recurrent neural networks and proposed the concept of "fake safe assets", which is crucial for understanding the performance of complex financial products under different market conditions [5]. Siade et al. (2020) used a Gaussian process machine learning model combined with group theory to plan groundwater distribution. Although this is in a non-financial field, its methodology provides a potential interdisciplinary application perspective for financial models [6]. Arpaci et al. (2024) developed and verified the Investment Strategy Scale (ISS), which is used to evaluate short-term and long-term investment strategies, and demonstrated the practical application of quantitative tools in investment decision-making [7]. Ryou et al. (2020) used a hidden Markov model to create a momentum investment strategy; their research results confirm the effectiveness of machine learning in identifying changes in market status [8].

Li and Sun (2020) discussed the intelligent stock investment strategy based on support vector machine parameter optimization algorithm and emphasized the key role of parameter selection in improving SVM performance [9]. Han and Yao (2023) explored the portfolio selection and risk prediction of financial market based on the SVM algorithm, and verified its ability to process financial data in a complex market environment [10]. Silva et al. (2024) used the SVM model to pre-select stocks for portfolio optimization, and their research further confirmed the practicability and efficiency of SVM in screening high-performance stocks [11]. Together, these studies demonstrate the wide application and far-reaching impact of SVM and other machine learning tools in modern finance, especially in regard to quantitative investment strategies and asset allocation.

Given the constant changes in financial markets and the rapid growth of data information, traditional investment strategies face many challenges in terms of efficiency and effectiveness [12]. The introduction of machine learning provides a new perspective and method for quantitative investment and asset allocation, especially the support vector machine algorithm that shows unique advantages in pattern recognition and prediction. The significance of this study is that the application of SVM algorithm in quantitative investment strategy can deepen the understanding of the function of SVM algorithm in financial markets and improve the scientific and practical aspect of strategy. The research results can provide financial practitioners with effective investment decision support, increase the accuracy of asset management and risk control capability, and contribute theoretical and practical experience to the development of the financial technology field.

2. RELATED THEORIES

2.1 Quantitative Investment and Asset Allocation

2.1.1 Introduction and Advantages of Quantitative Investment

Quantitative investment is an investment method based on a mathematical model, statistical analysis and computer technology. Its core function is to identify the buying and selling opportunities in the market by means of algorithms. Compared with traditional investment methods based on judgment and experience, quantitative investment can provide greater efficiency and accuracy when processing large amounts of data [13]. Quantitative investment greatly reduces the impact of human emotions or biases on investment decisions through programmed automation, and at the same time improves the execution speed and timeliness of transactions.

One of the main advantages of a quantitative investment strategy is its ability to execute complex mathematical models to capture small movements in the market, which is nearly impossible for traditional investors. For example, it can use high-frequency data to develop trading algorithms that are able to complete many transactions in a very short period, profiting from small movements in the market [14]. Typically, quantitative investment strategies have a high degree of transparency because each step of the investment decision is verifiable and based on pre-set rules.

Quantitative investing also allows investors to effectively manage risk by diversifying their investments. Using advanced mathematical models and historical data, quantitative strategies predict correlations and volatility between different assets to help build an optimal portfolio. The automated nature of a quantitative investment strategy enables it to quickly adapt to changes in the market environment and adjust the portfolio in time to cope with market fluctuations.

2.1.2 Asset Allocation

Asset allocation is a core part of investment management and refers to the process of allocating investment amounts among different types of investment assets (such as stocks, bonds, cash, etc.). This process is based on an assessment of the expected returns and risks of different asset classes with the aim of optimizing the risk-return relationship of a portfolio [15]. The theoretical basis of asset allocation is derived mainly from modern portfolio theory, which emphasizes the diversification of investment to reduce risks and improve the overall efficiency of the portfolio.

In practice, asset allocation involves not only choosing different asset classes, but also investing in assets in different geographical regions, industries, and risk levels. The right asset allocation can effectively diversify the risk of a particular market or economic cycle and hedge against potential market uncertainty and economic volatility. Based on an investor's risk tolerance, investment horizon, and financial goals, asset allocation can help investors achieve their long-term financial and investment goals.

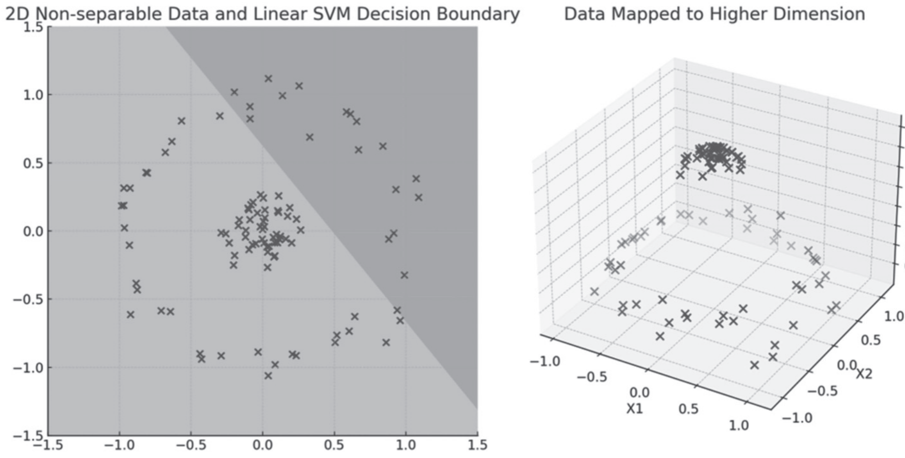

Figure 1 Mapping of low-dimensional to high-dimensional space.

With the development of financial markets and the proliferation of investment instruments, asset allocation strategies are constantly evolving. The traditional static asset allocation is more concerned with the one-time asset allocation decision, while the dynamic asset allocation strategy adjusts the asset allocation ratio according to the changes in market conditions, and emphasizes adaptability and flexibility. Factor investment has become a part of the modern asset allocation strategy, which systematically allocates assets by identifying basic factors that affect asset prices, such as value, scale, momentum, etc., to obtain excess returns.

Therefore, intelligent asset allocation is the key to achieving investment objectives, controlling risks and improving investment efficiency. Through scientific asset allocation, the expected returns from the investment portfolio can be effectively improved without increasing the investment risk.

2.2 Machine Learning

2.2.1 Machine Learning Concepts

Machine learning is a branch of artificial intelligence that involves the use of algorithms and statistical models that enable computer systems to automatically learn from data and improve without requiring explicit programming instructions. Through the training process, machine learning models can recognize complex patterns and relationships to provide predictions or decision support that are often difficult for human analysts to understand [16]. This field uses large amounts of data to train models so as to make accurate predictions or decisions based on new, unseen data.

Generally, there are three types of machine learning: supervised learning, unsupervised learning and reinforcement learning. Supervised learning predicts outcomes by training datasets containing known answers, while unsupervised learning deals with unlabeled data, aiming to discover underlying patterns in the data. Reinforcement learning trains the model to seek the best strategy among a series of options through a reward mechanism.

Machine learning has a unique advantage when working with large data sets, adapting to new data and continuously optimizing its performance over time. Therefore, machine learning technology is widely used in many fields such as image recognition, natural language processing, medical diagnosis and financial market prediction, and has become a key driving force in the development of modern science and technology. Through the effective use of machine learning, the accuracy and efficiency of task processing can be greatly improved.

2.2.2 Support Vector Machine

Machine learning is a technique that enables computer systems to learn from experience and make predictions and decisions about data without needing to be explicitly programmed. At its core is the development of algorithms that enable the system to automatically extract information and use the data for learning. In financial quantitative analysis, bioinformatics, market trend prediction and other fields, machine learning has shown outstanding application value [17].

Support vector machines are a supervised learning model widely used in machine learning for classification and regression analysis. The goal of an SVM is to find a hyperplane to classify different data points. In two dimensions, this hyperplane can be a straight-line separating data points belonging to different categories.

The strength of an SVM is its powerful kernel technique, which allows data to be mapped from the original space into higher-dimensional feature spaces, so that data that is linearly indivisible in the original space can be linearly separated by hyperplanes in the new space. This mapping is achieved by selecting suitable kernel functions, commonly used kernel functions include linear kernel, polynomial kernel, radial basis function (RBF), and sigmoid kernel.

For example, consider a simple linearly indivisible case where data points cannot be separated by a straight line in a low-dimensional space. By applying nonlinear mapping to a high-dimensional space, these data points can be effectively separated by a hyperplane. Let $\mathbf{x}_i$ and y_i be the training sample and label, $y_i \in \{1, -1\}$. The goal of SVM is to find the parameters $\mathbf{w}$ and b that separate the hyperplane $\min_{\mathbf{w},b} \frac{1}{2}\mathbf{w}^T\mathbf{w}$, subject to $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1, \quad \forall i$.

As shown in Figure 1, the figure on the left shows data sets that cannot be effectively separated by a linear SVM in a two-dimensional space. The figure on the right shows the same data set after nonlinear mapping (here, radial basis

Table 1 SVM stock selection model and asset allocation feature labels.

Category	Indicator	Content
Financial Ratios	P/E Ratio	Market value divided by net profit
	P/B Ratio	Market value divided by book value
Stock Price Technical Indicators	Relative Strength Index	Reflects the strength of stock prices
	Moving Average	30-day stock price move
Market Sentiment Indicators	Trading Volume	Reflects market activity
Macroeconomic Indicators	GDP Growth Rate	Annual economic growth rate

function, RBF) to a higher dimensional space. In the mapped three-dimensional space, the data points can be effectively separated by a hyperplane (although the hyperplane is not shown in the figure), showing that with appropriate kernel functions, support vector machines can handle linearly indivisible data problems in the original space. This feature makes an SVM particularly valuable in a variety of practical applications, especially when the feature relationships are complex and not easy to identify intuitively.

3. USE MACHINE LEARNING SVM ALGORITHM TO CONSTRUCT MULTI-FACTOR QUANTITATIVE STOCK SELECTION MODEL AND ASSET ALLOCATION

3.1 Extraction of SVM Stock Selection Model and Asset Allocation Feature Labels

When constructing a stock selection model and asset allocation strategy of an SVM, feature label extraction is one of the key steps. In order to ensure that the model can perform consistently under a variety of market conditions, the time frame for feature extraction is usually set to the last five years (2018–2023), covering bull, bear and market volatility periods, as shown in Table 1.

(1) Price/Earnings ratio: The equation is market value divided by net profit. For example, a company has a P/E ratio of 15.67 in 2018, 14.32 in 2019, 16.48 in 2020, 13.89 in 2021, and 17.05 in 2022.

(2) Price-to-book ratio: The equation is market value divided by book value. For example, the price-to-book ratio of a company over the years is 1.23, 1.35, 1.28, 1.22, 1.31, respectively.

(3) Relative Strength index: a technical indicator that reflects the strength of stock prices, and the equation involves the average return and average loss ratio. For example, the RSI of a company over the years are 37.58, 45.76, 42.89, 50.27, and 48.63, respectively.

(4) Moving averages: Commonly used and include 30-day moving averages. Suppose that the 30-day moving average of a company's stock price is 22.56 in 2022 and 23.89 in 2023.

(5) Trading volume: Indicators reflect market activity. For example, a company's average monthly trading volume was 254 million shares in 2018, 268 million shares in 2019, 277 million shares in 2020, 260 million shares in 2021, and 282 million shares in 2022.

(6) GDP growth rate: an important macroeconomic indicator that affects the overall market performance. Suppose annual GDP growth from 2018 to 2022 is 2.98%, 3.01%, −0.32% (recession), 2.15%, and 3.25%, respectively.

By constructing the key feature labels used in the stock selection model of an SVM (Table 1), financial ratios, stock price technical indicators, market sentiment indicators and macroeconomic indicators are covered. Financial ratios such as P/E and P/B provide insight into a company's financial position; stock price technical indicators, including relative strength index and moving average, reflect the market performance and trend of the stock; market sentiment shows the activity of the market through trading volume; macroeconomic indicators such as GDP growth rate reveal the impact of the economic environment on the market. Together, these characteristics constitute a comprehensive data framework designed to assess the future performance of stocks and ensure the validity and comparability of the data in the model through standardized processing. This analytical framework helps to assess the potential value of stocks under different market conditions, thereby supporting investment decisions.

3.2 Division of SVM Stock Selection Model and Asset Allocation Training set and Test set

When constructing a stock selection model and asset allocation strategy of an SVM, the division of the training set and the test set is very important, because it is directly related to the generalization ability and practical application effect of the model. This research dataset covers comprehensive financial market data from 2018 to 2023, with data from 2018 to the end of 2021 used as a training set and data from 2022 used as a test set. This partitioning allows the model to learn on historical data and validate its predicted performance on recent data.

The training set included four years of financial data, market sentiment indicators, stock price technical indicators, and macroeconomic data, which were carefully preprocessed and feature-engineered to form thousands of features for training

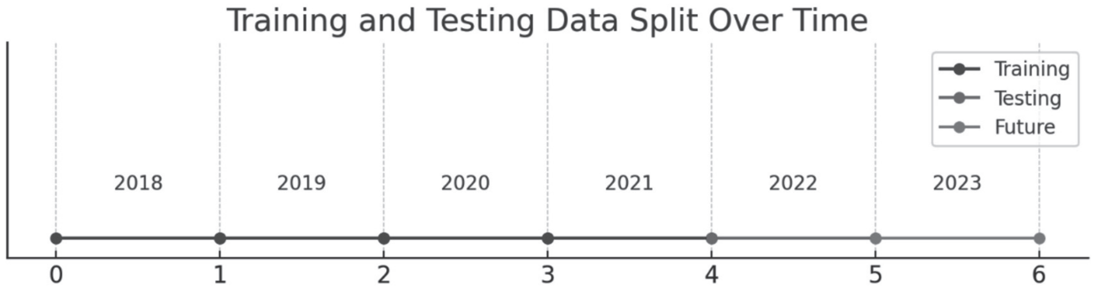

Figure 2 Time division of training set and test set.

Table 2 Evaluation results of different parameters.

Kernel Function	C	γ	Degree	Cross-Validation Average Accuracy (%)
Linear	0.01	N/A	N/A	82.3
	0.1	N/A	N/A	84.6
	1	N/A	N/A	85.2
	10	N/A	N/A	85.0
Polynomial	1	0.1	2	83.7
	1	1	3	85.9
	1	0.01	4	84.1
	10	0.1	3	86.4
Radial Basis	1	0.01	N/A	87.5
	1	0.1	N/A	88.3
	1	1	N/A	86.7
	10	0.1	N/A	89.1

SVM models. The test set includes a year's worth of similar data to assess the model's performance and accuracy in practice.

To ensure the representativeness and independence of the training and testing data, the time series method is used to divide the data to prevent data leakage. This means that all data points in the training set precede any data points in the test set, ensuring fair and valid model evaluation. The ability of the model to adapt to new data can be tested, while overfitting is controlled. This segmentation strategy also simulates the situation in the real investment environment, where investors make investment decisions based on past data and experience and verify the validity of these decisions in the future market environment.

As shown in Figure 2, the training set covers data from 2018 to 2021 to build and tune the SVM model, while the 2022 data is used as a test set to evaluate the model's actual performance and predictive power. The 2023 data is labeled 'future data', indicating that it can be used to further validate and test the robustness and fitness of the model in real time. This division ensures the effectiveness of model training and testing, while also simulating the temporal and forward-looking nature of real-world investment decisions.

3.3 SVM Stock Selection Model and Asset Allocation Parameter Optimization

Parameter optimization is a key step taken to improve the performance of the SVM stock selection model and asset allocation strategy. The goal of optimization is to select the right combination of parameters to achieve the highest prediction accuracy and the best risk-adjusted rate of return. Common SVM parameters include kernel types, regularization parameters, kernel parameters (such as the role of γ in radial basis functions), and relaxation variables. To systematically evaluate the effects of different parameter configurations on model performance, an extensive grid search and cross-validation were conducted, using the following parameter grids:

(1) Kernel function type: linear, polynomial, radial basis

(2) C (regularization parameter): 0.01, 0.1, 1, 10

(3) γ (free parameter of kernel function, used only for radial basis and polynomial kernel): 0.01, 0.1, 1, 10

(4) Degree of polynomial kernel (valid only for polynomial kernel): 2, 3, 4

The effect of each parameter combination is evaluated using 50-fold cross-validation, which divides the data set into five parts and takes turns using four of them for training and the remaining one for testing. This effectively evaluates the model's performance on unseen data and reduces the risk of overfitting.

As shown in Table 2, the radial basis kernel function performs best for a range of C and γ parameters. When C is 10 and γ is 0.1, the average accuracy of cross-validation reaches 89.1%, which shows better performance than other parameter combinations.

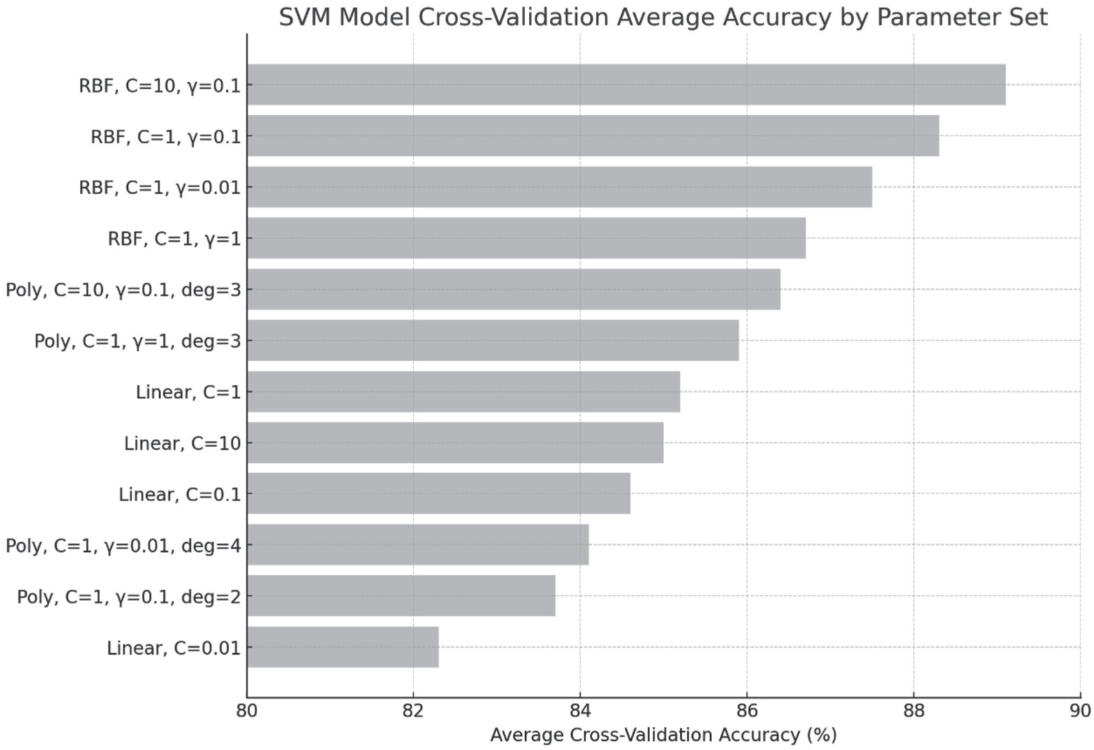

Figure 3 Average accuracy of cross-validation of models in each phase.

As shown in Figure 3, the average accuracy of model cross-validation in each phase varies significantly according to the combination of parameters. The figure clearly shows the effect of different kernel functions and parameter settings on model performance. The radial basis kernel function achieves the highest average accuracy of 89.1% when C is 10 and γ is 0.1, indicating that this combination of parameters is the most effective in the data set and problem setup in this study. This visual representation helps to understand and select the model parameter configuration that is most suitable for the practical application, thereby optimizing the performance of the stock selection model and asset allocation strategy.

4. BACK TEST RESULTS AND ANALYSIS OF SVM MULTI-FACTOR QUANTITATIVE STOCK SELECTION MODEL AND ASSET ALLOCATION

4.1 SVM Multi-Factor Quantitative Stock Selection Model and Asset Allocation Stock Selection Logic

In the SVM multi-factor quantitative stock selection model and asset allocation, a comprehensive examination of multiple financial indicators, market data and macroeconomic factors is adopted to carry out stock selection and asset allocation. The model is based mainly on the algorithm framework of an SVM and learns to identify those stocks and asset classes that may bring excess returns through training data. The stock selection logic is as follows.

(1) Fundamental analysis: Use financial ratios such as price-earnings ratio, price-to-book ratio, ROE (return on equity), and operating income growth rate to evaluate the value and growth potential of the company. Models predict future performance by analyzing the relationship between historical data and stock performance for these indicators.

(2) Technical analysis: Use technical indicators such as moving average, Relative Strength index (RSI), MACD, etc., to capture market trends and reversal signals. These technical indicators help determine the entry and exit timing and enhance the market adaptability of the model.

(3) Market sentiment analysis: Consider indicators such as trading volume and price volatility to assess changes in market sentiment. Market sentiment is often a predictor of short-term market movements and is particularly critical for capturing investment opportunities.

(4) Macroeconomic factors: macroeconomic data such as GDP growth rate, inflation rate and unemployment rate are incorporated into the model to analyze the impact of these macroeconomic indicators on the overall performance of the market.

The asset allocation strategy is based on the above stock selection results and the correlation between and volatility of asset classes. The model not only selects the stocks that are expected to perform the best, but also reduces risk by diversifying investments, including allocation among stocks in different industries and geographies, and dynamic adjustment among different asset classes such as stocks, bonds, and cash.

4.2 Setting of SVM Multi-Factor Quantitative Stock Selection Model and Asset Allocation Back Test Environment

When setting up the back test environment of the SVM multi-factor quantitative stock selection model and asset allocation, a set of detailed frameworks is adopted to simulate the real investment environment to ensure the reliability and practicality of the back test results. The back test environment includes the extensive collection and collation of historical data covering stock prices, trading volumes, financial statements, macroeconomic indicators, and more, covering a time span from 2018 to 2022.

(1) The integrity and consistency of data must be ensured as this is essential for the accuracy of the back test. All data is taken from reliable financial databases, including exchange data, financial news agencies and publications from macroeconomic statistics departments. In the data pre-processing stage, missing value processing, outlier detection and correction, and data standardization are carried out to ensure that data from different sources and formats can be uniformly processed and used for model training and testing.

(2) To implement the back measurement method, an event-driven back test system is adopted, which can simulate the factors such as order execution, transaction cost, market impact and time delay in actual trading. The back test system is configured with various fees that may be encountered in actual trading, including but not limited to commission fees, stamp duty, trade slip points, etc., to simulate the impact of an actual trading environment on strategy performance.

(3) The back test is set to the daily frequency, which means that the model will re-evaluate and adjust the position according to the latest data at the end of each day. This not only captures short-term fluctuations in the market, but also makes it possible to better assess the adaptability of the strategy to rapid market changes. Throughout the back test, various risk metrics such as maximum retracements, Sharpe ratios, etc., are also monitored to fully evaluate the performance of the strategy under different market conditions.

4.3 Return Test and Risk Evaluation Indicators

4.3.1 Strategy Return Index

Among the return and risk evaluation indexes, the strategy return index is one of the key factors used to evaluate the performance of investment strategy.

(1) Total rate of return

Total return measures the absolute return of an investment strategy over the entire back test period and is calculated with Equation (1).

$$\text{Total Return} = \frac{\text{Portfolio Value}_{\text{end}} - \text{Portfolio Value}_{\text{start}}}{\text{Portfolio Value}_{\text{start}}} \tag{1}$$

Where Portfolio $\text{Value}_{\text{start}}$ and Portfolio $\text{Value}_{\text{end}}$ represent the total asset value of the policy at the beginning and end of the back test, respectively.

(2) Annualized Return

The annualized rate of return converts the total return of the strategy into an annual equivalent value, which is suitable for comparing the performance of the strategy over different periods of time. The calculation is shown in Equation (2):

$$\text{Annualized Return} = (1 + \text{Total Return})^{\frac{365}{d}} - 1 \tag{2}$$

d indicates the number of days in the back test period.

(3) Cumulative income

The daily cumulative value of the strategy during the entire back test period reflects the long-term performance of the strategy, and is calculated with Equation (3):

$$\text{Cumulative Return} = \prod_{t=1}^{T}(1 + r_t) - 1 \tag{3}$$

r_t 0 indicates the daily return rate, and T indicates the total number of days of back test.

(4) Maximum retracement correction return

The maximum retracement modified return is a comparison of the annualized rate of return with the maximum retracement to evaluate the return per unit of risk taken, calculated with Equation (4):

$$\text{Calmar Ratio} = \frac{\text{Annualized Return}}{\text{Maximum Drawdown}} \tag{4}$$

The maximum retracement is the maximum loss of the strategy from the highest peak to the lowest trough.

4.3.2 Strategy Risk Indicators

In a quantitative investment strategy, understanding and managing risk is equally important. The following are some commonly used strategic risk indicators and the respective equations used to calculate the potential risks of investment strategies:

(1) Maximum retracement

The maximum retracement is a measure of the maximum loss a strategy can suffer. It shows the largest drop in strategy from peak to trough, calculated with Equation (5):

$$\text{Maximum Drawdown} = \min\left(\frac{\text{Portfolio Value}_{\text{min}} - \text{Portfolio Value}_{\text{max}}}{\text{Portfolio Value}_{\text{max}}}\right) \tag{5}$$

where Portfolio $\text{Value}_{\text{max}}$ and Portfolio $\text{Value}_{\text{min}}$ represent, respectively, the highest value of the portfolio during the period under consideration and the subsequent lowest value.

(2) Volatility

Volatility is the standard deviation measuring the fluctuation of investment return, reflecting the instability and risk of strategy return. The calculation done with Equation (6):

$$\text{Volatility} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(r_i - \bar{r})^2} \tag{6}$$

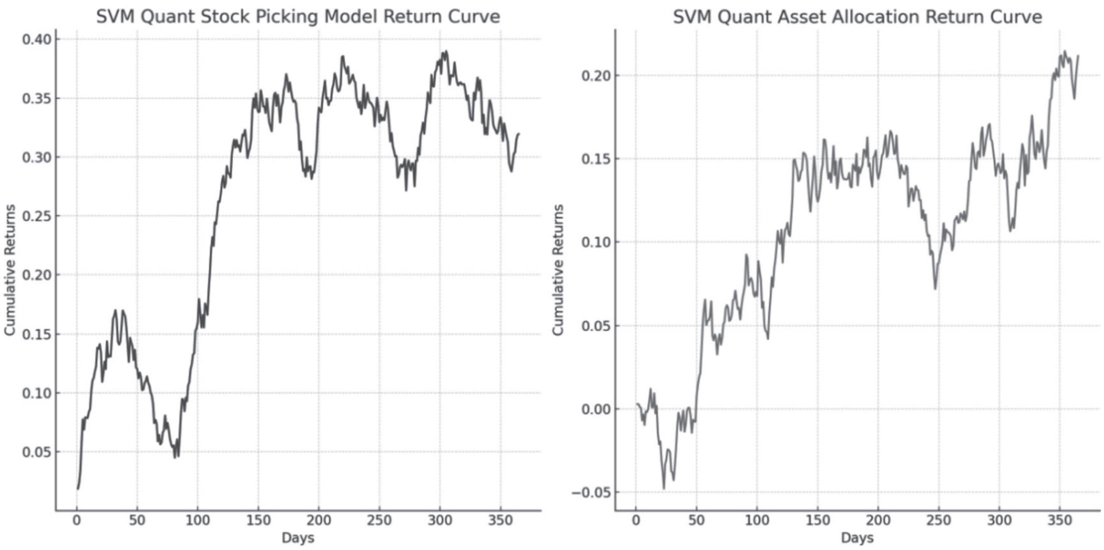

Figure 4 SVM quantitative stock selection model and asset allocation return curve.

where r is the daily return rate, $\bar{r}$ is the average daily return rate, and N is the total number of days during the back test period.

(3) Sharpe ratio

The Sharpe ratio is used to assess the excess return per unit of total risk, calculated with Equation (7):

$$\text{Sharpe Ratio} = \frac{\text{Annualized Return} - \text{Risk-Free Rate}}{\text{Volatility}} \tag{7}$$

where Risk-Free Rate can usually be used as a reference for long-term treasury bond rates.

(4) Sotino ratio

The Sotino ratio is like the Sharpe ratio but focuses only on downside risk and is calculated with Equation (8):

$$\text{Sortino Ratio} = \frac{\text{Annualized Return} - \text{Risk-Free Rate}}{\text{Downside Deviation}} \tag{8}$$

Downside Deviation is the standard deviation of adverse return fluctuations, considering only fluctuations in returns below a certain target or requirement.

4.3.3 Comprehensive Policy Indicators

When evaluating the performance of a quantitative investment strategy, in addition to the individual income indicators and risk indicators, several comprehensive indicators are commonly used to evaluate the performance of the strategy. These indicators combine the two aspects of return and risk and can indicate the efficiency and effect of the strategy more comprehensively. The following are several key strategic indicators and their respective equations:

(1) Sharpe ratio

A common measure of investment efficiency is the Sharpe ratio, which represents the amount of excess return a portfolio receives for each unit of total risk taken. The calculation Equation is shown in Equation (9) below.

$$\text{Sharpe Ratio} = \frac{\text{Annualized Return} - \text{Risk-Free Rate}}{\text{Volatility}} \tag{9}$$

where Annualized Return is the annualized rate of return of the strategy, Risk-Free Rate is the risk-free rate, which is usually replaced by the annualized rate of return of short-term Treasury bonds, and Volatility is the standard deviation of the return of the strategy.

(2) Sotino ratio

The Sotino ratio focuses on the assessment of adverse risks and considers only the volatility below the target rate of return, which is more suitable for the evaluation of the downside risk of the strategy. This is calculated with Equation (10):

$$\text{Sortino Ratio} = \frac{\text{Annualized Return} - \text{Risk-Free Rate}}{\text{Downside Deviation}} \tag{10}$$

where, Downside Deviation is the volatility calculated when the strategy's return is below the target return or the average return.

(3) Kalmar ratio

The Kalmar ratio is another measure of portfolio performance that focuses specifically on performance under extremely adverse conditions. It assesses the risk-adjusted return of the strategy by correlating the annualized return to the maximum retracement. This is calculated with Equation (11):

$$\text{Calmar Ratio} = \frac{\text{Annualized Return}}{\text{Maximum Drawdown}} \tag{11}$$

where, Maximum Drawdown is the maximum money retracement during the strategy back test.

4.4 SVM Quantitative Stock Selection Model and Asset Allocation Strategy Back Test Results

As shown in Figure 4, the return curve of SVM quantitative stock selection model is on the left, and the return curve of SVM quantitative asset allocation is on the right. Figure 4 indicates that the money growth trajectory of the two strategies during the back test shows different earnings and money volatility.

Table 3 Results of VM quantitative stock selection model and asset allocation strategy back test.

Category	Indicator Name	Stock Selection Model Backrest Results	Asset Allocation Backrest Results
Return Metrics	Total Return Rate	28.5%	34.2%
	Annualized Return Rate	5.7%	6.8%
	Cumulative Return	142.5%	171.0%
Risk Metrics	Maximum Drawdown	−15.3%	−12.9%
	Volatility	14.6%	11.8%
	Downside Risk	9.7%	7.5%
Composite Metrics	Sharpe Ratio	0.39	0.58
	Sortino Ratio	0.59	0.91
	Calmar Ratio	0.37	0.53

The return curve shows that the SVM quantitative stock selection model has some volatility, which may be due to the model's dependence on the performance of a single market or a specific class of stocks, making it more susceptible to short-term market fluctuations. Specifically, the graph shows that the stock selection model has grown rapidly during certain periods, which is most likely because the model has captured the upside opportunities of certain highly volatile stocks. However, this high volatility also carries significant downside risk, as shown by several sharp declines in the chart.

In contrast, the return curve of the asset allocation model is smoother, indicating that diversified asset allocation can effectively spread risks and reduce the impact of single-market or asset fluctuations on the overall portfolio. The asset allocation model optimizes the balance between return and risk by assigning weights among different asset classes (such as stocks, bonds, commodities, etc.) and therefore shows relatively stable growth throughout the back test period.

The back test of the SVM quantitative stock selection model and asset allocation strategy covers various return and risk indicators, and also includes a comprehensive set of performance indicators to evaluate the effectiveness and robustness of the strategy.

As shown in Table 3, in terms of total return and annualized rate of return, the total return of the asset allocation strategy was 34.2%, which exceeded the 28.5% of the stock selection model. Asset allocation strategies can take advantage of the asymmetry of risk and return between different asset classes, adapt to market changes through dynamic adjustment, and optimize returns. At the same time, the annualized return rate of the asset allocation strategy is 6.8%, which is higher than the 5.7% of the stock selection model, indicating that it has stronger long-term sustainable growth ability.

In terms of cumulative return and risk indicators, the cumulative return of asset allocation strategy reached 171.0%, which was significantly higher than the 142.5% of stock selection model. This indicates an asset allocation strategy that reduces reliance on a single market by diversifying investments. In regard to risk, the maximum retracement of the asset allocation strategy was −12.9%, lower than the −15.3% of the stock picking model, showing better capital protection ability. The volatility and downside risk of the asset allocation strategy were 11.8% and 7.5%, respectively, lower than that of the stock selection model, indicating greater resistance to market fluctuations.

Regarding risk-adjusted performance, the Sharpe ratio and the Sotino ratio further demonstrate the advantages of asset allocation strategies. The asset allocation strategy has a Sharpe ratio of 0.58, higher than the 0.39 of the stock selection model, which gives more excess return per unit of total risk taken. At the same time, the Sotino ratio also improved from 0.59 to 0.91, highlighting the superior performance of asset allocation strategies when factoring in the downside risks.

While stock selection models provide substantial returns, with regard to risk control and return stability, asset allocation strategies perform better after adjusting for risk. The increase in the Kalmar ratio, from 0.37 to 0.53 in the stock selection model, provides further evidence that the multi-asset allocation strategy offers a higher annualized return when considering maximum retracement, demonstrating the stability of its long-term performance. These results validate the effectiveness of multi-asset allocation strategies in diversifying risk and achieving solid growth.

5. CONCLUSION AND PROSPECTS

5.1 Conclusion

This research proves the effectiveness and feasibility of using machine learning technology in the field of financial investment by applying support vector machine algorithm to the construction and optimization of quantitative stock selection model and asset allocation strategy. Experimental results show that the SVM algorithm can effectively identify potential investment opportunities, and significantly improve the overall performance of investment strategy through accurate model parameter adjustment. In the field of asset allocation, the SVM model shows strong risk control ability and robust return performance.

The back test results show that, compared with traditional investment strategies, SVM-based strategies perform better according to several key indicators: total return, annualized return, and risk-adjusted return ratio. Asset allocation strategies also showed better performance on risk metrics such as maximum retracements and volatility, underscoring the potential of machine learning to improve portfolio stability and reduce risk.

This paper not only demonstrates the application value of machine learning SVM in quantitative investment, but also provides a new perspective and methodological basis for future research in this field. The results of this research are expected to provide technical support for practitioners in the financial industry to help them make more scientific and accurate decisions in the complex market environment.

5.2 Outlook

It is anticipated that in the future, quantitative investment strategies based on machine learning, especially support vector machines, will be widely used in asset allocation and stock selection models.

(1) Due to the constantly changing market environment and the rapid development of data technology, further research will be conducted to explore more complex and high-dimensional data sets to improve the predictive accuracy and operational flexibility of the model. Introducing more kinds of data, such as social media sentiment analysis, real-time data on macroeconomic indicators and the impact of global events, can provide a more comprehensive view of the model and enhance its ability to adapt to different market conditions.

(2) Considering the challenges of model overfitting and real-time trading, future research will focus on developing more effective algorithmic tuning strategies and parameter optimization techniques. This includes leveraging emerging technologies such as deep learning to strengthen the adaptability and stability of strategies. More complex asset allocation methods, such as dynamic adjustment strategies, will also be explored to better respond to rapid market movements.

(3) In the future, the research will focus on regulatory compliance and ethical issues to ensure that all algorithms are implemented in compliance with regulatory requirements and maintain market fairness and transparency. It is expected that the aforementioned measures will drive the research on applying machine learning to quantitative investment in a wider range of practical applications, and provide investors with scientific, efficient and safe investment decision-making tools.

FUNDING

No funding was received.

Conflict of Interest

The authors have no financial or proprietary interests in any material discussed in this article.

Data Availability Statement

The data used to support the findings of this study are all in the manuscript.

Author Contributions

Zhanyong Wu, Xiaohang Ma and Yanxue Li wrote the main manuscript text, prepared figures, tables and equations. Zhanyong Wu, Xiaohang Ma and Yanxue Li reviewed the manuscript.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

1. Zhang, R., Yi, C., & Chen, Y. (2020). Explainable machine learning for regime-based asset allocation. In 2020 IEEE International Conference on Big Data, 5480–5485. https://doi.org/10.1109/BigData50022.2020.9378332
2. Sahu, B. R. B., & Kumar, P. (2024). Portfolio rebalancing model utilizing support vector machine for optimal asset allocation. Arabian Journal for Science and Engineering, 1–27. https://doi.org/10.1007/s13369-024-08850-9
3. Lee, T. K., Cho, J. H., Kwon, D. S., & Sohn, S. Y. (2019). Global stock market investment strategies based on financial network indicators using machine learning techniques. Expert Systems with Applications, *117*, 228–242. https://doi.org/10.1016/j.eswa.2018.09.005
4. Bradrania, R., & Neghab, D. P. (2022). State-dependent asset allocation using neural networks. European Journal of Finance, *28*(11), 1130–1156. https://doi.org/10.1080/1351847X.2021.1960404
5. Kaczmarek, T., Bedowska-Sojka, B., Grobelny, P., & Perez, K. (2022). False safe haven assets: evidence from the target volatility strategy based on recurrent neural network. Research in International Business and Finance, *60*, 101610. https://doi.org/10.1016/j.ribaf.2021.101610
6. Siade, A. J., Cui, T., Karelse, R. N., & Hampton, C. (2020). Reduced-dimensional Gaussian process machine learning for groundwater allocation planning using swarm theory. Water Resources Research, *56*(3), e2019WR026061. https://doi.org/10.1029/2019WR026061
7. Arpaci, I., Aslan, O., & Kevser, M. (2024). Evaluating short- and long-term investment strategies: development and validation of the investment strategies scale (ISS). Financial Innovation, *10*(1), 63. https://doi.org/10.1186/s40854-023-00573-4
8. Ryou, H., Bae, H. H., Lee, H. S., & Oh, K. J. (2020). Momentum investment strategy using a hidden Markov model. Sustainability, *12*(17), 7031. https://doi.org/10.3390/su12177031
9. Li, X. T., & Sun, Y. (2020). Stock intelligent investment strategy based on support vector machine parameter optimization algorithm. Neural Computing & Applications, *32*(6), 1765–1775. https://doi.org/10.1007/s00521-019-04566-2
10. Han, X. Y., & Yao, D. Q. (2023). Exploration on portfolio selection and risk prediction in financial markets based on SVM algorithm. International Journal of Information Technology and Web Engineering, *18*(1), 1–16. https://doi.org/10.4018/IJITWE.332777
11. 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*, 105014. https://doi.org/10.1016/j.frl.2024.105014

12. Gao, Y., Sun, L. (2024). Credit risk evaluation of science and technology finance based on artificial intelligence and Bayesian algorithm. Engineering Intelligent Systems, *32*(5), 445–455, 2024.
13. Fonseca, A. R., Leles, M. C. R., Moreira, M. G., Vale-Cardoso, A. S., Pereira, M. V. L., Sbruzzi, E. F., & Nascimento, C. L. (2021). Testing the application of Support Vector Machine (SVM) to technical trading rules. 2021 15th Annual IEEE International Systems Conference (SysCon), 1–8. https://doi.org/10.1109/SysCon48628.2021.9447068
14. Ammer, M. A., Ahmed, Z. A. T., Alsubari, S. N., Aldhyani, T. H. H., & Almaaytah, S. A. (2023). Application of artificial intelligence for better investment in human capital. Mathematics, *11*(3), 612. https://doi.org/10.3390/math11030612
15. Li, Z. W., Han, J., & Song, Y. P. (2020). On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning. Journal of Forecasting, *39*(7), 1081–1097. https://doi.org/10.1002/for.2677
16. Suprihadi, E., & Danila, N. (2024). Forecasting ESG stock indices using a machine learning approach. Global Business Review, 09721509241234033. https://doi.org/10.1177/09721509241234033
17. Nti, I. K., Adekoya, A. F., & Weyori, B. A. (2020). Efficient stock-market prediction using ensemble support vector machine. Open Computer Science, *10*(1), 153–163. https://doi.org/10.1515/comp-2020-0199