DEPARTMENT OF
Industrial Engineering &
Decision Analytics

Robo Advising for Retail Investors

By Kwok Ho Kai, Kyro (20762783)

Final Project Report for

IEDA 4500 - Engineering Foundations in FinTech Spring 2024

2024/05

Table of Content

List of Figures

List of Tables

1. Introduction

1.1 Overview

This report delves into the trading strategy development and implementation of a robo-advising system crafted for retail investors, highlighting a growing trend in the intersection of technology and finance. Robo-advisors have become a popular choice in financial advising due to their ability to offer automated, algorithm-driven services with minimal human intervention. Their Informed investment decisions are often with better outcomes than traditional advising methods.

This project aims to demonstrate these capabilities to deliver a system that simplifies investing and enhances decision-making for retail investors through daily trading signals and strategic guidance focused on SPY which break down barriers between complex financial strategies and retail investors and empower them with tools and insights that were traditionally available only to large financial institution.

1.2 Background and Motivation

Robo-Advisors vs Traditional Advisors

In today's world, nearly everyone is seeking ways to grow their money within the financial markets. However, the complexity of these markets often necessitates guidance from professional advisors to make informed investment decisions. The appearance of robo-advisors offers algorithm-driven services that operates without the typical human errors associated with traditional advising. Unlike human advisors, who may have personal biases and emotional influences that could affect their judgment, robo-advisors provide a consistent, unbiased approach to investment advice.

Traditional advising often requires face-to-face interactions or phone calls, limiting accessibility and convenience for many potential investors. In contrast, robo-advisors are available 24/7, making them highly accessible to a broader audience and bring them a lower minimum investment, which removes barriers that often exclude small-scale investors from receiving professional advice. Furthermore, the automated processes inherent in robo-advisors allow for rapid portfolio adjustments in response to market changes, a critical advantage over traditional advisors who may need significant time to analyze new information and react accordingly. Thus, the benefits of robo-advisors generally outweigh those of traditional advising methods, particularly in terms of accessibility, speed, and objectivity.

Characteristics of Retail Investors

Retail investors typically enter the market with limited levels of knowledge and understanding of financial concepts, and many possess limited awareness of how to manage risk effectively. They often lack the time or resources to actively learn about or manage their investments. Consequently, retail investors frequently seek straightforward, understandable investment solutions that can be managed with minimal personal involvement.

1.3 Objective

The insights derived from the background and motivation highlight the necessity for transformative solutions in financial advising. The project is structured around three primary objectives designed to cater specifically to the needs of retail investors.

Data-Driven Strategy Development

The first objective is to develop a data-driven strategy that leverages quantitative analysis and algorithmic processing to provide insightful advice to retail investors. This involves the use of backtesting and strategy optimization tools to ensure that the strategies are robust and can adapt to changing market conditions. In this way, retail investors could benefit from scientifically tested and optimized strategies and increase the likelihood of achieving higher returns.

Daily Trading Signals

The second objective is to offer simple and straightforward daily trading signals such as ‘Long', ‘Short’ or ‘None'. These signals are generated based on up-to-date market data and algorithmic forecasts. Therefore, retail investors could make informed decisions quickly without the need to continuously monitor market fluctuations.

Automated Diversification through ETFs

The third objective is to implement a feature within the robo-advisor that focuses on providing advice on ETFs, particularly SPY, which is designed to automatically diversify the investor's portfolio across different asset classes. This allows retail investors to gain exposure to a broad market through a single transaction to enhances risk management. It also eliminates the need for investors to conduct extensive research or make complex decisions about selecting individual securities.

2. Methodology

2.1 Strategy Development

The trading strategy adopted in this project centers around the application of mean reversion strategies, specifically leveraging the relationship between the VIX futures and the VIX index to generate actionable trading signals for SPY.

The VIX, often referred to as the "fear index," measures the expected volatility of S&P 500 index options. In figure 1, it is tested with around -0.7 correlation coefficient and inversely correlated with the market performance shown in figure 2. Typically, as the S&P 500 declines, the VIX rises, indicating heightened market fear or uncertainty. Although the VIX index itself is not tradable, VIX futures provide a means for trading based on forecasts of the VIX index's values.

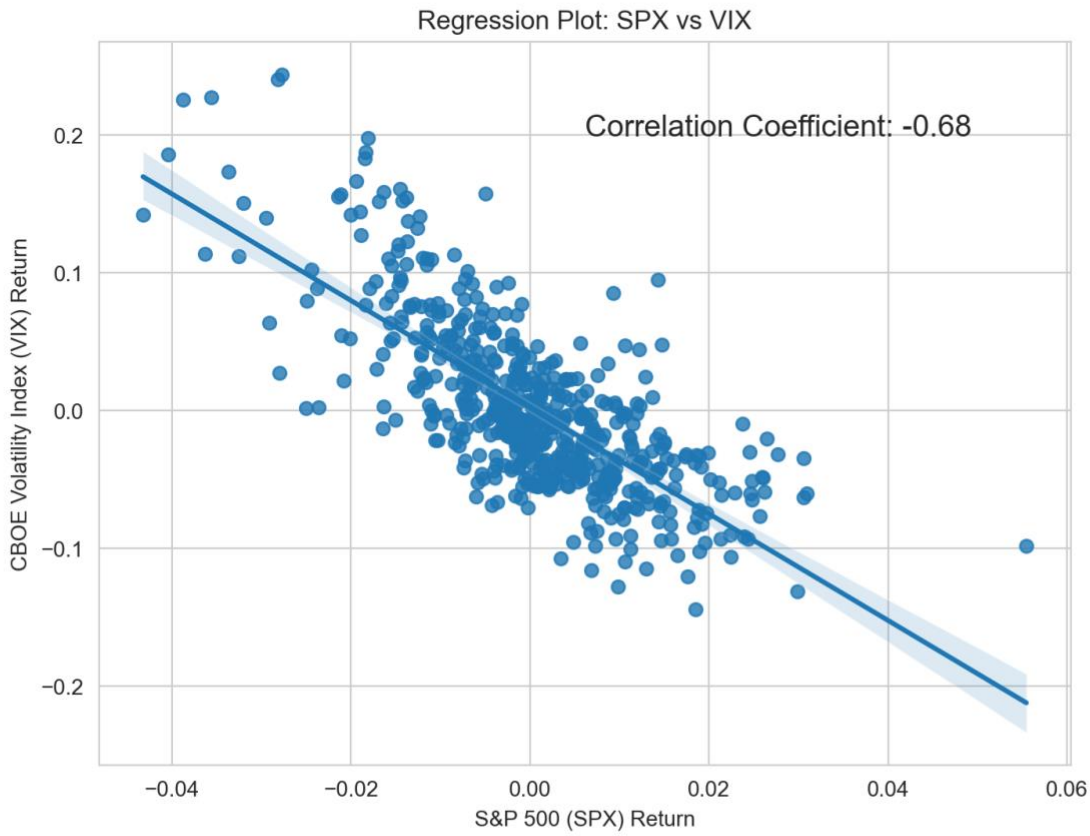

Figure 1: Correlation Plot between SPX and VIX

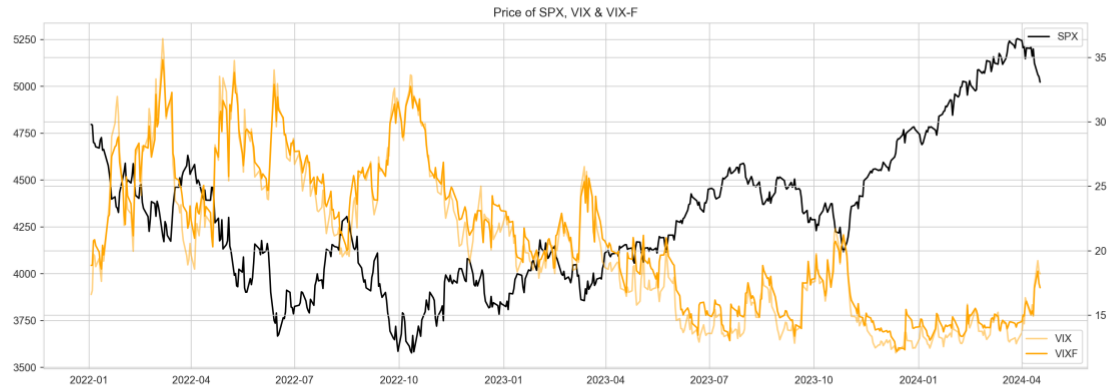

Figure 2: Time Series Plots of SPX, VIX and VIX-F

2.1.1 Mean Reversion

The strategy involves monitoring the spread between the VIX futures and the VIX index. This spread is tracked over time to identify extreme values that suggest a potential mean reversion scenario. The spread is calculated as follows:

$$\text{Spread} = \text{VIX 1 month Future} - \text{VIX Index}$$

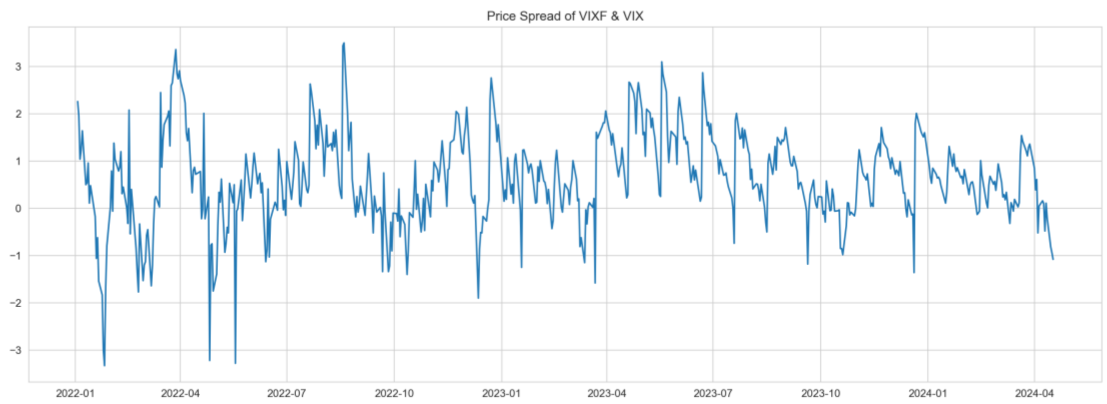

Figure 3:Spread of VIX-F and VIX

To assess whether the spread is stationary—meaning it has a tendency to revert to its mean, I carried out a cointegration test. The null hypothesis is set that the spread is non-stationary. However, a p-value close to zero provide confidence to reject the null hypothesis, confirming that the spread exhibit mean-reverting properties.

Pair	**P-Value**
SPX&VIX	0.965
SPX&VIXF	0.945
VIX&VIXF	**0.0001**

Table 1: P-Value of Cointegration Test of Different Pair

2.1.2 Rationale Behind

Institutional investors such as mutual funds, hedge funds, and proprietary trading firms, predominantly dominate the trading landscape of options and futures, including VIX futures (VIXF). These institutions wield significant influence over market movements due to their substantial capital and voluminous trading activities. This dominance provides a strategic advantage in predicting and understanding market volatility trends by tracking their trading behaviors.

The proposed strategy centers on monitoring the movements in VIX futures as a proxy for institutional sentiment toward anticipated market volatility. Typically, these institutional players engage in trading VIX futures for hedging against expected market upheavals or for speculative purposes. When these entities anticipate an increase in short-term market volatility, they tend to augment their positions in VIX futures. This increased demand propels the price of VIXF upwards, consequently widening the spread between VIX futures and the VIX index.

According to the historical mean reversion tendencies as confirmed through rigorous cointegration analysis, the observed widening spread indicates it is likely to revert to its mean by tending itself to rise to correct this temporal imbalance, which trigger SPX to drop with the negative correlation between VIX and SPX.

2.1.3 Strategies Algorithm

Indicator Used

In refining the strategy to capitalize on the mean reversion attributes of the spread, two primary indicators are introduced to explain the trend of spread:

1. Simple Moving Average (SMA): This indicator is utilized to filter out daily noise while retaining the crucial trends in the spread. It is is calculated by averaging the spread values over a specified number of past days. The formula for SMA is given by:

$$SMA_n = \sum_{i=0}^{n} Spread_i$$

- SMA_n: The simple moving average over n periods
- $Spread_i$: The value of the spread at period i
- n: The number of periods over which the average is calculated.

2. Rolling Z-Score: It is used to determine how unusually high or low the current spread is relative to historical spread data to capture the concept of relatively high and low. It is calculated for each point in time by standardizing the spread using the mean and standard deviation of the spread over a rolling window. The formula for the Rolling Z-Score is given by:

$$Rolling\ Z\_Score_i = \frac{Spread_i - SMA_{n,i}}{\sigma_{n,i}}$$

- $Rolling\ Z\_Score_i$: The Z-Score at period i
- $Spread_i$: The value of the spread at period i
- $SMA_{n,i}$: The average of the spread over the past n periods at period i
- $\sigma_{n,i}$: The standard deviation of the spread over the past n periods at period i

Strategies Logic

1. Mean Reversion Strategy: When the rolling Z-score exceeds a predefined upper threshold, it indicates that the spread is significantly higher than usual, suggesting an overestimation of future volatility compared to current market sentiment. Conversely, a rolling Z-score below the lower threshold indicates an underestimation. Based on these triggers, the strategy involves shorting the market when the spread is high (as indicated by a high rolling Z-score) and going long when the spread is low.

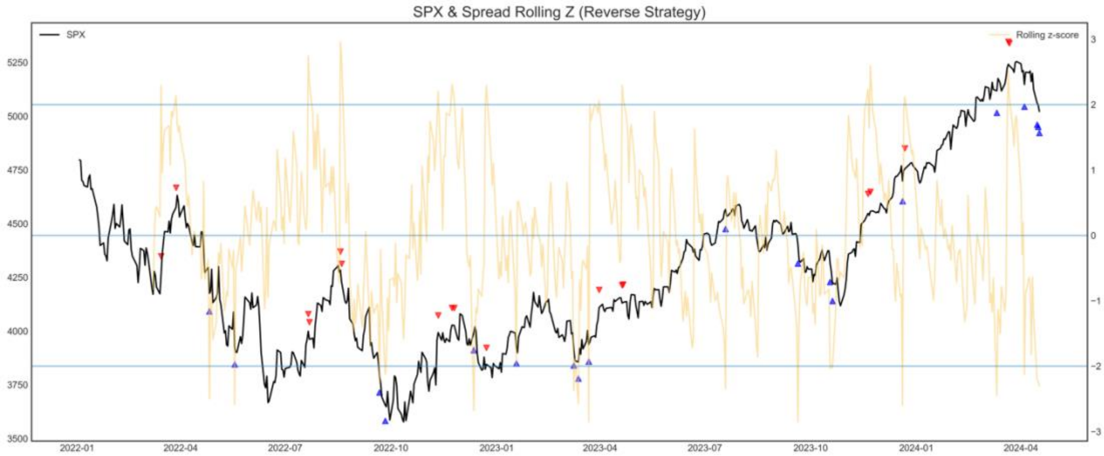

Figure 4: Example of Reverse Strategy

2. Local Extrema Strategy: This strategy focuses on identifying local maxima and minima of the rolling z-score. A local maximum suggests that a peak has been reached and the spread might start to decrease (mean revert), indicating a potential downtrend in market volatility. Similarly, a local minimum suggests an upcoming uptrend. The strategy involves shorting the market when a local maximum is confirmed and going long when a local minimum is detected. Besides, to confirm a local maximum or minimum, the strategy examines the rolling z-score over a specified number of periods to avoid the use of future data and ensures that the trading signal is based on fully available information.

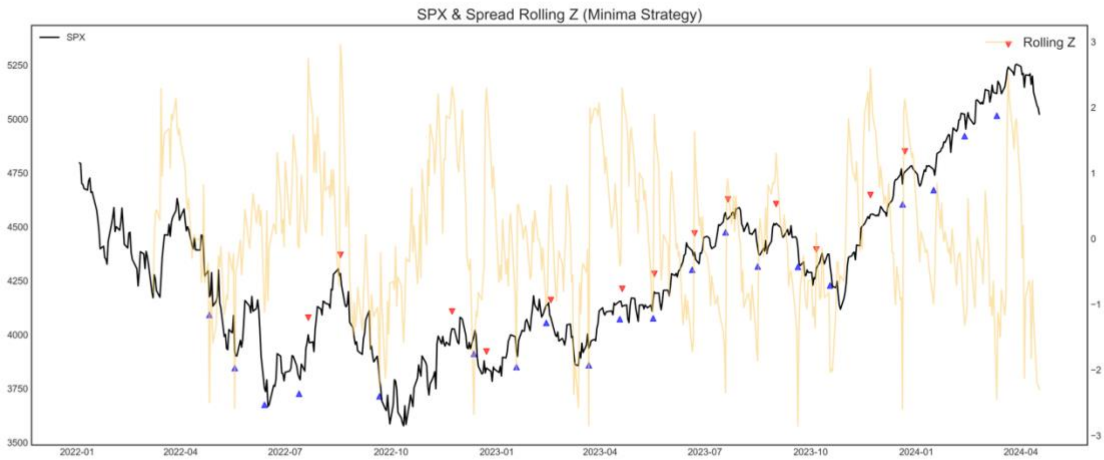

Figure 5: Example of Local Minima Strategy

2.2 Backtesting and Optimization

Sampling

Before backtesting, sampling is an important step in algorithm optimization to obtain statistically significant results and avoid overfitting. The data is divided into an in-sample training set for training and optimizing the hyperparameter, and an out-of-sample testing set for evaluating performance.

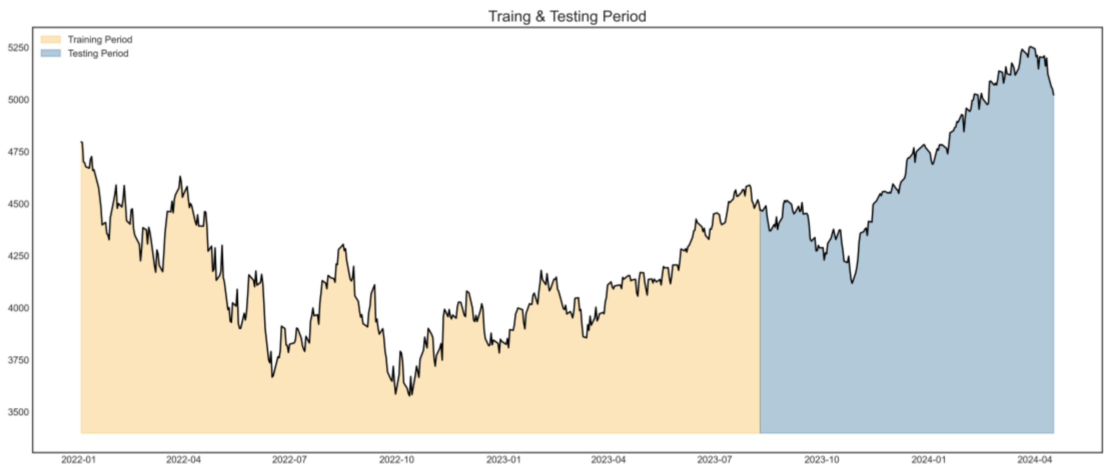

Figure 6: Splitting into Training and Testing Sample

Backtesting

Backtesting is a critical part of developing and optimizing a trading strategy. It could evaluate the performance of a trading model and strategy without risking real capital by using historical data to simulate how the strategy would have performed in the past. Additionally, backtesting can provide insight into the strengths and weaknesses of the strategy and provide confidence in the approach. The hyperparameter is firstly trained on the training set and then used to generate trading signals for the testing set and measure the profit and loss (PnL) for each trade.

Optimization

Hyperparameter optimization is a critical step in the development and refinement of algorithmic trading strategies. This process involves fine-tuning the pre-defined parameters that govern the strategy's behavior to improve performance and robustness.

It is employed by systematic grid search approach which explore a comprehensive grid of hyperparameter combinations, including the window size of SMA and Rolling Z-Score, holding dates of trading positions, upper threshold and lower threshold for reverse strategy, and the confirmation window for local minima strategy.

Performance Metrics

Performance Evaluation Metrics are important in backtesting as they help to quantify the performance of a trading strategy to measure the effectiveness of a trading strategy and compare it to other strategies. The report will mainly focus on sharpe ratio, maximum drawdown, and long-short ratio.

- Sharpe Ratio: This measures the risk-adjusted return of the strategy. It considers both the return and the volatility (risk) of the returns. A higher sharpe ratio indicates the strategy generated higher returns with lower exposure to risk.

- Long-Short Ratio: This measures the ratio of long trades to short trades. A ratio close to 1 indicates the strategy was neutral and not biased towards a bullish or bearish market. This indicates the robustness and versatility of the strategy across different market conditions.

3.3 Automation

The automation of the robo advising ensures consistency, efficiency, and real-time application of the strategy to capitalize on market opportunities. Here is the outlined methods used for automating data collection, signal generation and sending the signals through discord social media.

3.3.1 Data Crawling

The historical and daily updated data required for this strategy are sourced from the Chicago Board Options Exchange (CBOE) website, which is the official institution that tracks the VIX and VIX Futures Index. To automate the data collection process:

1. Expiration Date Mapping: Each day, the system checks for the VIX futures contract with the expiration date closest to the current date. This involves a mapping process where today's date is compared against the set of available futures expiration dates to identify the most relevant contract.

2. Data Retrieval: Once the appropriate expiration date is identified, the URL for the data is located within the HTML content of the CBOE webpage. The URL is then used to programmatically download the data using Python’s pandas library with the pd.read_csv(url) function, which loads the data directly into a DataFrame for further processing. More code details could be seen in GitHub repository in appendix.

```python
def run():
    expiry_date_dict = {
        '2024-05': '2024-05-22',
        '2024-06': '2024-06-18',
        '2024-07': '2024-07-17',
        '2024-08': '2024-08-21',
        '2024-09': '2024-09-18',
        '2024-10': '2024-10-16',
        '2024-11': '2024-11-20',
        '2024-12': '2024-12-18'
        }

    closest_expiry_date = str(get_closest_expiry_date(expiry_date_dict))
    url = 'https://cdn.cboe.com/data/us/futures/market_statistics/historical_data/VX/VX_'+closest_expiry_date+'.csv'
    vixf_raw = get_vixf_data(url)
    df = get_data(vixf_raw)
    return df
```

Figure 7: Implementation of Data Crawling

3.3.2 Data Transformation

With the data automatically crawled and loaded into the system, the next step involves cleaning and transforming this data in accordance with the strategy's logic and optimized hyperparameters:

1. Calculate the Spread: The first transformation step computes the spread between the VIX futures and the VIX index. This is the foundational data point from which the strategy generates its signals.

2. Smooth the Spread: Utilizing the optimized window sizes determined through hyperparameter optimization, the spread is smoothed using both the Simple Moving Average (SMA) and Rolling Z-Score. This smoothing helps to mitigate the impact of daily volatility and short-term fluctuations in the spread data.

3. Transform Indicators into Signals: Based on the strategy logic and the optimized hyperparameters, the transformed spread data and indicators are used to generate trading signals. For instance, if the Rolling Z-Score exceeds a certain threshold, a sell signal might be triggered according to the reverse strategy logic; similarly, identification of a local maximum or minimum would trigger buy or sell signals based on the local extrema strategy.

3.3.3 Integration with Discord Bot

To ensure that the trading signals generated by the automated trading system are readily accessible to retail investors, a Discord bot has been developed. This bot acts as a bridge between the strategic outputs from the Python environment and the end-users on Discord, a popular platform among investors for real-time communication and updates.

Bot Registration and Setup

The bot was registered on the Discord Developer Portal, where it was assigned a unique token. This token is crucial as it authenticates the bot's actions on the Discord server, ensuring secure and authorized interactions.

Discord API Utilization

The bot was created by leveraging the Discord API, which allows for programmable interactions with Discord servers and channels. Initially, I familiarized myself with the Discord API documentation to understand the capabilities and methods available for sending messages and managing channels.

Automation Script

In Python, a script was developed to link the bot with the trading strategy outputs. Using the discord.py library, the script uses the bot's token to log into Discord, navigate to the specific server and channel, and post messages.

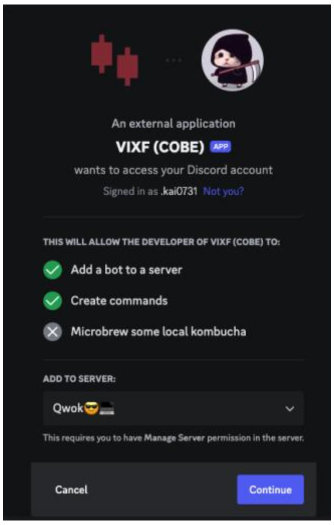

Figure 8: Screenshot of Adding Discord Bot into Discord Public Channel

3. Result and Analysis

3.1 Performance Evaluation

Reverse Strategy

Figure 9 heatmap shows how the combination of Z-score and SMA window sizes impacts the Sharpe Ratio. A deeper blue color, indicating a higher Sharpe Ratio, is more prevalent in certain regions of the heatmap. This result successfully eliminates the small SMA windows and Rolling Z-Score window which is in the red area to prevent the algorithm underfit. The regions of deep blue suggest that larger window sizes tend to stabilize the strategy performance by smoothing out noise and short-term fluctuations. This is critical in avoiding overfitting to market 'noise' rather than capturing underlying trends.

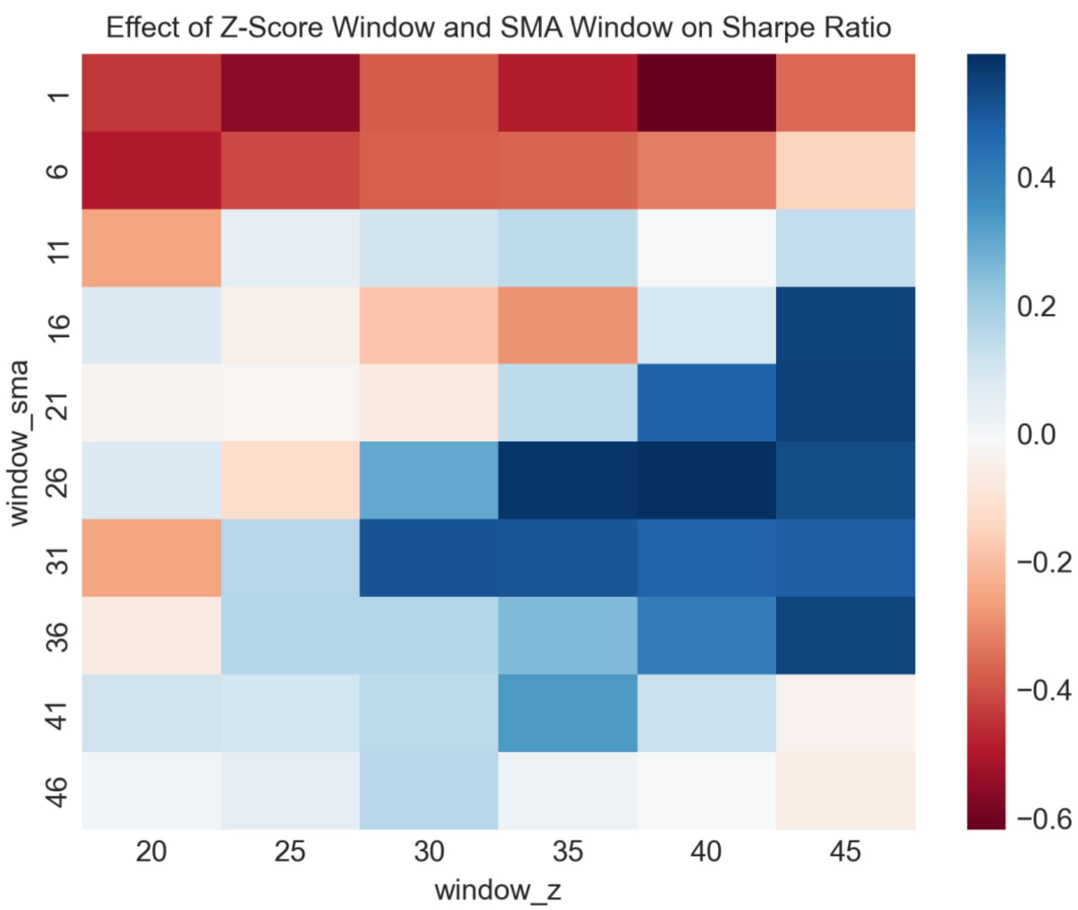

Figure 9: Heatmap of Different Window Size on Sharpe Ratio in Reverse Strategy

Figure 10 provides a visual summary of how different thresholds for entering long and short positions affect the strategy's performance. The strategy performs better with lower thresholds around -1.9 and upper thresholds around 2.7. This configuration allows the strategy to enter long positions more readily and short positions with more caution, which is advantageous in a market that exhibits more bullish than bearish periods. This approach aligns with the historical behavior of the US market, where bullish periods are more common.

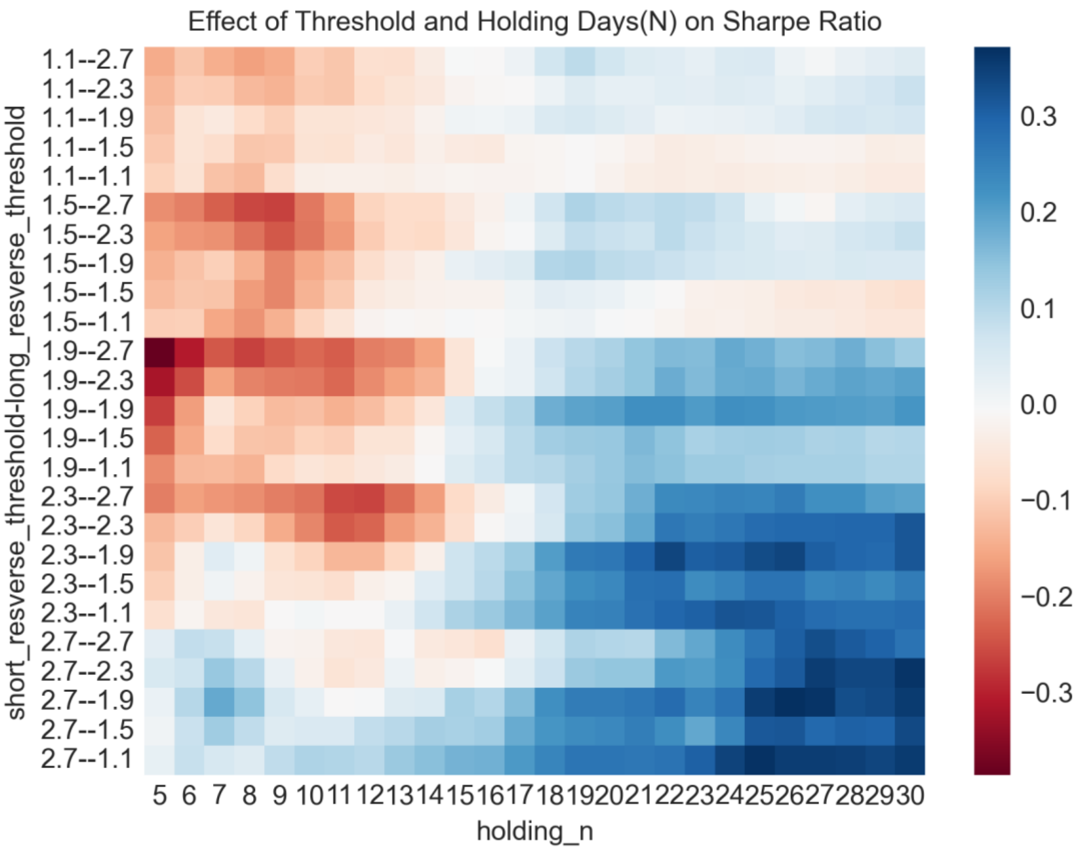

Figure 10: Heatmap of Different Threshold on Sharpe Ratio in Reverse Strategy

Figure 11 provides a comprehensive overview of the reverse strategy performance. The top graph illustrates the strategy position over time. The middle graph displays the equity curve for the strategy, showing the cumulative returns over time for strategy (orange line) and both long (blue line) and short (red line) positions for comparison.

The table in last row provide some key performance statistics.

As observed, the strategy performed well in in-sample, with 1.3 sharpe ratio. It can capture reversal timing effectively is highlighted by its consistent outperformance compared to the index. This suggests that the strategy's parameters are well-tuned to identify potential market turnarounds at the right moments, allowing it to capitalize on these opportunities before the broader market reacts. Besides, the long-short ratio is close to 1 which underlines the strategy's market-neutral stance. It is important to show the strategy robustness regardless of the general market direction.

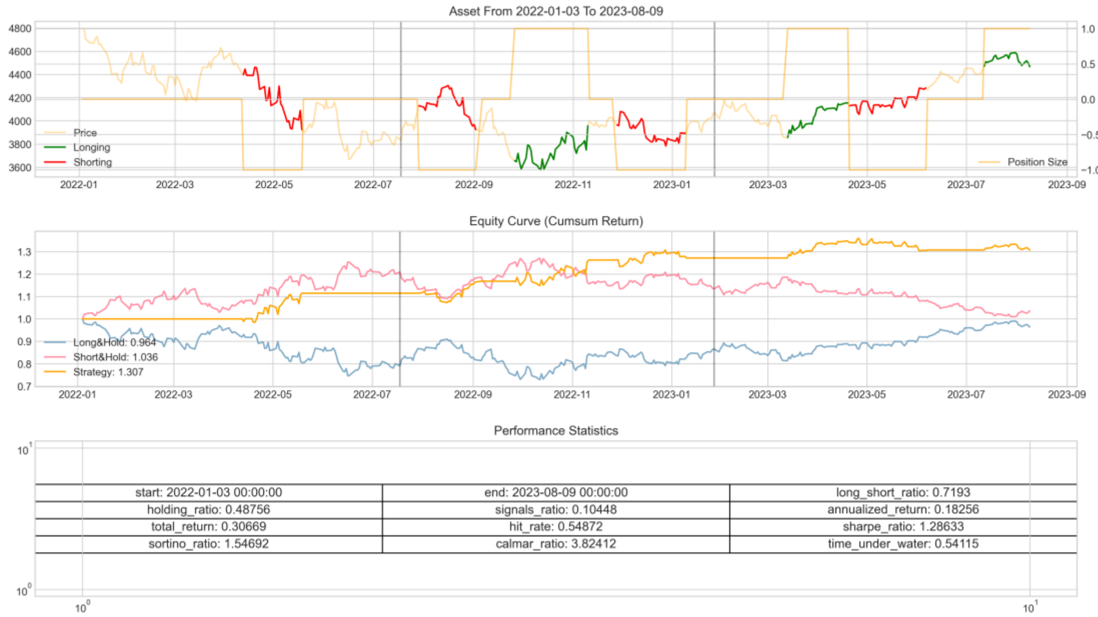

Figure 11: Backtesting Result of Reverse Strategy in In-Sample

Figure 12 shows the reverse strategy backtesting result in out-of-sample. The sharpe ratio is decrease to 0.7 and cannot beat the index. The difference in performance between the training and testing phases may indicate overfitting during the optimization stage. The strategy could be too closely tailored to the historical data, failing to generalize to new, unseen market conditions encountered during the testing phase. To address this issue, conducing additional out-of-sample testing and cross-validation to ensure that the strategy is robust and not overly fitted to specific historical data sets is recommended.

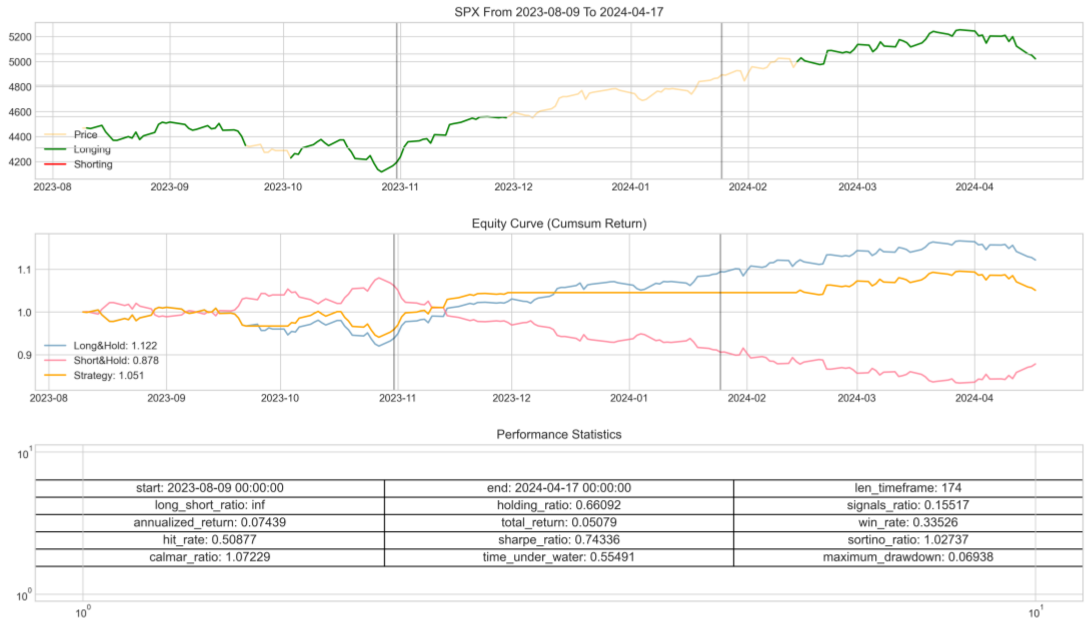

Figure 12: Backtesting Result of Reverse Strategy in Out-of-Sample

Figure 13 shows the local minima strategy backtesting result in out-of-sample. The strategy achieves a Sharpe Ratio of 2.4 indicates that the strategy not only provides high returns but does so with a controlled level of risk, which is ideal for risk-averse investors. The strategy also success in capturing the downturn in the previous month and adjusting positions accordingly is a key strength. It demonstrates the strategy's responsiveness to market signals and its capability to pivot quickly based on market conditions. Besides, the relatively balanced long-short ratio (1.25) suggests that the strategy maintains an effective equilibrium between its bullish and bearish positions, contributing to its overall stability and performance consistency.

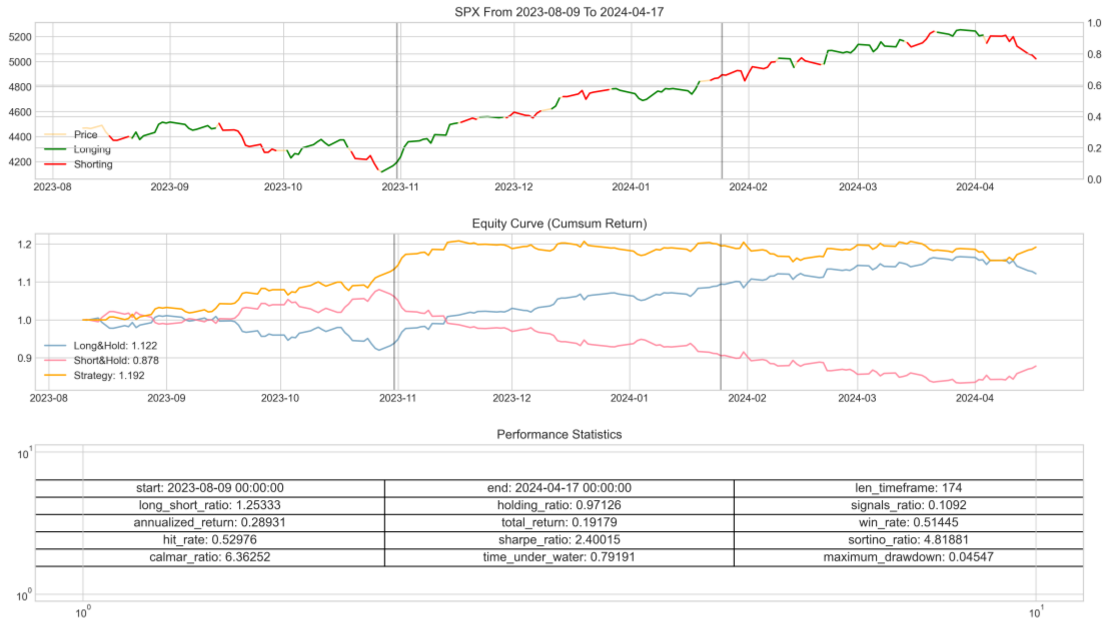

Figure 13: Backtesting Result of Local Minima Strategy in Out-of-Sample

In summary, the reverse strategy demonstrated a promising in-sample performance with a Sharpe Ratio of 1.3, effectively capturing market reversals and outperforming the index. However, the strategy's performance diminished in the out-of-sample testing, with a reduced Sharpe Ratio of 0.7 and failure to beat the index. This decline suggests potential overfitting to historical data, highlighting the necessity for further testing and parameter adjustments to enhance the strategy's robustness and adaptability to new market conditions.

Contrastingly, the local minima strategy showcased a robust out-of-sample performance with a well Sharpe Ratio of 2.4. It effectively capitalized on market downturns and exhibited a balanced approach with a long-short ratio of 1.25. These results underscore the strategy’s capability to deliver high returns while maintaining controlled risk levels, which is crucial for long-term sustainability and appeal to risk-averse investors.

3.2 Robo Advising Outcome

The integration of Robo Advising into a Discord channel represents a dynamic approach to delivering actionable investment strategies to retail investors. The relevant strategy information could be sent to a dedicated Discord channel before the market opens each day.

Content of Updates

- Market Data Overview: A brief summary of key market and indicators value.
- Strategy Positions: Clear indications of the advised positions (Long, Short, or None) based on the latest updated market data. This information guides investors on potential entry or exit points.

Figure 14: Screenshot of Robo Advising Information in the Discord Channel

By receiving daily, data-driven strategy recommendations, retail investors are better equipped to make decisions that align with current market conditions. This reduces the likelihood of emotional or ill-informed trading decisions. Besides, the advice provided by the Robo Advising bot includes not only position recommendations but also insights into potential market shifts. This can help investors manage risks better by adjusting their positions in anticipation of market movements.

Video Demonstration

For a practical demonstration of the robo advising outcome discussed, please refer to the following video: "[Demo] Robo Advising through Discord", available at: https://www.youtube.com/watch?v=z_t32dOK5uM

4. Discussion

4.1 Limitation and Improvement

Insufficient Backtesting

Only one sample for backtesting and hyperparameter optimization has been carried out. While the results are promising, there is a potential risk of overfitting.

It could enhance backtesting procedures by implementing additional backtesting using various sampling methods such as rolling-walk sampling and multiple randomized sampling to enhance model robustness and reduce the risk of overfitting.

Limited Underlying Asset

The strategy has been applied exclusively to the S&P 500 Index (SPX). This narrow focus may not sufficiently capture broader market dynamics.

It could expand strategy to additional asset classes by applying the trading strategy to other underlying assets such as commodities, interest rates, forex, and cryptocurrencies to diversify insights and improve market coverage.

Limited Information from Signals

Current signals are classified simply as long, short, or none. While this categorization ensures clarity, it may lack depth in terms of actionable information.

It could refine signal information by introducing an additional hyperparameter for bet size and optimize it to maximize the growth rate. This adjustment will provide investors with more detailed decision-making information, including insights on bet size and its associated risk-reward ratio.

Lack of Full Automation

While the script successfully automates data scraping, signal generation, and the robo-advising process, it still requires manual intervention to initiate the run. This manual step could lead to delays or inconsistencies in execution, potentially affecting the timeliness and effectiveness of the trading strategy.

It could implement full automation using cloud services to schedule and run the job automatically every day. This improvement will ensure that the trading system operates consistently without human intervention, increasing efficiency and reducing the risk of human error. By setting up automated workflows in the cloud, the system can also scale more effectively and adapt to increased data loads or computational demands over time.

4.2 Conclusion

The comprehensive analysis and implementation of 2 strategies and technologies in investment management demonstrate the profound impact and necessity of adapting to modern financial tools and methods. Here, I reflect on the key insights and successes outlined in the report:

Importance of Professional Institution Sentiment

The implied volatility and futures products provide useful insights into market behavior. These financial instruments, which are commonly utilized by professional institutions, provide insight into projected market volatility and future price fluctuations, which is critical for both short-term trading and long-term investing plans.

Using these data, investors can better understand the attitude and positioning of professional institutions, resulting in more informed and strategic investing decisions.

Insight of Feature Engineering on VIX Futures

Employing SMA and Rolling Z-Score in this context not only aids in trend identification but also enhances the robustness of the strategies by providing a clearer understanding of the underlying momentum and potential reversals in the market.

Success of Robo Advising

The introduction of Robo Advising brings in an age of change in financial services. Robo Advising simplifies access to complex investment methods that were previously reserved for professional traders and institutions by moving away from traditional, human-centric decision-making processes and toward a data-driven approach.

This technology-driven strategy assures that investment decisions are made using detailed data analysis and trends, rather than subjective human judgment. Robo Advising's achievement in giving quick, accurate, and practical financial advice demonstrates how technology may improve the efficiency and efficacy of market participation.

Reference

1. C. Krauss, "Statistical arbitrage pairs trading strategies: Review and outlook," Journal of Economic Surveys, first published May 09, 2016. [Online]. Available: https://doi.org/10.1111/joes.12153

2. E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst, "Pairs trading: Performance of a relative-value arbitrage rule," Review of Financial Studies, vol. 19, no. 3, pp. 797-827, 2006. [Online]. Available: https://academic.oup.com/rfs/article_abstract/19/3/797/1593449

3. F. D'Acunto, N. Prabhala, and A. G. Rossi, "Promises and Pitfalls of Robo-Advising," The Review of Financial Studies, vol. 32, no. 5, pp. 1983-2020, May 2019. [Online]. Available: https://doi.org/10.1093/rfs/hhz014

4. F. S. Lhabitant, Handbook of Hedge Funds. John Wiley & Sons, 2012. [Online]. Available: http://radoudoux.free.fr/Handbook_of_Hedge_Funds.pdf

5. J. Mencía and E. Sentana, "Valuation of VIX derivatives," Journal of Financial Economics, [Online]. Available: https://doi.org/10.1016/j.jfineco.2012.12.003

6. R. E. Whaley, "Understanding the VIX," Journal of Portfolio Management, vol. 35, no. 3, p. 98, Spring 2009. [Online]. Available: https://www.researchgate.net/profile/Robert-Whaley2/publication/277429711_Understanding_the_VIX/links/5f7f0538299bf1b53e1615fa/Understanding-the-VIX.pdf

Appendix

Source Code

The source code for the implementation discussed in this report, including data processing, backtesting, optimization, robo advising automation script and discord bot script, is available on GitHub. Details are as follows:

Repository Name: Robo-Advising-for-Retail-Investors

Access Link: https://github.com/KyroKwok2021/Robo-Advising-for-Retail-Investors

Additional Notes:

1. The TOKEN and CHANNEL_ID required for the Discord bot are not defined within the source code due to the privacy of this information. To run the Discord bot script located in the discord folder, you must provide your own TOKEN and CHANNEL_ID.
2. The backtesting implementation utilizes two self-written packages: visualization and signal_oscillator, which are part of a designed framework to facilitate signal generation, backtesting, and visualization in a structured and convenient manner. These packages are not publicly available as they have not been published due to personal and confidential reasons. This may result in errors when attempting to run the backtesting code from the repository.

Video Demonstration

For a visual demonstration of the robo advising outcome, refer to the following details:

Title: [Demo] Robo Advising through Discord

Uploaded by: Kwok Ho Kai

Uploaded date: 2024/05/28"

Available at: https://www.youtube.com/watch?v=z_t32dOK5uM

Disclaimer

This report summarizes a quantitative trading strategy backtesting for educational and informational purposes only. It is not investment advice or an endorsement to take any action based on its contents.

Trading involves substantial risk of loss which should be carefully considered before entering markets. You are solely responsible for your own trading and investment decisions. Past performance does not guarantee future results.

The information provided in this report is subject to change at any time without notice. Any views or opinions expressed may not reflect those of any entity. This report and its contents are intended for the original recipient only.