Stock Portfolio Optimization

A
Project Report

submitted for partial fulfillment

Bachelor of Technology degree
in
Computer Science and Engineering

By
Vivek Bhardwaj (1805210063)
Yashraj Singh Bhadauria (1805210066)
Umang Dubey (1805210060)

Under the supervision of
Prof. Maheshwari Tripathi
Dr. Promila Bahadur

Department of Computer Science and Engineering
Institute of Engineering and Technology, Lucknow
Dr. A.P.J. Abdul Kalam Technical University, Lucknow, Uttar Pradesh.
June 2022

Contents

Declaration

We hereby declare that this submission is our own work and that, to the best of our belief and knowledge, it contains no material previously published or written by another person or material which to a substantial error has been accepted for the award of any degree or diploma of university or other institute of higher learning, except where the acknowledgement has been made in the text. The project has not been submitted by us at any other institute for the requirement of any other degree.

Submitted by: Date: 25 May 2022

(1) Name: Vivek Bhardwaj
Roll No : 1805210063
Branch: Computer Science & Engineering
Signature:

(2) Name: Yashraj Singh Bhadauria
Roll No : 1805210066
Branch: Computer Science & Engineering
Signature:

(3) Name:Umang Dubey
Roll No : 1805210060
Branch: Computer Science & Engineering
Signature

Certificate

This is to certify that the project report entitled “Stock Portfolio Optimization” presented by Vivek Bhardwaj, Yashraj Singh Bhadauria and Umang Dubey in the partial fulfillment for the award of Bachelor of Technology in Computer Science and Engineering, is a record of work carried out by them under our supervision and guidance at the Department of Computer Science and Engineering at Institute of Engineering and Technology, Lucknow.

It is also certified that this project has not been submitted at any other Institute for the award of any other degrees to the best of my knowledge.

Prof. Maheshwari Tripathi
Department of Computer Science and Engineering
Institute of Engineering and Technology, Lucknow

Dr. Promila Bahadur
Department of Computer Science and Engineering
Institute of Engineering and Technology, Lucknow

Acknowledgement

We would like to express our special thanks of gratitude and appreciation to all those who gave us the opportunity to complete this project. First and foremost, we wish to express our sincere gratitude to our supervisors, Prof. Maheshwari Tripathi ma’am and Dr. Promila Bahadur ma’am, for their guidance, insightful suggestions, invaluable encouragement and practical advice that helped us tremendously at all stages in my project work and research related to it. During this we came to know about many new things which we are really thankful to them and learned many things regarding real world issues related to project management. Without their support and guidance, this project would not have been possible.

We would like to express our gratitude to the Head of Department - Dr. Divakar Singh Yadav Sir, Project Coordinator - Dr. Pavan Kumar Tiwari Sir , faculty members of CSE Department and friends who motivated us and helped us at each step in completing this project work within the limited time frame.

We would also like to express our sincere regards to all the Authors of all the references and other literary work referred to in this project work.

Yashraj Singh Bhadauria
Roll No : 1805210066
Signature :

Vivek Bhardwaj
Roll No : 1805210063
Signature :

Umang Dubey
Roll No : 1805210060
Signature :

Abstract

The Stock Portfolio Optimisation aimed at solving the problem of a retail investor who wants to invest his/her capital in a way to get expected returns with an acceptable amount of risk. It's always better to keep a diversification in the portfolio then relying on a single peg. The project aims at suggesting alternatives to the previous methods of diversification like 60:40 bond-equity split, equal distribution of funds etc.

We have done a comparative analysis of existing algorithms of Mean Variance Optimisation, Risk Parity, Hierarchical Risk Parity while capturing the shortcomings of all and the time when they shine. The idea is to avoid following a single algorithm for all users and to do personalisation in selecting the best way which matches with the risk profile of the user.

The main focus is on the newly introduced Hierarchical Risk Parity algorithm which distributes the risk among the assets and tries to establish better distribution based on their degree of correlation.

The project also aims to provide users with comparative analysis of the past trends in the portfolio value variation according to different weight allocation methods available which augments the understanding of the investor to make better decisions.

This report consists of experimental data with various combinations of stocks from multiple sectors analyzed on the performance in a way to see the adaptability of the portfolio to random shocks in the market. A graphical representation of the portfolio trends and distribution matrix is provided to the user for choice.

List of Figures

Chapter 1
Introduction

The practice of investment management has been transformed in recent years and in computational investment the most common financial issue is portfolio creation or portfolio optimization. Investment managers must develop portfolios on a daily basis that integrate their risk and return opinions and forecasts. His momentous knowledge was to understand that different levels of risk correspond to different optimal portfolios in terms of risk-adjusted returns., thus the idea of “efficient frontier”. One suggestion was that it is once in a while ideal to designate all resources to the ventures with most noteworthy anticipated returns. Instep, we ought to take under consideration the relationship over elective speculations in order to construct a diversified portfolio.

CLA is a quadratic optimization algorithm intended specifically for portfolio optimization problems with inequality constraints.
Bailey and López de Prado provide a description and open-source implementation of this approach. Surprisingly, most budgetary experts appear to be unaware of CLA, relying instead on general-purpose quadratic programming strategies that do not guarantee the correct arrangement or a halting time.

Despite Markowitz's brilliant theory, CLA solutions are rather dependable due to a number of practical issues. Small variations in anticipated returns cause CLA to generate quite diverse portfolios, which is a big concern. Even with a tiny input, the deviation is rather substantial. Many authors have chosen to ignore them entirely and concentrate on the covariance matrix. Risk-based asset allocation schemes, such as "risk parity," have resulted as a result of this. However, lowering the return estimates improves but does not eliminate the instability issues. Quadratic programming approaches necessitate the "inversion of a positive-definite covariance matrix (all eigenvalues must be positive)." When the covariance matrix is numerically ill-conditioned, i.e. has a high condition number, this inversion is prone to big mistakes.

The Black-literman model begins from an impartial position utilizing advanced portfolio hypothesis (MPT), and after that takes extra input from investors' sees to decide how the extreme resource allotment ought to veer off from the starting portfolio weights. It at that point experiences a handle of mean-variance optimization (MVO) to maximize anticipated return given one's objective hazard tolerance. The MPT model is said to be limited because it only uses historical market data and then assumes the same returns in the future.
The Black-literman approach allows the investor to apply their own viewpoints before optimizing the asset allocation recommendation.

Optimal capital allotment across resources is seemingly one of the most broadly examined points in quantitative money. Markowitz looked to probability and statistics to further his insights; Markowitz devised a method that allows an investor to trade off risk tolerance and reward expectations analytically, resulting in the optimum portfolio that optimizes return while reducing risk through diversification. Due to its intuitive appeal and theoretical feature as the pareto-optimal in-sample allocation (Kolm, Tütüncü, and Fabozzi 2014), mean-variance optimization (MVO) is the foundation of most applied portfolio optimization algorithms.

Despite its popularity, there are nevertheless a number of criticisms of MVO that are worth highlighting :

A. GIGO (garbage in,garbage out)
B. Allocations to specific asset classes
C. Risk diversification.

Since the global financial crisis in 2008 risk management has become more important, risk parity portfolio design tries to allocate the stocks on the basis of risk associated with them instead of considering them as a single portfolio entity.

Sensitivity to measurement error, which leads to poor out-of-sample performance, and covariance matrix instability are also key practical concerns. Due to MVO's sample dependence, solely risk-based allocation methods have emerged, which ignore error-prone return estimates. In large N portfolios, risk parity and minimum variance approaches exemplify this strategy, however they are often prone to the instability-critique. CLA has a number of disadvantages (inversion of covariance matrix- the inverse of covariance matrix can change significantly for small changes in portfolio, dependency on the estimation of stock-returns , consider correlations between all the assets and leads to very large dependency graph and not all assets are related to each other).

All of these drawbacks render CLA and other similar allocation algorithms inappropriate for real applications, which is where Hierarchical Risk Parity (HRP) comes in, since it attempts to address and improve on the aforementioned issues. There are three major phases to Hierarchical Risk Parity: -

A. Hierarchical Tree Clustering
B. Matrix Seriation
C. Recursive Bisection

The main idea of HRP is to allocate weights to a portfolio of securities based on

1) the clusters formed by securities (determined on how each security correlates to the portfolio)
2) the volatility of each cluster (more volatile clusters receive lesser weighting, and vice versa)

Hierarchical clustering is used to place our assets into clusters suggested by the data and not by previously defined metrics. This ensures that the assets in a specific cluster maintain similarity. The objective of this step is to build a hierarchical tree in which our assets are all clustered on different levels

A major source of quadratic optimizers' instability: A complete graph with 1/2N(N-1) edges is coupled with a matrix of size N. Because there are so many edges linking the graph's nodes, weights can rebalance with perfect freedom. Because there is no hierarchical structure, modest estimation errors will result in completely different answers. HRP uses a tree structure instead of a covariance structure to achieve three goals: It fully leverages the information included in the covariance matrix, b) weights' stability is recovered, and c) the solution is intuitive by construction, unlike standard risk parity approaches. In deterministic logarithmic time, the method converges.

HRP is robust, visible, and adaptive, allowing the client to present constraints or alter the tree structure without jeopardizing the algorithm's appearance. These characteristics are derived from the fact that HRP does not require invertibility of covariance. HRP can undoubtedly compute a portfolio using an ill-degenerated or even a single covariance framework, which is a remarkable feat for quadratic optimizers. "Monte Carlo tests appear to suggest that HRP has a lower out-of-sample change than CLA or traditional chance equality techniques," says the study.

A single optimizer/approach cannot be utilized for every set of assets, every set is unique in itself and markets are complex , so investors need to cautiously do the stock allocation in order to survive in the market.It is naive to think that one method is best for everything until the end of time. Monte Carlo simulation helps in quickly comparing a variety of optimization methods to find which is most robust in your particular case.

After calculating the expected returns and covariance matrix the data is fed to the simulator which inturn calculates the optimized portfolio using a variety of algorithms. Large number of simulated inputs for this purpose and optimized portfolios are created for all of them.

These optimized portfolios on simulated inputs are error estimated with optimized portfolios using original inputs and various insights are acquired using the same.The insights are presented in a graphical format.

Chapter 2
Literature review

2.1 Diversification of portfolio

2.1.1 Introduction

Diversification is a portfolio allocation strategy that aims to minimize distinguishing risk by holding assets that are not ideally positively correlated. Correlation is simply the relationship that two variables share, and it is measured using the correlation coefficient, which lies between $-1 \leq \rho \leq 1$.

For a better diversified portfolio the key point is to hold assets that are not are not ideally positively correlated.

i) When a portfolio is diversified, the risk associated with it becomes very low.

ii) Systematic chance alludes to the chance that's common to the whole showcase, not at all like a peculiar hazard, which is particular to each resource. Enhancement cannot lower efficient chances since all resources carry this risk.

2.2 Markowitz's Modern Portfolio theory[5][8]

Harry Markowtiz gave the idea of Portfolio Selection based on diversification in his Paper "Portfolio Selection". Investors generally consider expected return a desirable thing while the variance is undesirable. Generally high return portfolios are associated with higher risks and low risks are associated with lower return portfolios but Markowitz suggested a way that can help to achieve an optimal stage of diversification.

2.2.1 Acceptable risk:

Investors prefer a less risky portfolio compared to a riskier one. An investor can analyze the risk associated with a portfolio based on statistical measures like variance and standard deviation.

2.2.2 Efficient frontier:
On a coordinate plane, the efficient frontier rates portfolios. The risk is shown on the x-axis, while the return is shown on the y-axis—annualized standard deviation is used to evaluate risk, while compound annual growth rate (CAGR) is used to assess return.

The investor would choose securities that are at the efficient frontier's right end. Securities near the right end of the efficient frontier are predicted to have a high degree of risk with a high potential return, making them suited for risk-averse investors.

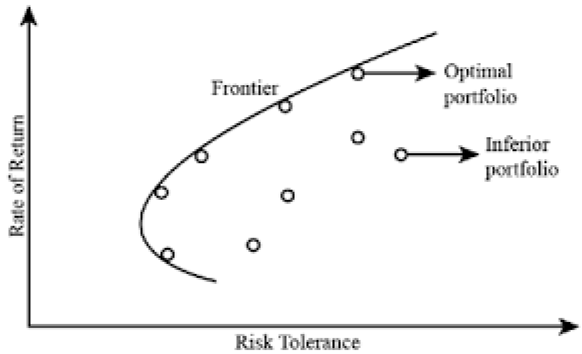


Fig 2.1:- Efficient frontier with Risk tolerance to risk - return

2.2.3 Sharpe Ratio (SR):
(*Expected return per unit of risk*)
It compares the performance of a portfolio with risk free returns. Any value greater than 1 is considered to be good by the investors.

$$S_a = \frac{E[R_a - R_b]}{\sigma_a}$$

2.3 Risk Parity[4][6][7]

2.3.1 Post Modern Portfolio Theory

The typical application of Modern Portfolio Theory (MPT) combines asset classes based on their projected returns, risks, and correlations, and then determines the best managers in each asset class once the asset allocation mix is defined. PMPT, on the other hand, is distinct in three ways: first, alpha and beta returns are separated; second, the sizes of alpha and beta are changed to more acceptable values; and third, considerably more diversified portfolios of each are produced. As a result, a PMPT portfolio will not only have returns and risks that are more tailored to the investor's goals, but it will also be far more diversified than a standard portfolio.

2.3.2 Weighting based on individual risks.

Advocates of the Risk Parity approach suggest that equally weighting asset classes by their risk (volatility) contribution to the portfolio would be a more efficient way to asset allocation. This method effectively provides the same volatility risk budget to each asset class; in other words, under the Risk Parity weighting scheme, each asset class contributes roughly the same predicted fluctuation in the portfolio's dollar value. Under the Markowitz framework, Risk Parity weighting could be seen as ideal if all asset classes have about the same Sharpe Ratios and correlations.

Investors do not need to create expected return assumptions to form portfolios, which is a major advantage of Risk Parity weighing over mean–variance optimization. Only asset class covariances must be specified, which can usually be predicted more precisely than expected returns based on historical data (Merton, 1980). Certainly, covariance estimates can influence portfolio allocation; nevertheless, it's unclear whether low-quality covariance estimates would skew portfolio returns downward.

When asset allocations are modified to the same risk level, the portfolio can achieve a better Sharpe ratio and be more robust to market downturns, according to the risk parity strategy.
The risk parity portfolio aims to confine each asset (or asset class, such as bonds, equities, real estate, etc.) to contribute equally to the portfolio total volatility.
Add formulas for relative risks

2.4 Hierarchical Risk Parity[1],[2],[3]

The risk-based portfolio optimization method Hierarchical Risk Parity (HRP) has been proven to construct diversified portfolios with robust out-of-sample features without the use of a positive-definite return covariance matrix.

2.4.1 Improvements from the existing algorithms.

Hierarchical risk parity is a random stock market shock by removing the rigorous analytical approach for calculating weights and instead relying on a more approximate machine learning-based approach (hierarchical tree clustering). On the other hand, it produces stable weights. In addition, older algorithms such as CLA perform inversion of the covariance matrix. This is a very unstable task and prone to have a large impact on performance with small changes in the risk model.. By completely removing the dependence on the inversion of the covariance matrix, the hierarchical risk parity algorithm is fast, robust, and flexible.

In fact, HRP can compute a portfolio based on a singular covariance matrix, an impossible feat for quadratic optimizers. The algorithm operates in three stages: Tree clustering, quasi-diagonalization and recursive bisection.

2.4.2 Three steps of Algorithm.

Tree Clustering stage is characterized by breaking the portfolio into various hierarchical clusters.Here we calculate the tree clusters based on a matrix mapping time series data with the number of stocks in the portfolio. The aim is to create a correlation distance matrix using Euclidean Distance.

The Quasi Diagonalisation stage refers to the matrix serialization method. It arranges the stock covariance matrix in a way that higher covariance is present at the diagonal and lower covariance at off diagonal elements.

Recursive Bisection stage is characterized by assigning the weights to the large cluster in a top down manner by recursively distributing the weights between the children.

2.4.3 Conclusion:

We get to know that relying on an approximate machine learning based approach (hierarchical tree-clustering) instead of fully analytical approach Hierarchical Risk Parity produces weights which are stable to random shocks in the stock-market.

2.5 Monte Carlo Simulation[2],[9]

Since portfolio optimization decisions via algorithms are based on calculated inputs which might be unstable and might even underperform compared to a naive equal allocation between assets a few times. Many times the blind allocation of assets using any algorithm without a deep analysis can lead to losses , just for those cases Monte Carlo Simulation is a tool that comes to rescue.

2.5.1 Multiple Simulation to predict better

Monte Carlo Simulation is used to predict the probability of various outcomes of an event. It involves simulating input values using random seed to achieve an average of multiple results being generated.

2.5.2 Five Step process

MCOS consists of five steps, as outlined below. The input variables are the array of expected outcomes (μ), and the covariance matrix of the expected outcomes (V). These variables may incorporate priors, following the Black-Litterman method or similar.

2.5.3 Provides Comparative Analysis

Estimation of error between portfolio created using original inputs and simulated inputs over various algorithms using this approach presents a holistic comparison between various optimizers for a particular set of assets on which optimization is to be performed.

Chapter 3
Methodology

When a user visits our web app we take information from the users: the amount they want to invest and a list of assets (in the form tickers array). After taking user information we fetch information of assets selected by users from historical Asset data and give this and user info to Analysis Stage.

In the Analysis Phase, we create a Risk Model(covariance Matrix based on the data and analyze forecasted returns). The next Stage is Optimizers where we use input from the Analysis phase and run all three optimization algorithms on the data respectively. The output of the optimizer phase will be three optimized portfolios according to the respective algorithm.

Another Phase is Simulation and that input is coming from the optimizers phase. In the simulation phase Simulation is applied to input data over 100's of time, and out of this phase helps to predict which portfolio we should choose according to the risk-to-return user wants.

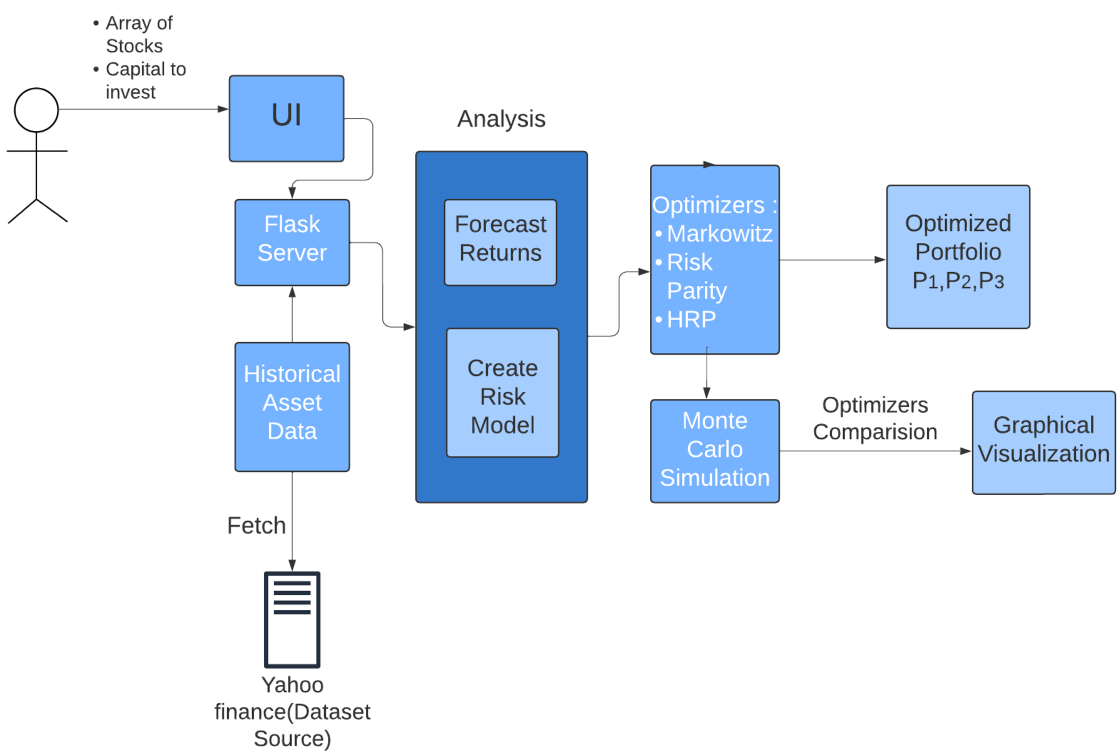


Fig 3.1 Flow Diagram of Web App

3.0 Project Characteristics

Before beginning with the detailed explanation of methodology used for creating the project here are some significant characteristics which provides insights into what purpose the project serves and how it can be used :

1. This project allows users to optimize their stock portfolios using 3 established portfolio optimization techniques.
2. Project's UI takes in an array of stocks and the capital to be invested as input and provides users with an allocation mapping i.e. "how much amount to invest in which stock". The allocation mapping is provided for each optimization approach.
3. It provides users with a pie chart representation of allocation for each optimization technique which provides a pictorial understanding of how the portfolio is supposed to look with respect to each optimizer.
4. It also performs monte carlo simulations over input data which helps in error estimation of each optimizer that helps in estimating. how much deviation is there in the results when there is fluctuations in the market.
5. Monte Carlo simulation also helps in comparative analysis of various optimization approaches thus helping users decide which optimizer best serves his/her purpose.
6. Comparative analysis from simulation is presented in the form of a bar chart to the user which helps in easy comparison between optimizers pictorially.

3.1 User Portfolio

After Onboarding of users onto the Platform the first thing we take input from the user is, User Details and how much amount of money they want to invest. After Taking input from users the next step will be the "List of Assets User wants to select".

When users get Onboarded they have to provide an amount of money and a list of assets from tickers in which they want to invest their money. And the Next page team will show them some trending assets and information about them, if users want to select some of them and they can and can skip also. From the Historical Asset data extract information about the assets that the user has selected. And next will be to analyze the expected return and create a risk model.

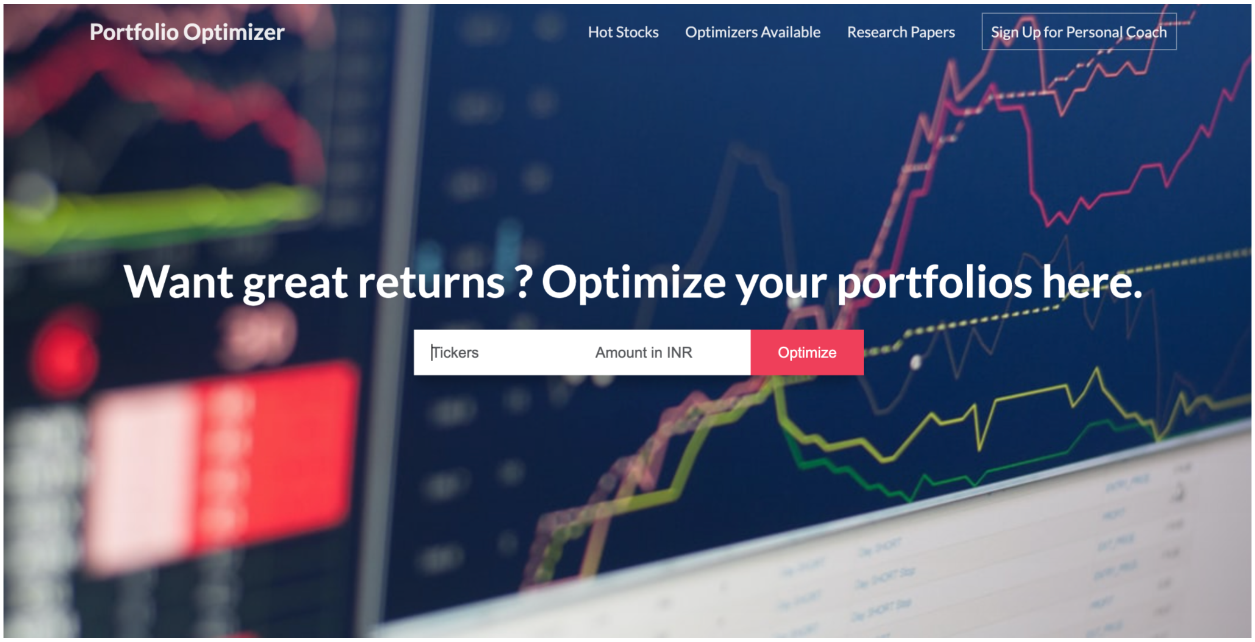


Fig 3.2 PortFolio Onboarding page of User Details

After taking a List of stock input Array and amount given by User Next Step will be to give input to the Analysis part.

3.2 Historical Asset Data

Yahoo finance Api is used to fetch all available data of the stocks selected by the user. The data is in the form of a time series array which contains all trading days data.
The close value for each day is being used for calculating the Expected returns and formulating a Risk Model.

3.3 Analysis

There are two prerequisites for the optimisers which are not generally available and needs to be calculated:

3.3.1 Forecast returns

"pypfopt" library is used to getting the mean historical return of the portfolio from the historical data.

Expected Portfolio returns : $$\mathrm{E}(R_p) = \sum_i w_i \, \mathrm{E}(R_i)$$ where Ri is the return on asset i and and wi is the weighting of component asset i.

3.3.2 Create Risk Model

Covariance matrix and Correlation matrix are being calculated using “numpy”.

$$\operatorname{cov}(r^A, r^B) = \sigma_{r^A, r^B} = \frac{1}{n} \sum_{i=1}^{n} \left(r_i^A - \mu^A\right) \left(r_i^B - \mu^B\right)$$

3.4 Optimizers

3.4.1 Markowitz Portfolio Theory

The modern portfolio theory (MPT) is a method that can be used by risk-averse investors to construct diversified portfolios that maximize their returns without unacceptable levels of risk.

Users which are more concerned with drawback chance might incline toward the Post Modern Portfolio to modern portfolio theory. Another Advantage of MPT is that it can decrease instability. Markowitz model uses the correlation between the stocks to predict an optimal allocation of weight.

For MPT we use the Efficient Frontier approach and draw a line of risk free return.
The point where line and curve intersect become the market portfolio. The best optimized portfolio we get at this point.(Max return for minimum risk).

Basic Idea : Classical Mean-Variance Optimizers (Efficient Frontier)

Input : Risk Model and forecast return

Output from this we get Optimal Weight distribution .

Maximize Sharpe Ratio among portfolios on efficient frontier:

- Slope of capital allocation line

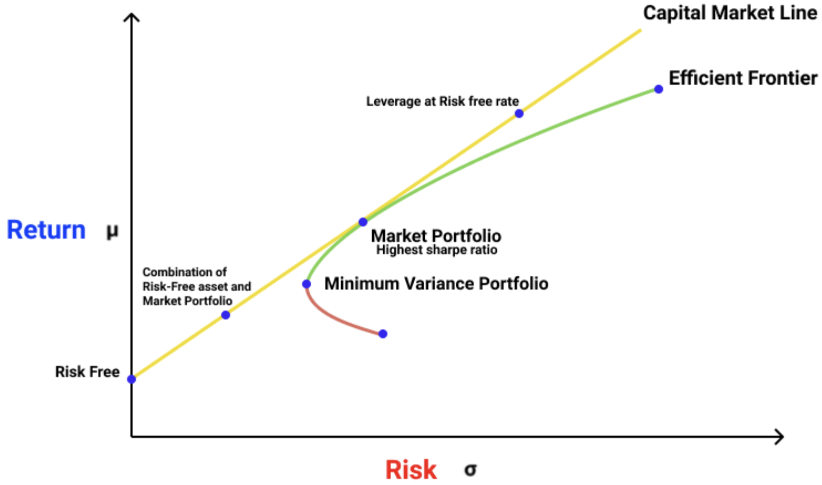


Fig 3.3 Capital Assignment Line for Asset Allocation

3.4.2 Risk Parity

In the Modern Portfolio Theory Risk is Assigned to whole portfolio but In Risk Parity Portfolio It focuses more on the Allocation of Risk rather than Allocation of Assets.

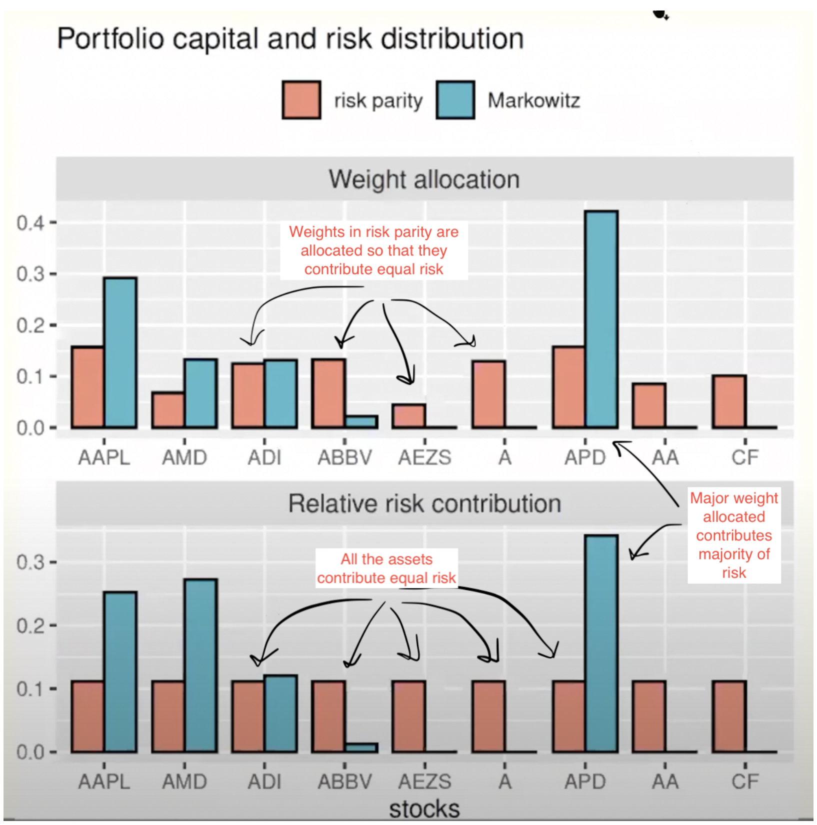


Fig 3.4 Comparison of Risk Allocation of RPP and MPT

The risk parity portfolio (RPP), also known as the equal risk portfolio (ERP), aims to "equalize" risk by ensuring that each asset's risk contribution is equal, rather than just having an equal capital allocation like the equally weighted portfolio (EWP).

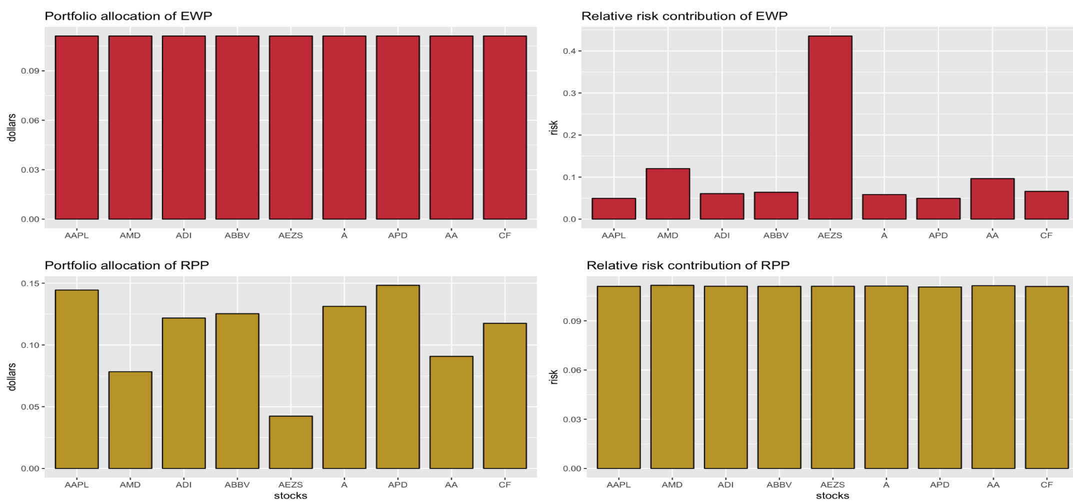


Fig 3.5 Comparison of EWP and RPP for portfolio Asset Allocation

3.4.3 HRP

Hierarchical Risk Parity Consists of 3 Stages as follows:-

Stage 1:- Hierarchical Tree Clustering:-

- Build a distance Matrix ← correlation matrix where d(i,j) = √ 1 / 2 (1 − ϱ (i,j))
- Followed by clustering using some heuristic over Euclidean distance (calc from dist[][]).

	a	b	c	d	e
a	0	17	21	31	23
b	17	0	30	34	21
c	21	30	0	28	39
d	31	34	28	0	43
e	23	21	39	43	0

Fig 3.6 Distance Matrix for HRP

Final step of Tree Clustering Visualized in the form of a nice cluster diagram called dendrogram. One such dendrogram can be seen below.

Usually in case of risk parity all assets are allocated with an equal risk contribution but in case of HRP the assets are divided into clusters on the basis of correlation among them. In the below figure the whole portfolio can be divided into multiple clusters each incorporating assets of a particular sector. These clusters as a whole are then treated as a single entity hence each of the clusters below contributes 25% to the risk and then this risk is further divided down the tree in a similar manner of clustering.

Fig 3.7 Hierarchical Tree Structure for Clustering

Stage 2:- Matrix Seriation(Quasi-Diagonalization)

In technical terms greater covariances are placed along the diagonal and smaller covariances are present around this. Also due to the fact that the off-diagonal elements are not perfectly zeros, this matrix is referred to as a quasi-diagonal covariance matrix. This step filters out the correlation between similar and dissimilar assets to understand the corresponding changes in a better way.

- Reorganizes rows & cols of covariance matrix. Larger Values are placed around the main diagonal
- After this stage similar Assets are together and dissimilar are far apart in the covariance matrix.

Stage 3: Recursive Bisection

This is the final and the most important step of this algorithm where the actual weights are assigned to the assets in our portfolio.

- Distribute the allocation through recursive bisection based on cluster covariance.

3.5 Monte Carlo Simulation on input to pick a better algorithm

3.5.1 Need for Simulation

Most of the time portfolio optimization decisions are taken by blindly picking up any of the above algorithms , in the hope that the results will always align with the requirements and lead to better results. But the results sometimes are unstable and even after optimization of the portfolio there are not any good gains which totally remove the point of optimization.

Different optimization approaches might cater to different sets of inputs and this methodology tries to exploit on that front trying out various algorithms for a given set of input and identifying which algorithm has what to offer , as , and it is unrealistic to expect that one method will dominate all under varied circumstances

3.5.2 Problem Formulation

The problem monte carlo is trying to solve here is formulated as a

> System with N random variables, where the expected value of draws is μ, and the variance of these draws is V ,the covariance matrix and we need to calculate ω.

These variables μ and V are typically unknown and are calculated using the historical stock price data which is acquired from yahoo finance.
In this method multiple simulated values of μ and V are calculated using a random seed and then Ledoit-Wolf shrinkage and then denoising to get $\hat{\mu}$ and $\hat{V}$ respectively.
Now $\hat{\omega}$ is calculated for the simulated values multiple times.
The mean of the calculated values gives the glimpse on how sensitive is portfolio w.r.t to any particular algorithm.

Thus MCOS method is used for estimating $\hat{\omega}$ while controlling for noise-induced and signal-induced instabilities.

3.5.3 The Monte Carlo Estimation method

There are basically five steps in this method , with input being the list of expected price of the set of stocks in the portfolio and the covariance matrix.

3.5.3.1 Calculating simulated predictions and covariance

The algorithm uses the original matrix X of size TxN i.e the matrix of time series data of N stocks for T days in order to obtain the simulated pair $\{\hat{\mu}, \hat{V}\}$ by multivariate using a random seed. Ledoit-Wolf shrinkage may be applied to X if required.
Now this simulated pair is used instead of the original $\{\mu, V\}$ corresponds to the true values.

3.5.3.2 Removing the noise

In this step noise reduction is done on the covariance matrix , this helps in preventing the instability of input data to the algorithms. This project uses a Kernel Density Estimate (KDE) algorithm to fit the Marcenko-Pastur distribution to the empirical distribution of eigenvalues. This helps in the separation of noisy eigenvalues from signal related eigenvalues.

3.5.3.3 Applying allocation algorithms

In this step a variety of portfolio allocation algorithms namely , Mean-variance optimizer , Risk parity optimizer , and Hierarchical risk parity optimizer are applied to the simulated values and $\hat{\omega}*$ is estimated.

3.5.3.4 Applying the Simulation

This step is the combination of all the 3 steps mentioned previously, now optimal allocation is done for simulated pairs $\{\hat{\mu}, \hat{V}\}$ via each of the optimizers $\hat{\omega}^*(i)$, where $\hat{\omega}^*$ is the optimal allocation via i-th optimizer.

This calculation is done multiple times , the no. of times is defined by the user , greater the number more accurate will be the predictions made.

3.5.3.5 Error Estimations

In this step the optimal allocation is done for original pair $\{\mu, V\}$ for each optimizer ,i.e. $\omega^*(i)$, where ω^* is the optimal allocation via i-th optimizer and the results are compared with estimated $\hat{\omega}*$.

The estimation error may be evaluated in the following terms for each i :

- the mean difference in expected outcomes, $(\omega*(i) - \hat{\omega}*(i))'\mu$
- the mean difference in variance, $(\omega*(i) - \hat{\omega}*(i))'V(\omega*(i) - \hat{\omega}*(i))$
- the mean difference in Sharpe ratio, $(\omega*(i)-\hat{\omega}*(i))'\mu \;/\; \sqrt{(\omega*(i)-\hat{\omega}*(i))'V(\omega*(i)-\hat{\omega}*(i))}$

3.6 Optimized Portfolio

After the analysis the user will receive multiple barcharts and pie charts with each having a significance of its own alongwith the details which looks similar to

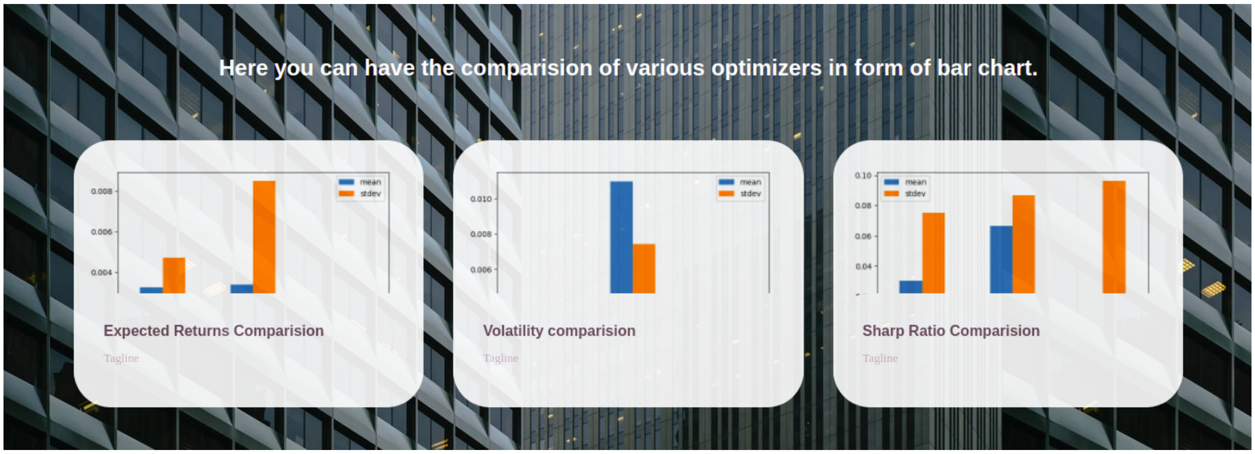


Fig 3.8 : Error estimations of various portfolios w.r.t multiple estimates

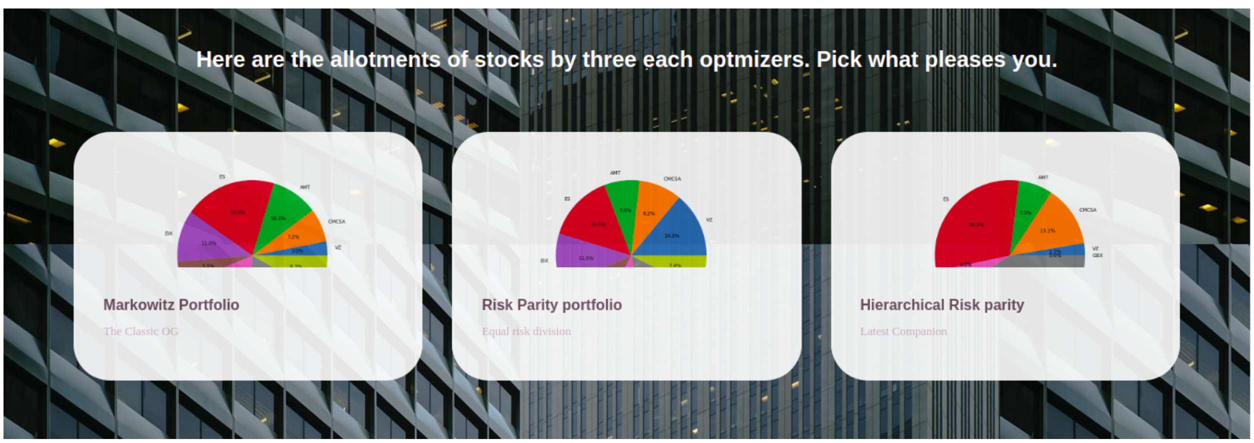


Fig 3.9: Pie Chart for allocation via different optimizers

3.7 Graphical Visualization Explanation

A graphical representation of the different error estimates is being plotted as a bar graph and presented to the user for estimation.

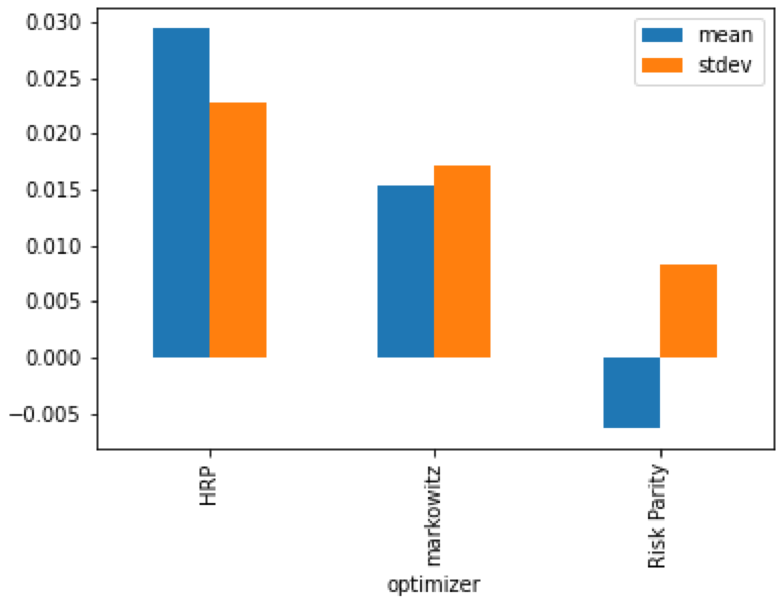


Fig 3.10: Expected Outcome Error Bar Chart

A bar chart is presented to the user representing the mean error in outcome subjected to slight variations in the correlation and forecast and standard deviation for the same. A highly positive variation in case of HRP signifies that the outcomes were majorly on the positive side to the predicted value of outcome. A negative deviation on risk parity signifies a little lesser outcome in majority than the predicted value.

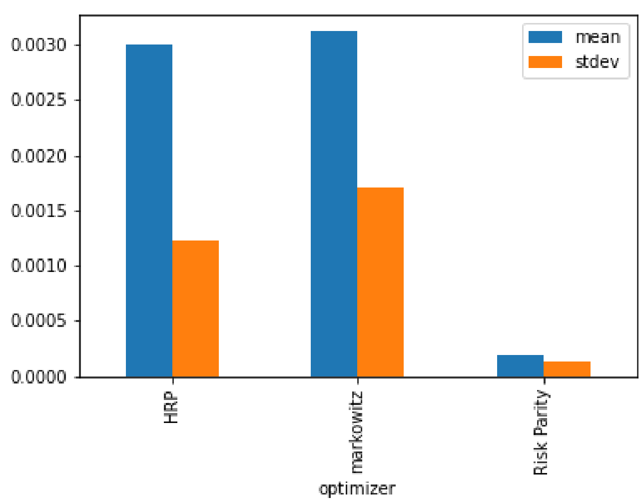
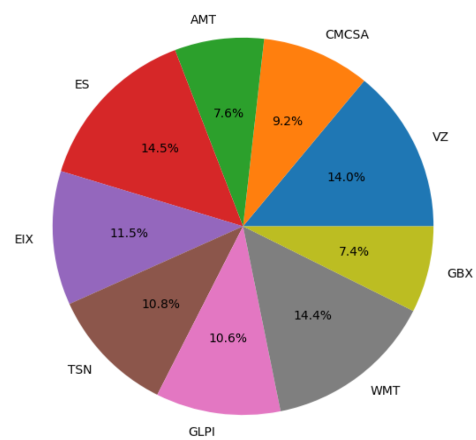


Fig 3.11: Variance Error Bar Chart

The variance error signifies the error in volatility of the portfolio when subjected to expected variations in correlation and stock price forecast as it can be clearly seen that Risk Parity has a very low deviation from the predicted variance and may be presented as a sign of better adaptability to fluctuations in the market. On the other hand Markowitz has a higher variance error which makes it the most susceptible to market fluctuations.



3.12 Risk Parity based allocation of funds

The above pie chart represents the relative amount of total funds a user has to invest in a particular fund e.g. here user is supposed to invest 1400 USD for VZ ticker from overall capital to be allocated.

3.8 Out of current Scope

1. The platform created takes just the array of stocks and amount to be invested , no information about the user is taken into account.
2. Portfolio optimization is independent of the users personal risk appetite and is not personalized for different users as per their choices.
3. Portfolios are optimized using only historical data and market semantics are not taken into consideration.
4. The platform design doesn't suggest stocks users can invest in.

Chapter 4
Experimental Results

4.1 Objective: To analyze the performance of a stock portfolio after various optimization techniques of portfolio allocation.

4.2 Method : A portfolios of multiple stocks belonging to different sectors is chosen:

A dataset containing the price history of these stocks is being taken starting from 1990 till 2020 to create optimized portfolios using the methodology above.The data for the analysis has been acquired from yahoo finance api.

Following optimizers along with some a few random allocations have been used to distribute capital among the assets :

1. HRP Optimiser
2. MarkowitzOptimizer
3. RiskParityOptimizer

Monte Carlo simulation is performed with 50 iterations and the bar chart below provides contrast between various forementioned optimization algorithms for the following errors :

1. Expected Outcome Error.
2. Variance Error.
3. Sharpe Ratio Estimate Error.

Afterwards a line chart was created for the three optimized portfolios by using data from 2020 till present to see how our portfolios would have performed if we would have invested our money using optimized portfolios.

4.3 Analysis

Portfolio : [`VZ,CMCSA,AMT,ES,EIX,TSN,GLPI,WMT,GBX`]

Here are rebalanced portfolio returns for each of the optimisers for the portfolio under analysis

1. Annualized mean return by HRP : 14.84%
2. Annualized mean return by Markowitz : 14.05%
3. Annualized mean return by RiskParity : 14.18%

It is pretty evident from the above numbers that HRP performs better than the other two followed by Risk parity.

To get even more insights about optimized portfolio, below are the few barcharts created using the above method *(i.e Monte Carlo Simulation)* that form the basis for our analysis.

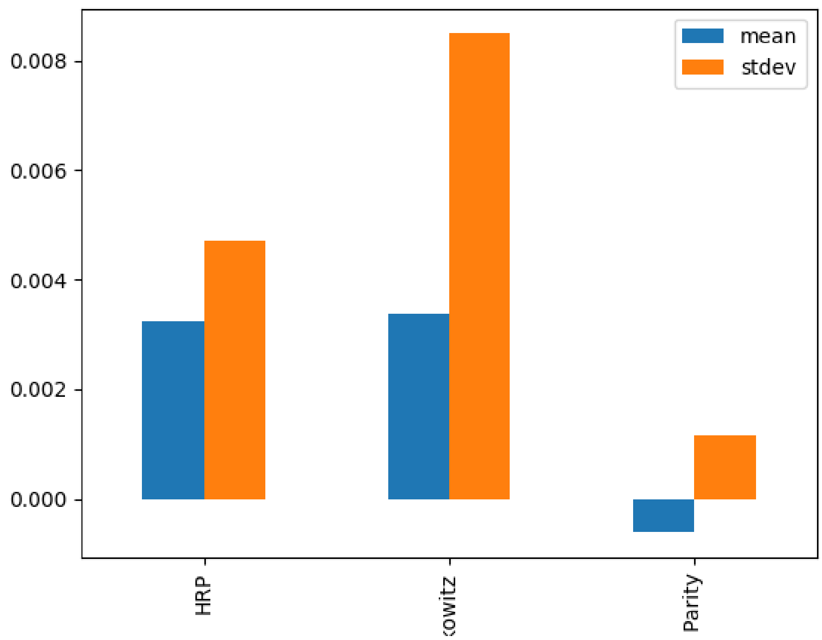


Fig 4.1 Expected Outcome error estimator for mean and standard deviation

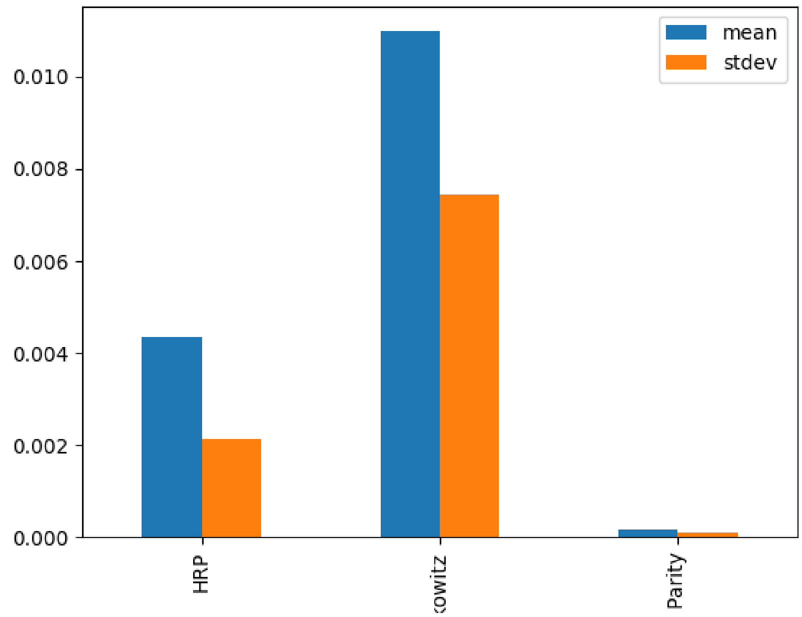


Fig 4.2 Variance error estimator for mean and standard deviation

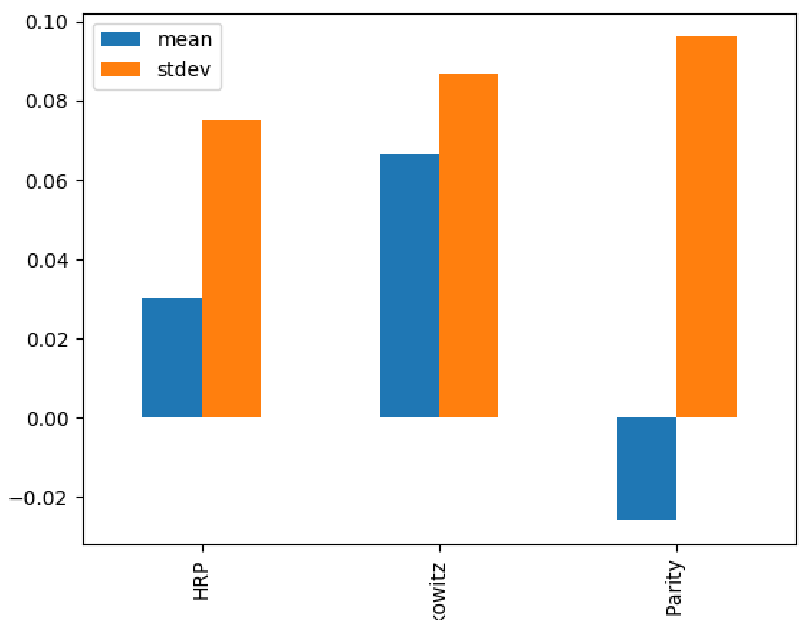


Fig 4.3 Sharpe Ratio Error Estimate

From fig 4.1 we can see that the expected outcome error is on the positive side and nearly comparable in case of HRP and Markowitz (i.e 0.3 %) so we could safely pick either but since the expected portfolio return of HRP is decently greater than Markowitz even after +0.3% HRP is way better followed by risk parity . Similar is the case when we look into fig 4.3

Also fig 4.2 makes a case against Markowitz as even after having high volatility compared to other two it shows very high positive error in variance compared to the other two which can be interpreted as high susceptibility to slight changes in the market. HRP shows higher variance error when compared to Risk parity which makes Risk parity a much more secure option.

Due to high variance error we put Markowitz portfolio out of the picture and choose between HRP and Risk Parity , since HRP seems to provide a higher return with slightly more risk we try to use it for our portfolio.
(A person expecting even lesser risk error can go with Risk parity).

Below is the rebalanced portfolio for stocks based on HRP.

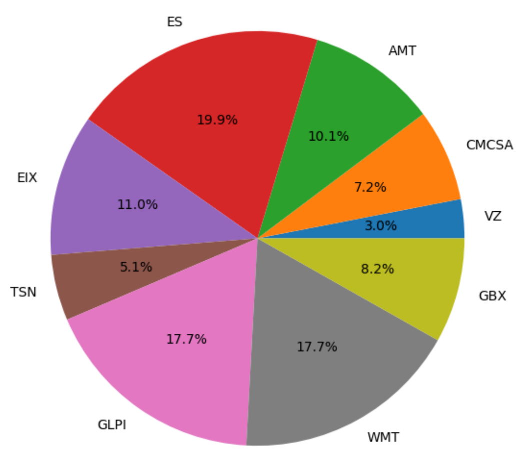


Fig 4.4 Distribution of stocks by amount to be allocated predicted by HRP

4.3 Observations:

Here is the plot of used portfolio value for using different optimizers for time duration 2020 - present (06-2022)

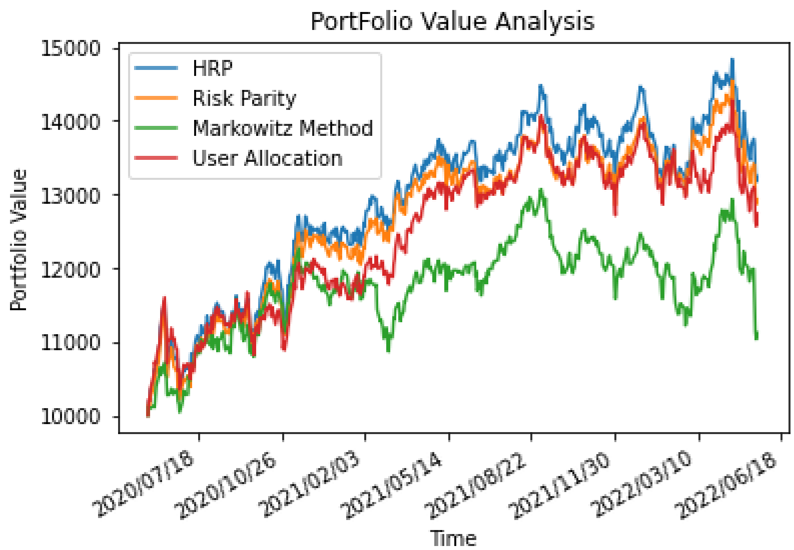


Fig 4.5 Performance of various portfolio allocation strategies over a period of two years.

Following observations can be made :

1. The results are in coherence with the analysis made above and our HRP-optimized portfolio was able to beat others in terms of returns.
2. HRP and Risk Parity based portfolios were able to retain much more value during and after market fluctuations from in and after 10/2020 (COVID onset).
3. Initially Markowitz did fairly decently but eventually showed a decline with time as expected from analysis.
4. HRP and Risk Parity beats the user defined portfolio , which surprisingly worked better than Markowitz.

4.3 Conclusions :

Analysis performed over barcharts created for a variety of errors using Monte Carlo simulation can help better decide which optimized portfolio to be used for the user's purpose and allows the user to perform qualitative comparative analysis between various optimization approaches.

4.4 Libraries/Tools :

- Pandas for stock data manipulation with data frame
- Matplotlib for plotting graphs
- Numpy for array manipulation and calculation of covariance, correlation
- Scipy for calculations

Chapter 5
Conclusions

5.1 Conclusions

Portfolio optimization is a excellent tool for investors looking to maximize their risk-to-reward Ratio. They can do so with the assistance of this project, which can help them choose the best algorithm as per their need and risk appetite. This can help them choose the best strategy knowing the tradeoffs

The decision would always be based on the investors' risk appetite and expected rate of return. Finally while any model or theory has advantages and disadvantages, portfolio managers can maximize the benefits of the portfolio maximizing technique if they employ it diligently.

This model can also help developers to build upon it , create better UI and extensions to distribute the capability of this project just by simple plugins.

This is a simple yet powerful tool which can help retail investors ramp up their investment journey after going through the basic strategies / algorithms available for diversification.

Investors simply have to provide the set of stocks they need to invest in , the amount to be invested and choose their strategy according to the analysis we present. The tool will help them with what amount of stock they need to buy.

Right now the tool supports HRP, MPT and Risk Parity as algorithms but the list can be ever growing and can be built upon by developers.

5.2 Future Work

Many different adaptations, tests and experiments have been left for the future due lack of user Traffic and data sets. Future work concerns deeper analysis of some methods, new curiosity and new proposals as per time.

Here are a few things that can be onboarded as an extension to the project :

1. Optimize portfolio with respect to the volatility provided by the investor.
2. Inclusion of multiple categories of assets like gold, bonds etc in the portfolio.
3. Improvement in the operational UI which makes it easy for non technical users to utilize the capabilities of this project.
4. Improvement in Trending Assets section on the home page which allows naive users to pick in their favorites from trending assets in order to try this tool.
5. Analyzing various deep learning algorithms like LSTM and CNN to further optimize the performance of the portfolio.
6. Adding functionality to suggest users different assets based on their negative correlations with other assets.

References

1. Lopez de Prado, M. (2016). A robust estimator of the efficient frontier. Available at SSRN 3469961.
2. Pfitzinger, J., & Katzke, N. (2019). A constrained hierarchical risk parity algorithm with cluster-based capital allocation. Stellenbosch University, Department of Economics.
3. De Prado, M. L. (2016). Building diversified portfolios that outperform out of sample. The Journal of Portfolio Management, 42(4), 59-69.
4. Dalio, R. (2005). Engineering targeted returns and risks. Alpha Manager.
5. Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.2307/2975974
6. Costa, Giorgio, and Roy H. Kwon. "Risk parity portfolio optimization under a markov regime-switching framework." *Quantitative Finance* 19.3 (2019): 453-471.
7. Gambeta, Vaughn, and Roy Kwon. "Risk return trade-off in relaxed risk parity portfolio optimization." *Journal of Risk and Financial Management* 13.10 (2020): 237.
8. Ivanova, Miroslava, and Lilko Dospatliev. "Application of Markowitz portfolio optimization on Bulgarian stock market from 2013 to 2016." *International Journal of Pure and Applied Mathematics* 117.2 (2017): 291-307.
9. Michaud, Richard O., and Robert Michaud. "Estimation error and portfolio optimization: a resampling solution." Available at SSRN 2658657 (2007).