Proceedings of the 3rd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2024.26965

Optimizing Portfolio Hedging Strategies for F30 and SPY

Xinzhu Li[1*], Siman Guo[2], Xin Fan[3]

[1]*Aberdeen Institute of Data Science and Artificial Intelligence, South China Normal University, Guangzhou, China,*
[2]*School of Economics, Jinan University, Guangzhou, China*
[3]*School of Physics NorthEast Normal University, Changchun, China*
**Corresponding Author. Email: lxz_bamboo@163.com*

Abstract. In this paper, we propose three optimization methods for portfolio hedging strategies for two assets, F30 and SPY, and evaluate their effectiveness by maximizing the Sharpe ratio. The first approach, based on six different SPY weighting strategies combined with rolling window optimization, identifies the optimal portfolio that achieves the highest risk-adjusted return (Sharpe ratio of 1.5090) under minimized risk conditions. The second approach utilizes Recursive Feature Elimination (RFE) and Gradient Boosted Regression (GBR) models for feature selection and forecasting by introducing richer macro-market data features, which may lead to overfitting in high volatility environments (Sharpe Ratio of 1.3156), although it displays significant advantages in capturing non-linear changes in the market. The third approach directly optimizes the objective function by expanding the rolling window to smooth short-term market volatility and reduce trading frequency and cost, resulting in a final risk-adjusted return of 1.210. The results of the study show that these approaches exhibit better portfolio optimization capabilities under different market conditions, providing new ideas for achieving more robust risk management and return optimization.

Keywords: Portfolio Hedging, Sharpe Ratio, Rolling Window Optimization, Recursive Feature Elimination (RFE), Risk Management.

1. Introduction

Effective risk management is critical in the financial markets for investors who want to achieve desired returns while minimizing potential losses. Portfolio hedging is a common strategy to minimize risk by balancing positions across various assets. This study focuses on portfolio hedging strategies for the F30 and SPY assets. SPY, an ETF tracking the S&P 500, has been widely studied in risk management; however, relatively little research has been conducted on the F30 asset in risk hedging. A deeper understanding of the synergistic effects of these two assets will provide new insights for more effective risk minimization strategies.

The primary objective of examining the hedging strategy for the F30 and SPY portfolios is to determine more effective ways to minimize portfolio risk. Traditional optimization methods, such as the Markowitz mean-variance model [1], are widely utilized in portfolio management. However,

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2024.26965

these methods typically assume static market conditions and constant asset correlations, which limits their adaptability to real-world dynamic market changes. Additionally, these methods presume that asset returns follow a normal distribution, whereas actual financial market returns often exhibit heavy-tailed distribution characteristics, potentially leading to suboptimal hedging strategies [2].

Previous research has indicated that traditional methods like mean-variance optimization struggle to handle nonlinear market changes and high volatility [2,3]. While some studies have attempted to incorporate machine learning techniques into portfolio optimization, these approaches still encounter challenges related to model interpretability and stability in rapidly evolving market environments. The current optimization strategies are often carried out in isolation, and lack of systematic exploration of the combination effects of different strategies, which limits the overall effect of portfolio optimization to a certain extent [4].

To address these challenges, this paper presents three methods for optimizing F30 and SPY portfolio hedging strategies. The method in this paper combines traditional risk minimization techniques, machine learning models and optimization of objective functions to find better hedging strategies. We choose the Sharpe ratio as the objective function and evaluation method because it provides a comprehensive measure of risk-adjusted returns, reflecting the balance between risk and return. By maximizing the Sharpe ratio, the return per unit of risk of a portfolio can be optimized, making it a suitable metric for comparing the effectiveness of different hedging strategies, especially for scenarios under high market volatility and uncertainty [5].

Risk minimization-based approach, we evaluate five different SPY weighting strategies and 15 different rolling window sizes, using an improved Markowitz model to deal with nonlinear changes and high volatility in the market, to find the optimal combination that can obtain the highest Sharpe ratio, providing new empirical support for the application of rolling windows in risk management.

Hence, this paper develops a machine learning algorithm namely Recursive Feature Elimination [6] and Gradient Boosted Regression [7] for capturing nonlinearity in the market data. The RFE method removes features with lesser prediction capabilities hence increasing the model's generality [8], while the GBR model is designed to change in the marketplace for a better investment approach.

An optimization-based approach is also developed, which employ a new designed objective that can adapt to changes in market positions. This enables a superior optimization framework over standard ones such as the Sharpe ratio that are inadequate in volatile markets. The new approach focuses on lower fluctions and a wider diversification in the asset's weights than the previous methods in order to reduce on the extreme weighting impact on the risk.

2. Literature review

The final say is that since the financial markets have adapted complexity, machine learning when used for portfolio optimization and objective function optimization has offered investors more robust solutions. By using the machine learning methods, the dimensionality of the data has increased for analyzing the complexity or non-linearity of the microstructure of the financial data. This capability helps in better processing of high dimensionality data along with better accuracy to foresee future trends. On the same note, improving the solution to the objective function incorporating regulation terms and penalties as well as constraints enhances portfolio resilience and flexibility in changing economic environments [8].

The use of machine learning in enhancing the portfolio predictions is incredibly effective in the current world. In their study, Gu, Kelly, and Xiualso established that gradient augmented regression (GBR) model outperforms standard linear models in terms of the predictive accuracy on stock returns and the ability to handle risk, especially in datasets with sparsity and non-linear

DOI: 10.54254/2754-1169/2024.26965

relationships. Likewise, Feng et al [7]. Feng et al discussed about the advantage of using the feature selection methods including the RFE for the financial forecasting. This make the model to be stable in predicting the results while at the same time decreasing the number of features that are used in model [6].

In this case, an important and popular type employed for tuning objective functions is critical to reduce the problem of extreme asset weights on portfolio allocation. In 1996, Tibshirani was the first to extend a regularization procedure in the Lasso perspective making it less centrifugal in variable selecting and reducing high weight of assets [9]. This approach helps to avoid concentration of assets, and liabilities, in the same investments within a given period.

Moreover, Dai and Wen expounded a generalized sparse and stable portfolio optimization technique based on lasso with extra term to improve the stability of the portfolio [10]. This method solves the problem of portfolio sparse and also make the weight less volatile in time series, which also can decrease the transaction cost. Further, by adopting L1 and L2 regularization, the extreme weights were controlled through Bertsimas et al. to boost the portfolio stability. Subsequently, DeMiguel et al. proposed a smoothing penalty approach to reduce variations in weights between successive portfolios and decrease costs of rebalancing still further [11,12]. More similar to what has been said above, optimization of portfolio construction can be enhanced through the use of optimization algorithms. Other researchers showed that SLSQP and TNC algorithms improve the speed when performing high-frequency trading constraint problems, and these algorithms are relevant in optimizing the objective function.

In summary, machine learning techniques combined with adjusted objective functions provide promising research directions for portfolio optimization. By introducing feature selection, regularization terms, and advanced optimization algorithms, portfolios become more robust to market changes while enhancing the flexibility and accuracy of forecasts. Although these methods show great potential, further research and improvements are needed to cope with rapidly changing markets and to improve the stability and interpretability of the models.

3. Optimized portfolio construction based on risk minimization

3.1. Data preprocessing

• **Delete anomalies and missing values:** Delete anomalies(N/A) and missing values in the data, such as when the return of f30 is 0.99.

• **Quantify the three states of Position:** Quantify the states in the Position column to 1, -1 and 0, which represent holding a long position, holding a short position and holding no position respectively.

• **Filter data:**

Accurate position data is essential for calculating the hedge ratio. In order to ensure that the data used is meaningful and effective, the position status of the current date is compared with the previous day, and invalid data caused by the position being 0 on the previous day is eliminated. This process helps to avoid the introduction of noise data when calculating the hedge ratio, thereby improving the accuracy of the results.

Using the same screening method to screen the data of long and short positions separately can ensure that these two different market behaviors are accurately distinguished during analysis, which is essential for understanding market dynamics and formulating hedging strategies. The data of long and short positions complement each other to form complete valid position data, which enables the long and short sides of the market to be fully considered when calculating the hedge ratio, thereby

DOI: 10.54254/2754-1169/2024.26965

drawing more reliable conclusions. By eliminating invalid data, the data quality can be significantly improved, making subsequent analysis and calculation more accurate and reliable; at the same time, noise interference is reduced, making the calculation results of the hedge ratio more stable and reliable. Accurate long and short position data helps to better understand market behavior and formulate more effective hedging strategies.

3.2. Calculating hedging ratios and portfolio return to minimize risk

The hedging ratio represents the number of derivative contracts used to hedge the risk of the underlying asset, and the optimal hedging ratio is determined by establishing a linear relationship between the two assets. Only by determining an appropriate hedging ratio can the risk of price changes in the underlying asset be effectively offset. The most commonly used hedging ratio calculation method is based on the minimum variance method. The derivation process is as follows:

$$V = X - \beta Y \tag{1}$$

$$Var(V) = Var(X - \beta Y) = (\sigma_X)^2 + \beta^2(\sigma_Y)^2 - 2\beta\rho\sigma_X\sigma_Y \tag{2}$$

In order to minimize the variance and find the optimal hedging ratio β, take the derivative of β and set it equal to 0. Next, for indicating the direction of hedging, a negative sign is added before the formula, and the formula for portfolio construction is built:

$$2\beta(\sigma_Y)^2 - 2\rho\sigma_X\sigma_Y = 0 \tag{3}$$

$$\beta = -\rho \times \frac{\sigma_X}{\sigma_Y} \tag{4}$$

$$Portfolio\ Construction = F30_{ret} + \beta \times SPY_ret \tag{5}$$

Here,

- V : The value of the hedging portfolio.
- X : The price of the underlying asset(F30).
- Y : The price of derivatives (SPY).
- β : The hedging ratio.
- $(\sigma_X)^2$: The variance of the underlying asset return.
- $(\sigma_Y)^2$: The variance of derivative returns.
- ρ : The correlation coefficient between the underlying asset and derivative returns.

3.2.1. Calculating hedging ratios based on six methods

This part adopts the rolling window data processing method and uses 6 different data acquisition methods to calculate β. The rolling window refers to a fixed-size window that slides on the data sequence, and each time a fixed-size window is slid for analysis. The first five methods use the same

DOI: 10.54254/2754-1169/2024.26965

window size (5, 10, 15, 20, 25, 30, 45, 60, 75, 90, 120, 180, 210, 240). The window sizes used by the sixth method are: 60, 90, 120, 180, 210, 240.

For each window size, the following 6 data processing methods are used. The resulting β value is used for portfolio construction on the next trading day.

Table 1. 5 methods & description for calculating hedging ratio

β	Methods	Description
β1	Hedging ratio calculated by all data	Calculate β using all data points without considering positions.
β2	Hedging ratio calculated by only valid position data	Calculate β using only valid position data; if valid data is insufficient (less than 2), β is marked as invalid.
β3	Hedging ratio calculated by valid position data equal to window size	Use only valid position data points; calculate β with a rolling window and apply it to the next trading day.
β4	Hedging ratio calculated by long and short valid position data	Calculate β separately for long and short positions; β values for long and short does not appear on the same day.
β5	Hedging ratio by long and short valid position data equal to window size	Extract long and short valid position data separately; calculate β using a rolling window after removing invalid data.

β6: Hedging Ratio by Linear Regression

To better estimate beta while integrating long, short, and non-position holding situations into a comprehensive framework beyond mere mathematical definitions, a new function is defined as follows

$$Y = \beta \bullet X_1 + \beta_l \bullet X_{long} + \beta_s \bullet X_{short} \tag{6}$$

Where:

• Y: A dependent variable that represents the return of F30.

• X_1: An independent variable that represents the returns of SPY when the investor does not hold F30.

• X_{long}: An independent variable that represents the returns of SPY when the investor holds a long position of F30.

• X_{short}: An independent variable that represents the returns of SPY when the investor holds a long position of F30.

Beta values are obtained for three distinct scenarios through multiple linear regression analysis. Since the raw β typically considers only the relationship between market fluctuations and a single asset, failing to account for risk variations under long and short positions. This static β is inadequate for providing precise risk hedging in complex market conditions, necessitating a dynamic adjustment of Beta values to capture risk more accurately under different investment states.

• β : This is a baseline risk measure when the investor holds no position in the target asset (e.g., f30). It represents a fundamental reference point that captures the interaction between SPY and f30 under normal market conditions without directional exposure.

• $\beta + \beta_l$: This adjusted Beta reflects the increased sensitivity of SPY to f30 returns when the investor holds a long position. It will capture the additional market exposure under long positions, thereby providing a more precise hedging ratio tailored to bullish scenarios.

DOI: 10.54254/2754-1169/2024.26965

• $\beta + \beta_s$: This adjusted Beta is used to characterize the interaction between SPY and f30 when the investor holds a short position. It aims to better reflect the additional market risk associated with short positions, providing an appropriate hedging strategy in bearish conditions.

This approach ultimately provides a new Sharpe Ratio value for the portfolio under this method.

Based on the above six different methods of calculating hedging ratio, considering 5 different aspects: all, positions, longs, shorts, positions*size, we can get 30 different return values for portfolio returns.

3.2.2. Evaluate portfolio return based on five methods

Table 2. Five methods & description for evaluating portfolio return

Method	Description
All	Selects all portfolio returns to calculate the average and volatility for each window.
Position	Uses data with valid positions to calculate the average and fluctuation in each window.
Long-Position	Selects data with a valid long position and multiplies the portfolio return by 1 to compute the average and fluctuation for each window.
Short-Position	Selects data with a valid short position and multiplies the portfolio return by -1 to compute the average and fluctuation for each window.
Position *Size	Utilizes valid data for both long and short positions, multiplying returns based on the previous day's F30% and recalculating averages and fluctuations in each window based on this adjusted return.

Using β1 as an example, the table below presents the five evaluation methods for the portfolio constructed based on the beta1 approach. Detailed results of all methods will be provided in Appendix.

Table 3. Average return based on β1

Window Size	All	Position	Long	Short	Position+Size
5	0.00023898	0.00231881	0.004903588	-0.001878331	0.000763092
10	0.000538909	0.002482681	0.007149608	-0.001685749	0.000757684
15	0.000632836	0.002563072	0.007594289	-0.00170217	0.000779438
20	0.000556116	0.002513004	0.007481606	-0.001661419	0.000759433
25	0.000520596	0.002509368	0.007249865	-0.001695206	0.000761745
30	0.000541143	0.002520534	0.007270662	-0.001702358	0.000758022
45	0.000420163	0.002599213	0.006938078	-0.001843446	0.000759797
60	0.000409845	0.002669437	0.007087944	-0.00189364	0.000772327
75	0.000439031	0.002541857	0.007103413	-0.001734828	0.000753859
90	0.000415034	0.002515833	0.007143157	-0.00169122	0.000740304
120	0.000340746	0.002489463	0.007300168	-0.001622164	0.000750695
150	0.000236533	0.002584848	0.00717313	-0.001753695	0.000768451
180	0.000158355	0.002561823	0.007004242	-0.001756043	0.000770987
210	6.05435E-05	0.002672651	0.007078975	-0.00186605	0.000793227
240	-2.44655E-05	0.002745751	0.006873969	-0.001982351	0.000820605

DOI: 10.54254/2754-1169/2024.26965

4. Optimized portfolio construction based on machine learning

4.1. Data preprocessing

Enriching the dataset by calculating and adding multiple new features is essential for improving the model's predictive accuracy. By incorporating a broader range of market indicators and technical variables, the model gains a more comprehensive understanding of market dynamics, allowing it to make more accurate predictions of stock returns under varying conditions.

4.1.1. Merging macroeconomic data

First, volume data for the VIX (volatility index) and SPY are merged into the main dataset. These market data can provide a reference to market sentiment and volatility and help to understand the market environment more fully.

• **VIX Daily Returns:** The daily returns of the VIX are calculated to provide information on the dynamics of market volatility.

• **VIX Rolling Standard Deviation:** Calculates the 10-day rolling standard deviation of the VIX to measure changes in short-term volatility.

4.1.2. Adding technical indicators

In order to capture the market trend and momentum of SPY and F30, we have calculated and added the following technical indicators:

• **RSI (Relative Strength Index):** measures the overbought or oversold state of the market and helps identify possible reversal points.

$$RSI = 100 - \frac{100}{1 + RS} \tag{7}$$

Where $RS = \frac{Average\ Gain}{Average\ Loss}$. The RSI is typically calculated using a 14-day rolling window to determine the average gain and loss over that period.

• **MACD (Moving Average Convergence Divergence):** Used to capture changes in price trends and help identify buy or sell times.

$$MACD = EMA_{long} - EMA_{short} \tag{8}$$

$$Signalline = EMA(EMAD) \tag{9}$$

Where EMA_{short} is typically the 12-day Exponential Moving Average, EMA_{long} is the 26-day Exponential Moving Average, and EMA(MACD) is usually a 9-day EMA of the MACD line.

• **Bollinger Bands:** Identifies extreme market behaviour by the position of price relative to the range of volatility.

DOI: 10.54254/2754-1169/2024.26965

$$Upper\ Band = SMA_{20} + (2 \times Standard\ Deviation) \tag{10}$$

$$Lower\ Band = SMA_{20} - (2 \times Standard\ Deviation) \tag{11}$$

Where SMA_{20} is the 20-day Simple Moving Average.

• **Momentum:** Measures changes in price momentum and captures short-term fluctuations in the market.

$$Momentum = Price_{current} - Price_{n\ period\ ago} \tag{12}$$

• **Moving Average Cross:** Identifies trend reversal points by the crossover of short-term and long-term moving averages.

$$Cross = SMA_{short} - SMA_{long} \tag{13}$$

where SMA_{short} is a 5-day Simple Moving Average, and SMA_{long} is a 20-day Simple Moving Average.

4.1.3. Calculating market characteristics

• **Moving Average:** 20-day rolling average of VIX and SPY volatility, smoothing market fluctuations and capturing long-term trends.

• **Volatility:** Calculates the 20-day standard deviation of the VIX and its returns to further quantify market volatility.

• **Correlation:** Computes the correlation between SPY and VIX to understand the relationship and co-movements between these two market indicators.

• **Cumulative Returns:** This analyses the net returns of both the SPY and VIX in order to determine the effects of VIX tracking to SPY in the long run.

• **RSI and VIX Combination:** This is accomplished by including the closing price of VIX with RSI to obtain a combined, more responsive signal to market sentiment shift.

As complexity is included with these technical indicators such as relative strength index (RSI), moving average convergence divergence (MACD) and Bollinger Bands; its capability to tap into market volatility, change in trend, movements of assets and its relationship with economic indicant is enhanced. This enhanced input makes the model more aware of general market behaviour and its trends, enabling it to respond flexibly and move precisely, so gaining robustness against market fluctuations.

DOI: 10.54254/2754-1169/2024.26965

4.2. Portfolio position setting

4.2.1. Parameterisation according to the technical

Positioning a portfolio on daily basis long or short, is primarily influenced by technical trade tools such as RSI, MACD, and Bollinger Bands. Here's how these indicators are calculated and used:

Relative Strength Index (RSI): This momentum oscillator examines the velocity as well as the alteration of the price using gains and losses. They help in establishing when a certain asset is overbought or when it is oversold, thus the need for use. For instance, when Relative Strength Index (RSI) is at 7 or above than 70 it is considered over bought and gives a sell signal. As for the other part of the equation, when the RSI is below 30 for a certain asset it means they are oversold and this is the right time to go on the market and buy.

Moving Average Convergence Divergence (MACD): It is an indicator used to point out stage flipping and is found by subtracting the 26 EMA from the 12EMA with nine EMA as the low pass filter. When the MACD crosses above the signal line and the difference exceeds 0.01, it indicates a buying signal.

Bollinger band: Bollinger Bands are created using a 20-day Simple Moving Average (SMA) along with a measure of volatility, typically set at two standard deviations from the SMA. A signal to buy is generated when the asset price drops below the lower band, while a signal to sell arises when the price moves above the upper band. Adjusting the width of the Bollinger Bands can help filter out market noise or minor price swings, thereby providing better flexibility to changing market conditions.

Figure1: Bollinger Bands applied to SPY and F30 price movements. The chart displays the closing prices of SPY (in orange) and F30 (in purple), with their respective Bollinger Bands on both sides. A sell signal is initiated when the price line crosses the upper band, while a buy signal is generated only when the price breaches below the lower band. This clearly shows how Bollinger Bands can serve as a summary indicator, helping to provide clear market signals and an understanding of the risk level when the width is adjusted according to different market conditions.

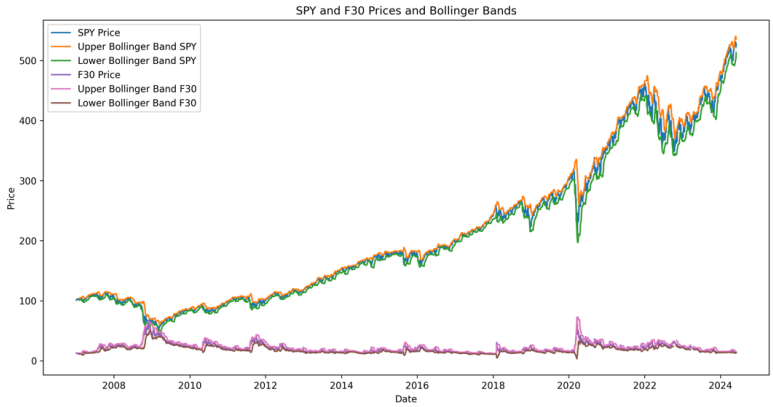


Figure 1. SPY and F30 prices and bollinger bands

4.2.2. Parameter optimization: grid search for technical indicators

• **RSI Thresholds:** Optimize settings to achieve the best levels for overbought and oversold thresholds (e.g., 20/80, 30/70, or even 40/60) that provide high probability buy and sell signals.

DOI: 10.54254/2754-1169/2024.26965

• **MACD Signal Difference:** The difference between the MACD line and the signal line is kept small (e.g., 0, 0.01, -0.01) to enhance precision in capturing market trends effectively. This allows the strategy to adapt better to a wide range of market trends.

• **Bollinger Bands Width:** The width of the Bollinger Bands can be adjusted to alter the sensitivity of signals under various market conditions. A narrower range might produce stronger signals in low-volatility markets, while a wider band can reduce false signals during periods of high volatility.

4.2.3. Signal combination and strategy optimization

After defining all the parameters, the signals generated by the RSI, MACD and Bollinger Bands are combined with different weights to form a comprehensive strategy to set up positions. To generate a valid trading position (long or short), all three technical indicators must provide consistent signals:

• **Long Position (Position = 1):** A long position will only be set up if all three indicators (RSI, MACD and Bollinger Bands) provide a buy signal.

• **Short Position (Position = -1):** A short position is opened only if all three indicators give a sell signal.

• **Invalid Position (Position = 0):** If the signals from the three indicators do not coincide (e.g. one indicator signals Buy and the other Sell), the position is considered invalid or neutral and the position is set to zero.

By applying this filtering mechanism, the strategy ensures that only the most robust and consistent signals lead to trading decisions, thus reducing the risk of false signals and improving the overall reliability of the strategy.

4.2.4. Cumulative return of the strategy

Figure2 below shows the cumulative return generated by the strategy after combining the signals from the RSI, MACD, and Bollinger Bands with consistent position setting. The steadily upward trend of the cumulative return line indicates that the strategy is effective over time. The approach of using only consistent signals from all three indicators helps to reduce the risk of false signals and enhance the robustness of the strategy. The result is a series of profitable trades that collectively produce a favourable cumulative return, demonstrating that the chosen parameters and the combination of signals effectively capture profitable opportunities in the market.

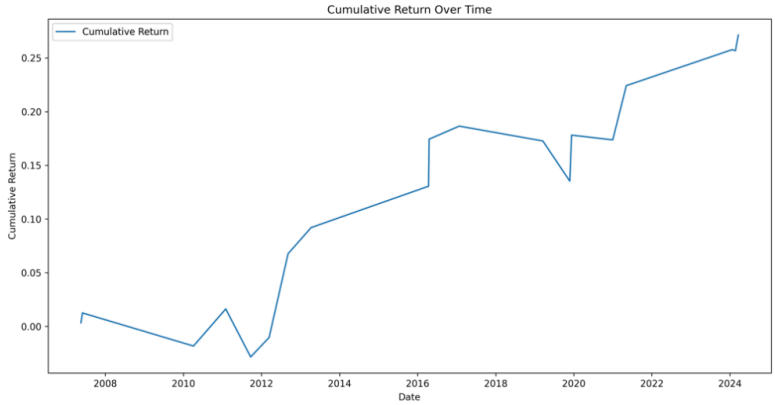


Figure 2. Cumulative return over time

DOI: 10.54254/2754-1169/2024.26965

4.3. Feature selection and model construction

4.3.1. Feature adding

To enhance the model's predictive capabilities, the dataset was expanded with a comprehensive set of features. These features include technical indicators, lagged variables, and macroeconomic factors that help capture various dimensions of market behavior. The features added to the dataset are as follows:

Technical Indicator adding:

• **RSI (Relative Strength Index):** RSI_SPY and RSI_f30 for SPY and F30, respectively. These indicators measure the market's overbought or oversold conditions.

• **MACD (Moving Average Convergence Divergence):** MACD_SPY and MACD_f30 capture the trend changes in SPY and F30 prices.

• **Bollinger Bands:** Upper and lower Bollinger Bands for SPY (BB_Upper_SPY, BB_Lower_SPY) and F30 (BB_Upper_f30, BB_Lower_f30) to identify extreme market behaviors.

• **Momentum Indicators:** Momentum_SPY and Momentum_f30 for SPY and F30, measuring the rate of price changes.

• **Moving Average Cross:** MA_Cross_SPY and MA_Cross_f30 for SPY and F30 to identify trend reversals.

Market Features adding

• **VIX-related Features:** MA_VIX_Close and Volatility_VIX_Close capture the VIX index's moving average and volatility, indicating market fear and uncertainty.

• **Correlation with VIX:** Corr_SPY_VIX_Close measures the correlation between SPY and VIX, providing insights into the relationship between market volatility and asset performance.

• **Cumulative Return:** Cumulative_Return_SPY is the cumulative return of SPY, reflecting its historical performance trend.

• **Combined Feature:** RSI_SPY_with_VIX An interaction feature that captures combined information between RSI and VIX to capture complex dependencies between these indicators.

Lagged Features adding

• **Lagged Prices for SPY and F30:** Lagged features have been added for both SPY and F30 with time lag of 1-3 days (SPY_Lag_1, SPY_Lag_2, SPY_Lag_3, f30w_Lag_1, f30w_Lag_2, f30w_Lag_3 .

4.3.2. Feature selection

Feature selection In this line, to figure out the most important features to predict F30 and SPY returns, we developed RFE combined with GBR. We split our data into 80-20 for training and testing such that the model development and evaluation framework was robust. The RFE technique is the elimination of the least influential features from the model at each step followed by retraining to reach a desired number of features. Refining iteratively means that only the most predictive features will be left, making the model more accurate and less complex. The following is a brief description of the RFE function used in our application:

$$L\left(y,\hat{y}\right)=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2 \tag{14}$$

DOI: 10.54254/2754-1169/2024.26965

Where y_i is the true value, $\hat{y}_i$ is the predicted value, n is the number of samples. With RFE, not only the number of features is systematically reduced, but at the same time focusing on those features that contribute the most to the revenue prediction.

Gradient Boosting Regressor, as a powerful integrated model, supports the feature selection process by minimizing the following objective function to optimize the model and ensure high predictive performance of the final model. The objective function of GBR consists of a loss function and a regularization term of the form:

$$\text{Objective} = \sum_{i=1}^{n} L\left(y_i, \hat{y}_i\right) + \sum_{j=1}^{T} \Omega\left(f_j\right) \tag{15}$$

Where the loss function $L\left(y_i, \hat{y}_i\right)$ measures the prediction error of the model, while the regularization term $\Omega\left(f_j\right)$ is used to control the complexity of the model to avoid overfitting. The regularization term is usually based on the complexity of the tree with the expression:

$$\Omega\left(f_j\right) = \gamma T + \frac{1}{2}\lambda \sum_{k=1}^{T} w_k^2 \tag{16}$$

Where T is the number of leaf nodes of the tree, w_k is the weight of the leaf node. γ, λ are hyperparameters that control the regularization strength. By optimizing the objective function, the GBR model is able to accurately assess the contribution of each feature to the target variable, thus supporting the feature selection process of RFE, and ultimately selecting the feature set that has the greatest impact on the predictive power of the model.

4.3.3. SPY & F30 feature importance result

Figure3 indicates that for SPY, the `RSI_SPY` is the most critical feature, underlining the importance of being able to identify overbought or oversold conditions that are pivotal to predict potential reversals. This ranking also shows that both `SPY` itself and `Volatility_VIX_Close` are very important features, indicating that both its own price changes and the broader market's volatility are crucial in forecasting SPY returns. The importance of the Bollinger Bands (`BB_Lower_SPY` and `BB_Upper_SPY`), `MACD_SPY`, and `Momentum_SPY` re.

The most influential feature for F30 is the asset's own price (f30w), followed by MACD_f30 and RSI_f30, which helps confirm the relevance of price, momentum, and trend indicators to predict F30's returns. The presence of Bollinger Bands (BB_Lower_f30 and BB_Upper_f30) and lagged SPY features (SPY_Lag_1 and SPY_Lag_2) implies that not only do F30's historical data matter but also that SPY recent performance would impact F30, reflecting similar market behaviours.

DOI: 10.54254/2754-1169/2024.26965

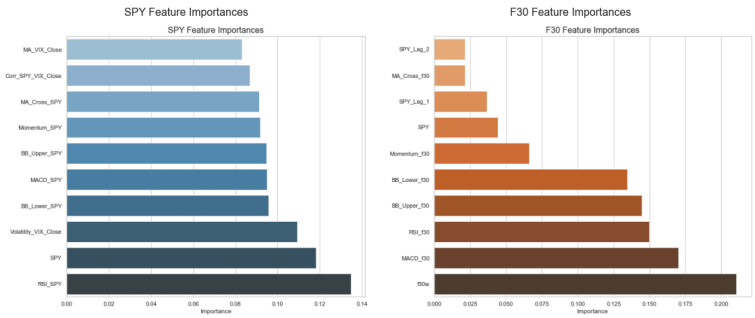


Figure 3. SPY & F30 feature importance

4.4. Hyperparameter optimization

We performed hyperparameter tuning and obtained the best-performing Gradient Boosting Regressor (GBR) models by using the RandomizedSearchCV function. This is essentially a randomized search over specified parameter grids so that we could effectively evaluate the effect of different hyperparameters in our model. In our search to identify the best combination of hyperparameters, we ran 50 iterations; each configuration was validated by using a 3-fold cross-validation technique based on our R^2 performance indicator.

Following the hyperparameter optimization phase, we proceeded to train the GBR models using the discovered optimal settings for predicting both SPY and F30. The hyperparameters that were fine-tuned during this stage encompassed:

• **n_estimators:** determines the number of weak learners (decision trees) in the

model. The number of boosting stages, ranging from 100 to 1000.The goal of the model is to minimize the following loss function:

$$L\left(\theta\right) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 + \lambda\sum_{j=1}^{m}\Omega\left(f_j\right) \tag{17}$$

Where y_i is the true value, $\hat{y}_i$ is the predicted value, n is the number of samples. $\Omega\left(f_j\right)$ is the regularization term. m is the number of trees, and λ is the regularization intensity.

• **max_depth:** controls the maximum depth of each decision tree, which varying between 3 and 10.

• **Learning_rate:** The step size shrinkage used to prevent overfitting, ranging from 0.01 to 0.2.The model's predictions are updated by the following formula:

$$\hat{y}_i^{(t+1)} = \hat{y}_i^{(t)} + \eta f_t\left(x_i\right) \tag{18}$$

Where □ is the learning rate, and $f_t\left(x_i\right)$ is the first t prediction of the tree in round t.

• **subsample:** The fraction of samples used for fitting the individual base learners, varying from 0.7 to 1.0.

DOI: 10.54254/2754-1169/2024.26965

4.5. Model performance evaluation

• Mean Square Error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{19}$$

• Average Absolute Error:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{20}$$

• Coefficient of determination (R^2):

$$R^2 = \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \tag{21}$$

R^2 Performance

As shown in the left panel of Figure4, the R^2 scores for both SPY and F30 are very high on the training set (blue bars), nearing 1.0, which indicates that the model fits the training data exceptionally well and captures most of the variance. On the test set (orange bars), the R^2 score for SPY stands at 0.64, which is a solid result given the complexities of financial markets. However, the R^2 score for F30 on the test set is 0.82. Such a high R^2 could suggest that the model is possibly overfitting, capturing patterns specific to the training data that may not generalize well to new data. Nonetheless, the results demonstrate that the model effectively captures significant trends, especially for SPY, while the higher R^2 for F30 warrants further investigation to confirm the model's generalizability.

MSE & MAE Performance

As depicted in the middle panel of Figure4, the MSE values for both SPY and F30 are very low on the training set, further confirming that the model fits the training data well. However, there is a notable increase in MSE values on the test set for both SPY and F30, especially for F30. This indicates that the model's predictions have a higher average error on the test data, suggesting a lack of generalization. The right panel of Figure4 shows that the MAE values for both SPY and F30 are low on the training set but increase slightly on the test set, particularly for F30.

DOI: 10.54254/2754-1169/2024.26965

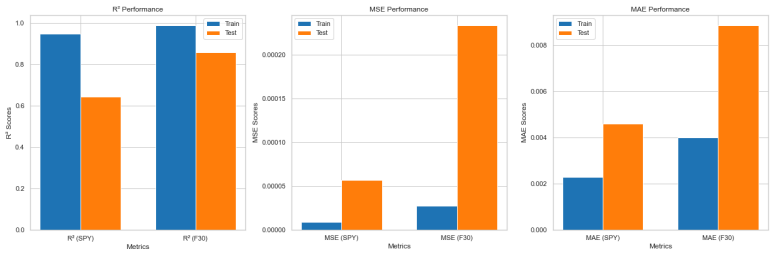


Figure 4. SPY & F30 performances

Figure5 illustrates the true versus predicted values for SPY and F30. The distribution of points in the scatterplot is more concentrated and closer to the red 45-degree diagonal line, indicating that the predicted values of the model have a high degree of agreement with the actual values. Although there is some deviation near some extreme values, the overall trend still reflects the predictive accuracy of the model.

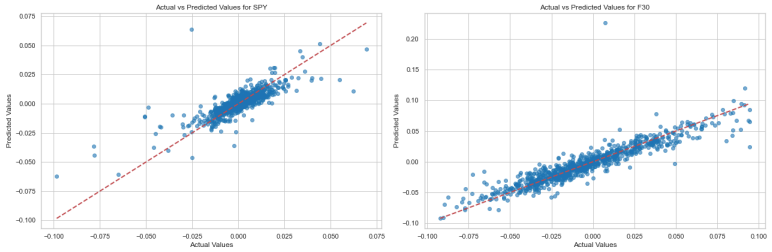


Figure 5. SPY & F30 true versus predicted values

In addition, as shown in Figure6, the residual analysis plot demonstrates the distribution of prediction errors for the SPY and F30 models. Most of the residuals are concentrated around the zero value and show a more symmetrical distribution, further proving the predictive ability of the models. The small number of extreme errors in the residual plots implies that the models are able to capture changes in returns more accurately in the vast majority of cases.

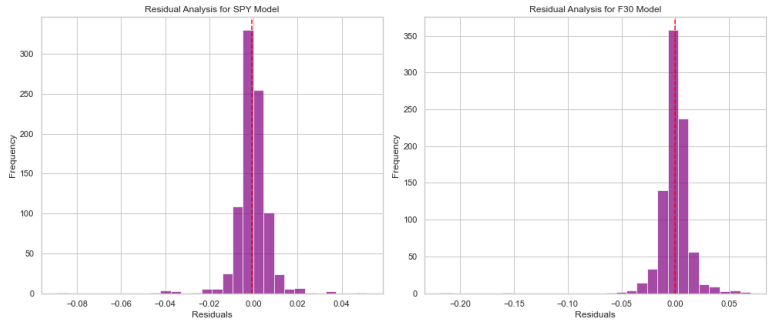


Figure 6. SPY & F30 residual analysis

Limitations and Issues

Overfitting: The significant drop in R^2 and the increase in MSE and MAE from the training to the test set suggest that the model is overfitting the training data. This overfitting limits the model's ability to generalize to unseen data, resulting in poor performance on the test set.

Imbalance Between Training and Testing Performance: The model's high performance on the training data compared to the testing data shows an imbalance. This discrepancy may weaken its

DOI: 10.54254/2754-1169/2024.26965

stability and generalization ability in real application scenarios where new data is usually encountered.

4.6. An optimal positioning of the individual instruments

4.6.1. Defining the portfolio return function

We make this adjustment in order to correctly calculate the daily return of the portfolio by taking into consideration the position (long or short) in each particular security. The purpose of this adjustment is to make sure that the strategy profits from long positions when prices increase and from short positions when they decrease.

Once a long position is held, the return of the portfolio is calculated as shown below:

$$Portfolio\ Return = \ Weight_{F30} \times F30\ Return + \ Weight_{SPY} \times SPY\ Return \tag{22}$$

When a short position is held, at which point the return of the portfolio is calculated by the formula:

$$Portfolio\ Return = \ -(Weight_{F30} \times F30\ Return + \ Weight_{SPY} \times SPY\ Return) \tag{23}$$

This approach defines how portfolio returns should be calculated under different market situations and forms the basis for weight optimization and risk management in the subsequent stages. On each trading day of the test dataset, the portfolio is constructed to ensure a dynamic adaptation in the predefined strategy with optimized overall balance between return and risk under any kind of market fluctuation.

4.6.2. Optimizing the weights of F30 and SPY assets

• **Define the objective function:** Optimization is targeting to achieve maximum Sharpe ratio. The Sharpe ratio is just the portfolio's return, less the risk-free rate, divided by the standard deviation of the portfolio. What it really represents is a metric to assess how effective the portfolio is in compensating investors for the risk they take.

• **Optimize daily weights on test set:** For each day in the test set, start the cycle from the second day, as the previous day's data is needed to calculate the covariance matrix. First, it is necessary to only consider the valid positions (long & short position), skipping invalid positions. The daily expected returns are then calculated based on the model's predicted SPY and F30 yields. The calculation of the covariance matrix depends on the number of days of data available: if the number of days is less than 2 (e.g., the second day), the unit matrix is used to avoid inaccuracies in the covariance matrix calculation; otherwise, the covariance matrix is calculated using the data of the predicted returns for all the days from the first day up to the current day with the formula:

$$\Sigma_t = \begin{bmatrix} \mathrm{Var}\left(R_{\mathrm{SPY},t}\right) & \mathrm{Cov}\left(R_{\mathrm{SPY},t}, R_{\mathrm{F30},t}\right) \\ \mathrm{Cov}\left(R_{\mathrm{F30},t}, R_{\mathrm{SPY},t}\right) & \mathrm{Var}\left(R_{\mathrm{F30},t}\right) \end{bmatrix} \tag{24}$$

DOI: 10.54254/2754-1169/2024.26965

5. Optimized portfolio construction based on adjusted objective function

5.1. Objective function design

To achieve a more stable and diversified portfolio allocation, the following objective function is established to optimize the weight distribution:

$$\underset{\mathrm{w}}{\operatorname{argmax}} \frac{\mathrm{w}^{\top}\mu}{1+\sqrt{\mathrm{w}^{\top}\Sigma \mathrm{w}}} - 0.01\sum_{i=1}^{n} w_i^2 - 0.1\sum_{i=1}^{n-1} |w_{i+1} - w_i| \tag{25}$$

$$\text{subject to } \sum_{i=1} w_i = 1 \tag{26}$$

Where:

$$\mu = [\mu_1, \mu_2, ..., \mu_n]$$

: The mean returns vector of the assets. Each

$\mu_i$is the expected return of asset i, reflecting its historical performance.

$\boldsymbol{w_i}$: The weight of the i-th asset, representing the proportion of funds allocated to the i-th asset in the portfolio.

$\boldsymbol{w}$: The weight vector representing the weights of all assets in the portfolio, denoted as $[w_1, w_2]$. The sum of the weights equals 1.

$\boldsymbol{\Sigma}$: The covariance matrix of asset returns, containing the variances and covariances between assets, which describe the interrelationships and overall risk of the portfolio. The covariance matrix is expressed as:

$$\Sigma\left(m\right) = \begin{bmatrix} \sigma^{(m)}_{\mathrm{SPY,SPY}} & \sigma^{(m)}_{\mathrm{SPY,f30w}} \\ \sigma^{(m)}_{\mathrm{SPY,f30w}} & \sigma^{(m)}_{\mathrm{f30w,f30w}} \end{bmatrix} \tag{27}$$

Where m denotes the data window size, the objective function employed in the optimization comprises three primary components, each fulfilling a distinct role in guiding the refinement of portfolio weights:

• **Return-to-Stability Ratio:** This ratio strives to maximize portfolio returns while accounting for risk, which accomplishes a good balance on expected returns against their standard deviation. What's more, the incorporation of "+1" ensures that low volatility does not disproportionately inflate values and makes the ratio present a stable property. Therefore, by intrinsically factoring risk into the return assessment, this component promotes portfolios that deliver consistent returns rather than merely pursuing elevated returns accompanied by heightened volatility.

• **Weight Regularization Penalty:** By imposing a penalty on the sum of squared weights, the whole function encourages a diversified asset allocation, which prevents the whole portfolio from over-concentration in any single asset. In this way, the unsystematic risk of the diversified portfolio will be mitigated, which also bolsters the portfolio's resilience and risk tolerance.

DOI: 10.54254/2754-1169/2024.26965

• **Smoothness Penalty for Weight Transitions:** This penalty targets discrepancies between adjacent weights and aims to prevent extreme variations, which will enhance the stability of the portfolio and further mitigate risk by promoting a balanced allocation.

The objective function prioritizes stable growth within a controlled risk framework and minimizes the overall sensitivity to volatility, as a result of which provides a more accurate representation of the portfolio's performance in reality.

Thus, adopting this method enables people to not only make optimization of portfolio returns and risks but also avoid potential risks linked to extreme weight configurations effectively, and then ultimately implement a more robust investment strategy.

5.2. Portfolio return calculation based on position dynamics

Before optimizing the portfolio construction, it is necessary to process the return of SPY and F30 since the position held by investors will affect the direction of their returns. When investors hold a long position, they expect the price of their assets to rise; When investors hold a short position, they expect the price of their assets to fall since they will earn more profit when repurchasing the assets at a lower price. So the return of the portfolio will be redetermined based on the daily valid position that invested held in the real market as follows:

• **Long Position:**

$$E(R_p) = w_{SPY} \bullet R_{SPY} + w_{f30w} \bullet R_{f30w} \tag{28}$$

• **Short Position:**

$$E(R_p) = -(w_{SPY} \bullet R_{SPY} + w_{f30w} \bullet R_{f30w}) \tag{29}$$

This method presents the calculation of portfolio returns across varying market conditions, which serves as the foundation for subsequent weight optimization and risk management. During each trading day within the data set, the returns of the portfolio will adapt dynamically within the predefined strategy no matter how volatile the market is.

5.3. Optimization of the portfolio construction

Initially, to ascertain the ideal weight allocation between f30w and SPY, the paper utilizes the scipy minimize function for optimization of the asset weights, whose goal is to maximize a tailored objective function (or, equivalently, minimize its negative counterpart), and iterates over different window sizes to compute daily portfolio returns and their corresponding risks.

At the outset, by imposing penalties on extreme weight values, the objective function achieves the prevention of excessive concentration in any single asset, which enables dynamic adjustments during the optimization process and fosters portfolio diversification and stability. Besides, the objective function integrates these penalties aiming to strike a balance between maximizing returns and controlling risks, which ensures the portfolio will be constructed with sufficient diversification.

In addition, the optimization process leverages the Sequential Least Squares Programming (SLSQP) and Truncated Newton Conjugate-Gradient (TNC) algorithms, whose proficiency in managing both linear and nonlinear constraints is performed well. In practice, each optimization

DOI: 10.54254/2754-1169/2024.26965

iteration starts with a set of diverse initial weight estimates, which will enhance the probability of converging to a global optimum, and ultimately these algorithms will iteratively refine these weight allocations to attain the optimal solution.

6. Results and discussion

6.1. Evaluation of three methods based on sharpe ratio

The Sharpe Ratio is such a widely recognized measure that it calculates the risk-adjusted return of an investment portfolio to determine its performance. To arrive at this, we first apply the portfolio weights on the optimized portfolio to the daily returns data. Such daily compounded returns summed up to estimate the annualized returns, which is represented by $E[R_p]_{annualized}$, using a particular formula.

$$E[R_p]_{\text{annualized}} = \left(\prod_{t=1}^{T} (1 + \text{ Portfolio Return }_t)\right)^{\frac{252}{T}} - 1 \tag{30}$$

Where T denotes the total number of trading days in the back testing period. The annualized volatility σ_p is then calculated by the following equation:

$$\sigma_p = \sigma_{daily} \times \sqrt{252} \tag{31}$$

Ultimately, Sharpe Ratio is calculated and optimized by the following formula:

$$Sharpe\ Ratio = \frac{E[R_p]_{\text{annualized}} - R_f}{\sigma_p} \tag{32}$$

6.2. Optimized portfolio construction based on risk minimization

The following table shows the sharp ratio results corresponding to different window sizes based on adjusted objective function method, detailed results can be found in Appendix A.

DOI: 10.54254/2754-1169/2024.26965

Table 4. Sharp ratio results based on valid position and assets weights

Window Size	β1	β2	β3	β4	β5	β6
5	1.2106	0.8198	1.0076	0.6262	1.0521	
10	1.3179	1.1432	1.0097	0.8266	1.1810	
15	1.3810	1.2070	1.1359	0.9145	1.4780	
20	1.3768	1.2281	1.1336	0.8510	1.5090	
25	1.3774	1.2333	1.1359	0.8771	1.4681	
30	1.3726	1.2397	1.1453	0.8721	1.4405	
45	1.3558	1.2444	1.2238	0.8010	1.4664	
60	1.3962	1.3750	1.2012	0.8225	1.4611	1.0779
75	1.3664	1.3715	1.2230	0.7916	1.3091	
90	1.3419	1.3454	1.2395	0.8075	1.3065	1.0258
120	1.3812	1.3772	1.2087	0.7605	1.2902	1.0807
150	1.4164	1.4103	1.2834	0.7524	1.2777	1.0799
180	1.4324	1.4283	1.2481	0.7907	1.3033	1.1159
210	1.4784	1.4370	1.3397	0.9820	1.2943	1.1453
240	1.4717	1.4398	1.3697	0.9469	1.2708	1.1354

This table shows that when window size=20, the optimal Sharp ratio=1.5090, which gives the best portfolio construction.

6.3. Optimized portfolio construction based on machine learning

The following table shows the optimized weights for selected F30 and SPY assets based on machine learning method, detailed results can be found in Appendix B.

Table 5. Selected F30 & SPY weights and sharp ratio

Date	F30 Weight	SPY Weight	Sharp Ratio
2022/11/15	0.6038	0.3961	1.2131
2022/11/16	0.6449	0.3550	1.2450
2022/11/17	0.6417	0.3582	1.2604
2022/11/18	0.6761	0.3238	1.2832
2022/11/21	0.7900	0.2099	1.2942
2022/11/22	0.7772	0.2227	1.3148
2022/11/23	0.7413	0.2586	1.3374
2022/11/25	0.6273	0.3726	1.2971
…	…	…	…
Average	0.6972	0.3027	1.3156

This table shows that the optimal average Sharp ratio=1.3156, which gives the best portfolio construction.

DOI: 10.54254/2754-1169/2024.26965

6.4. Optimized portfolio construction based on adjusted objective function

The following table shows the sharp ratio results corresponding to different window sizes based on adjusted objective function method, detailed results can be found in Appendix C.

Table 6. Sharp ratio results through various window sizes

Window Size	Sharp Ratio	Window Size	Sharp Ratio	Window Size	Sharp Ratio
5	0.7451	30	1.1571	120	1.1929
10	1.0258	45	1.1903	150	1.1614
15	1.1657	60	1.1852	180	1.1718
20	1.1173	75	1.1765	210	1.2100
25	1.1074	90	1.2045	240	1.1755

This table shows that when window size=210, the optimal Sharp ratio=1.210, which gives the best portfolio construction.

6.5. Discussion

The effectiveness of portfolio optimization strategies, whether aimed at minimizing risk, leveraging machine learning, or modifying the objective function, is influenced by market conditions.

Risk Minimization Approach:

I can say that this strategy makes use of a relatively narrow period of 20 days to compute the portfolio weights, making it easy to track short-term trends in the market. Its strength is that it can enter and exit trades quickly when responding to market fluctuations demonstrated well by a Sharpe ratio of 1.5090, which is the highest among the evaluated strategies. This makes Sharpe ratio higher so that this method is valuable especially for highly volatile environment. However, such constant changes may prove costly and add to the overall transaction and operational costs; therefore, the costs involved must be off stetted by the benefits to be derived.

Machine Learning Approach:

Here, features like Recursive Feature Elimination (RFE) and Gradient Boosted Regression (GBR), which capture essential market behavior patterns that are nonlinear. On the positive side, it has a great potential to drawn accurate predictions, though on the other side it has an issue with model's complexity and potential overfitting. The Sharpe ratio of 1.3156 obtained from the use of machine learning is slightly lower compared to the risk minimization strategy indicating that, although machine learning approach performs well under complex environments, its efficiency could be impacted in volatile and/or extreme condition in the market. They also identify some important points for future research, such as making these models more interpretable and stable.

Adjusted Objective Function Approach:

It is used in an attempt to filter out short-term market noise and to concentrate instead on longer-term trends, as a rolling 210-day window is used. This provided longer timeframe also limits the portfolio rebalancing thereby minimizing trading cost and increasing the portfolio stability. However, failure to react quickly to changes in the market to grab short-term gains which may be associated with higher risk lower returns hence the Sharpe ration of 1.210 is comparatively low than those of the other strategies. This analysis is best applied in extended investment horizons where the cost of trading is not very disruptive.

DOI: 10.54254/2754-1169/2024.26965

In other words, each strategy has its special benefits associated with various market conditions and investment goals and aims. This decision should therefore depend on the current market conditions, the investors willingness to take risks, and cost factors. The risk minimization strategy is especially effective for short-term investments in highly volatile markets.

The risk minimization approach is best for short-term, highly volatile markets.

The machine learning approach could be beneficial in more complex market conditions but needs more stability improvements.

The adjusted objective function approach is suitable for long-term investors focused on minimizing transaction costs.

Considering these factors helps investors choose the most appropriate hedging strategy for different market conditions to achieve better risk-adjusted returns.

7. Conclusion

This paper performs a thorough study of portfolio hedging strategies using F30 and SPY assets by integrating classic techniques for risk minimization with machine learning frameworks and optimization of the objective function. Using the developed methodology, maximization of the Sharpe ratio to find the best efficiency of the hedging strategy under different market conditions is carried out.

The paper shows three main achievements:

• We demonstrate that in markets characterized by high volatility, our strategy to minimize risk tremendously reduces the level of exposure to risk and increases the level of portfolio stability. This is achieved through various SPY-weighted strategies and the use of different rolling window sizes.

• The use of machine-learning techniques, especially Recursive Feature Elimination (RFE) combined with Gradient Boosted Regressors (GBR), significantly improves prediction accuracy and stability in portfolio optimization. It gives the capability to work with many nonlinear patterns associated with the movements in the various marketing areas and makes it easier to improve the flexibility of the various investment policies.

• By doing so. New methods in optimization and better objective functions that include smoothing penalties can be used to build a much more stable and cost-effective portfolio, actions that incorporate smoothing penalties and advanced optimization algorithms, a more stable and cost-efficient portfolio can be built. This makes the allocation of risk and return more acceptable as well as improving diversification and reducing the total risk in the portfolio when the markets turn sour.

Despite these promising results, certain limitations exist:

Risk minimization strategies: These rely on particular assumptions of market condition that may not apply in extraordinary situations.

Machine Learning Framework: It has great interpretability for the predictive performance, but it results in either the problems of feature engineering and model interpretability and the risk of overfitting.

Modified Objective Function Approach: Low frequency penalties augment portfolio von rick but distort the volatility of market prices. These modifications are only possible if right parameters are set, and it can easily manage itself to different parameters of market changes.

Future work should consist of enriching the database with more asset classes and years of market history in order to obtain a more stable strategy. Exploring deep reinforcement learning or other forms of learning models which, in some way, incorporate sophisticated learning mechanisms might be helpful for dynamic portfolio management research. Re-visiting other modes of regulation, or,

DOI: 10.54254/2754-1169/2024.26965

even where the multi-objective aims have been integrated into the portfolio, may also be of interest in an attempt to offer for new paradigms in respect of future and emerging markets.

References

[1] Markowitz, H. (1952). Portfolio selection. The Journal of Finance, 7(1), 77-91.
[2] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. Quantitative Finance, 1(2), 223.
[3] Tsay, R. S. (2005). Analysis of Financial Time Series. John Wiley & Sons.
[4] Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research, 259(2), 689-702.
[5] Tang, Y., Song, Z., Zhu, Y., Yuan, H., Hou, M., Ji, J., ... & Li, J. (2022). A survey on machine learning models for financial time series forecasting. Neurocomputing, 512, 363-380.
[6] Feng, F., He, X., Wang, X., Luo, C., Liu, Y., & Chua, T. S. (2019). Temporal relational ranking for stock prediction. ACM Transactions on Information Systems (TOIS), 37(2), 1-30.
[7] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. The Review of Financial Studies, 33(5), 2223-2273.
[8] Naik, N., & Mohan, B. R. (2019). Optimal feature selection of technical indicator and stock prediction using machine learning technique. In Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics: Second International Conference, ICETCE 2019, Jaipur, India, February 1–2, 2019, Revised Selected Papers 2 (pp. 261-268). Springer Singapore.
[9] Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267–288.
[10] Dai, Z., & Wen, F. (2018). A generalized approach to sparse and stable portfolio optimization problem. Journal of Industrial and Management Optimization, 14(4), 1651-1666. doi: 10.3934/jimo.2018025.
[11] Gunduz, H. (2021). An efficient stock market prediction model using hybrid feature reduction method based on variational autoencoders and recursive feature elimination. Financial Innovation, 7(1), 28.
[12] Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. Applied Stochastic Models in Business and Industry, 33(1), 3-12.