Diversified Portfolio Optimization for Indian Equities: Using Sharpe Ratio Maximization

A Dissertation Submitted to
SVKM's NMIMS (Deemed to be University)
In Partial Fulfillment for the Degree of

Bachelor of Science
in
Applied Mathematical Computing
By

Juhi Kariya (A005)
Diyanshi Shah (A010)
Tanushri Shetty (A011)

Under the Supervision of
Dr. Debasmita Mukherjee

NMIMS NSOMASA
Ground Floor, SBMP Phase I,
Irla, N. R. G Marg, Opposite Cooper Hospital,
Vile-Parle (West), Mumbai – 400056

April 2025

Certificate

This is to certify that the work described in this thesis entitled **"Diversified Portfolio Optimization for Indian Equities: Using Sharpe Ratio Maximization"** has been carried out by Juhi Kariya, Diyanshi Shah, and Tanushri Shetty under my supervision. I certify that this is their bonafide work. The work described is original and has not been submitted for any degree to this or any other University.

Date: 10th April 2025

Place: Vile Parle, Mumbai

Supervisor:

Dr. Debasmita Mukherjee

Acknowledgement

We take this opportunity to express our heartfelt gratitude towards our mentor, Dr. Debasmita Mukherjee, who supported us throughout the project. She guided us with her advice for topic selection. Dr. Mukherjee was very communicative, which made it easier for us whenever we needed help. She directed us on how we can improve our project. Without her guidance going about this project, would have been difficult.

We would also like to thank Nilkamal School of Mathematics, Applied Statistics & Analytics Mumbai for giving us this opportunity, and other faculties who helped us when we faced obstacles in our project.

Contents

1 Abstract

This project presents a comprehensive framework for portfolio optimization tailored to the Indian equity market. The objective is to construct a diversified portfolio that maximizes the Sharpe ratio while adhering to user-specified risk tolerance and investment budget constraints. A selection of 25 stocks is made from five key sectors—Technology/IT, Banking & Finance, Consumer Goods, Healthcare, and Infrastructure—ensuring diversification across market capitalizations (large-cap, mid-cap, and small-cap) to replicate real-world market conditions. The methodology begins by generating all possible five-stock portfolios, each comprising one stock from every sector. A filtering step eliminates high-volatility combinations that exceed the user's input risk tolerance threshold by more than 5%. For the remaining portfolios, an optimization algorithm employing Sequential Least Squares Programming (SLSQP) determines the optimal asset weights to maximize the Sharpe ratio, with constraints ensuring no single stock receives less than 1% allocation.

The framework is applied to historical data from 2020–2023 to identify the optimal portfolio, which is subsequently validated on out-of-sample data (2023–2025) to evaluate its real-world performance. Performance is benchmarked against an equal-weighted portfolio of all 25 stocks, and additional metrics such as the Sortino ratio and maximum drawdown are also calculated, providing further insights into risk-adjusted returns. A user-friendly interface, developed using Flask in Google Colab, is built that facilitates dynamic input of investment parameters. Our resulting model offers a practical and easy to use tool backed by data-driven decision making for investors in the Indian equity market.

2 Introduction

In today's world, investors are often overwhelmed by the vast amount of financial information available. With countless stocks, news updates, and opinions, making smart investment choices has become more confusing than ever. Many people want to grow their money wisely but don't have the time or tools to analyze everything themselves.

Portfolio optimization plays a pivotal role in filling this gap by enabling investors to allocate capital across assets in a way that maximizes returns for a given level of risk. In this project we develop a sector-diversified, Sharpe ratio-maximizing portfolio optimization model designed for the Indian equity market. By incorporating user-defined constraints and leveraging historical data, the model aims to provide a practical and intuitive yet easy to use investment decision-making tool suited to a wide range of investor profiles.

3 Literature Review

The field of portfolio optimization has undergone significant evolution, progressing from simplified linear models to modern AI-driven approaches. This progression is reflected in three key works that represent distinct methodologies and objectives across different eras.

In *A Linear Programming Algorithm for Mutual Fund Portfolio Selection*, Sharpe [1] extends Markowitz's portfolio theory by introducing a single-index model, where security returns are regressed on a common market index. This reduces computational complexity by minimizing the number of required covariance estimates and assumes uncorrelated residuals across securities. Sharpe reformulates portfolio variance using beta coefficients, allowing mutual fund managers to construct efficient portfolios via linear programming. Although less precise for low-risk portfolios due to potential residual correlations, the approach offers a practical alternative for high-return targeting under computational constraints.

Lee and Lerro [2], in *Optimizing the Portfolio Selection for Mutual Funds*, introduce a goal programming (GP) model to integrate investor preferences into the optimization process. Moving beyond efficiency frontiers, their approach balances conflicting goals — including maximizing returns, minimizing variance, ensuring dividend income, and limiting unexplained price variation — while honoring real-world constraints such as legal requirements and diversification. Using data from 61 firms across 10 industries, they show that GP-optimized portfolios outperformed actual mutual funds in 1969, particularly in bearish scenarios.

Abdi et al. [3], in *Prospective Portfolio Optimization With Asset Preselection Using a Combination of Long and Short Term Memory and Sharpe Ratio Maximization*, present a data-driven three-stage framework that blends deep learning with financial optimization. They use Long Short-Term Memory (LSTM) networks to forecast stock prices more accurately than traditional ARIMA models, validated using FTSE 100 data spanning 20 years. These forecasts guide asset preselection under cardinality constraints, followed by weight allocation aimed at maximizing the Sharpe Ratio. The model exhibits strong performance under diverse market conditions, demonstrating how machine learning can enhance both prediction and allocation.

These three studies collectively map the advancement of portfolio optimization — from Sharpe's computational simplification to Lee & Lerro's preference-based modeling and Abdi et al.'s predictive intelligence. Together, they showcase how theoretical foundations and modern technologies can intersect to inform smarter portfolio construction.

4 Data Description

This project takes inspiration from the principles of Modern Portfolio Theory (MPT), which emphasizes diversification to optimize returns while minimizing risk. At its core, MPT suggests that constructing a portfolio across various assets with different risk-return profiles can lead to better performance than investing in individual assets alone as this helps in reducing unsystematic risk.

To put this into practice, we curated a dataset of 25 companies listed on the National Stock Exchange (NSE) of India. These stocks are distributed evenly across five key sectors — Technology, Banking & Finance, Consumer Goods, Healthcare, and Infrastructure — with five companies representing each sector. This balanced approach helps simulate a realistic and diverse portfolio structure while keeping the dataset manageable and scalable.

4.1 Stock Selection Strategy

The stock selection was done with several goals in mind: realism, diversity, and computational feasibility. Here's how we approached it:

- **Diversification Across Sectors and Sizes :** To reflect real-world investment practices, we studied 2–3 prominent sectoral mutual funds in each domain to understand their allocation patterns across market capitalizations (large-cap, mid-cap, small-cap & other). Using these insights, we mirrored similar proportions in our dataset. This avoided an over-concentration in large-cap companies and allowed for exposure to mid and small-cap stocks, which often have different risk-return dynamics. The selection ensures representation across various segments of the market. Each sector includes a mix of company sizes — while not rigidly categorized, the distribution mimics fund behavior and prevents bias toward any single market cap.

- **Practicality and Scalability :** Given limitations in computational resources and data accessibility, especially when using the Yahoo Finance API (via the yfinance Python package), we intentionally restricted the dataset to 25 companies. This size allows for meaningful analysis without becoming unmanageable. Although this is a simplified and scaled-down representation, care was taken to make it reflective of actual portfolio management practices. The data incorporates companies with varied risk and return characteristics, laying the groundwork for evaluating performance through multiple metrics.

The final dataset of 25 selected stocks, along with their computed financial metrics, is summarized in the below table.

Company Name	Sector	Size	Expected Return	Volatility	Sharpe Ratio
Tata Consultancy Services	Tech/IT	Large Cap	0.0792	0.1879	0.0648
Infosys	Tech/IT	Large Cap	-0.1147	0.2109	-0.8614
Oracle Financial Services	Tech/IT	Mid Cap	0.1842	0.1502	0.7805
IndiaMart	Tech/IT	Other	0.1224	0.2496	0.2219
PVR Inox	Tech/IT	Small Cap	-0.3628	0.2917	-1.4734
State Bank of India	Banking & Finance	Large Cap	-0.3133	0.2493	-1.5257
ICICI Bank	Banking & Finance	Large Cap	-0.0841	0.1776	-0.8505
Nippon Life India Asset Management	Banking & Finance	Mid Cap	-0.3752	0.2464	-1.7946
Arman Financial Services	Banking & Finance	Small Cap	-0.1540	0.3454	-0.6397
Cholamandalam Investment and Finance	Banking & Finance	Other	0.4263	0.2873	1.2508
Hindustan Unilever	Consumer	Large Cap	0.0684	0.1893	0.0075
Mold-Tek Packaging	Consumer	Small Cap	0.2216	0.3163	0.4886
The Bombay Dyeing and Manufacturing Company Limited	Consumer	Mid Cap	-0.6088	0.3272	-2.0654
Polycab India	Consumer	Other	0.2220	0.2699	0.5742
United Breweries	Consumer	Large Cap	-0.3889	0.1542	-2.9566
Sun Pharmaceutical	Healthcare	Large Cap	0.0331	0.1483	-0.2282
Zydus Lifesciences	Healthcare	Mid Cap	0.3501	0.2264	1.2505
Gland Pharma Limited	Healthcare	Other	-0.6679	0.3790	-1.9393
AstraZeneca Pharma India	Healthcare	Small Cap	-0.1003	0.2021	-0.8276
Fortis Healthcare	Healthcare	Other	-0.2142	0.2603	-1.0805
NTPC	Infrastructure	Large Cap	0.1224	0.1653	0.3353
Honeywell Automation	Infrastructure	Mid Cap	-0.0676	0.2464	-0.5461
Oriental Cement	Infrastructure	Small Cap	0.0144	0.3254	-0.1616
Hindustan Aeronautics	Infrastructure	Other	0.4142	0.3299	1.0526
Reliance Industries	Infrastructure	Large Cap	-0.2137	0.2065	-1.3589

Figure 1: Summary of selected stocks and their risk-return metrics

4.2 Sector-wise Risk and Return Analysis

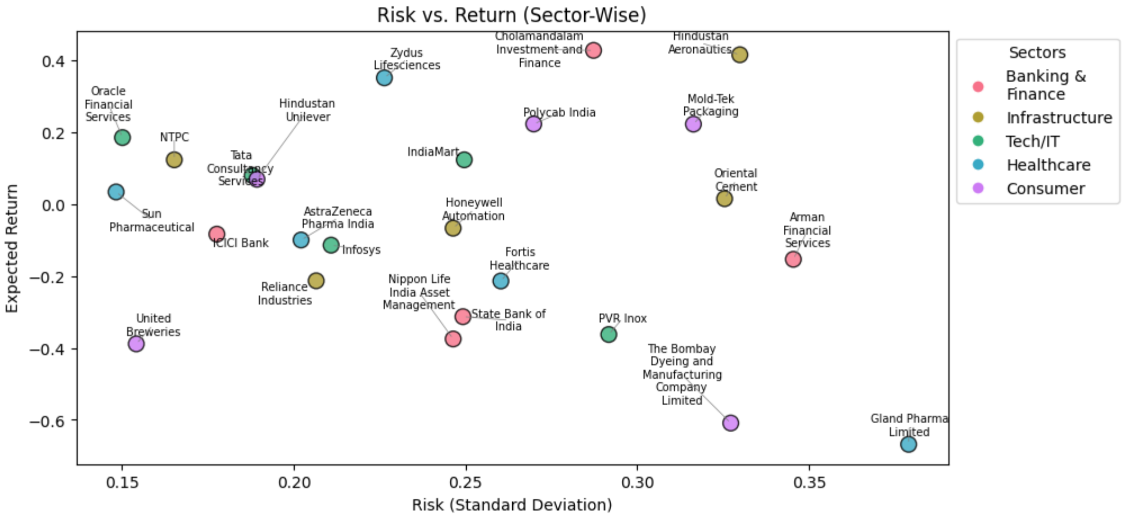

Figure 2: Sector-wise distribution of stocks based on expected return and risk

The scatter plot visualizes the relationship between expected return and volatility (risk) for the selected stocks, color-coded by sector. Several sectoral and stock-level insights can be drawn from this distribution:

- **Sector-wise Dispersion:** Sectors like Banking & Finance and Infrastructure show a wide dispersion across both risk and return, indicating considerable variability among constituent stocks. This suggests that intra-sector diversification may be beneficial when selecting companies within these sectors.
- **Tightly Clustered Sectors:** Healthcare stocks exhibit a more concentrated distribution, generally residing in the low-to-moderate risk zone with moderate to negative expected returns. This reflects a relatively defensive sector profile, often preferred in low-volatility portfolios.
- **High-Risk, High-Return Candidates:** Companies such as Polycab India, Mold-Tek Packaging, and Hindustan Aeronautics display high expected returns alongside above-average volatility, positioning them as potential high-risk, high-reward options.
- **Underperforming Outliers:** Gland Pharma Limited stands out as a clear outlier, with both the highest risk and a strongly negative expected return, making it an unattractive investment under conventional risk-return assumptions.
- **Stable, Low-Return Stocks:** Firms like ICICI Bank, Sun Pharmaceutical, and NTPC lie in the low-risk, low-return quadrant, often favored by conservative investors seeking stability over aggressive gains.

- **Mixed Sector Patterns:** The Tech/IT and Consumer sectors display a diverse spread, with some companies like Oracle Financial Services showing promising returns with moderate risk, while others like PVR Inox and United Breweries have significantly negative returns. This indicates that sector alone may not be a reliable indicator of performance, highlighting the need for stock-level analysis.

These trends provide useful guidance for constructing diversified portfolios, aligning stock selection with the investor's risk appetite, and identifying sector-specific behavior in the risk-return landscape.

5 Methodology

5.1 Generating and Filtering Stock Combinations

To ensure robust diversification, our portfolio strategy involves a meticulous process of selecting one stock per sector from the following five key sectors: Tech/IT, Banking & Finance, Consumer, Healthcare, and Infrastructure. Given that each sector comprises 5 stocks, the total number of possible combinations, each with a unique selection of one stock from each sector, is calculated as:

$$\text{Total Combinations} = 5^5 = 3125$$

However, many of these combinations may not align with an investor's specific risk tolerance. Therefore, we filter these combinations based on a preliminary assessment of portfolio risk to identify those that meet the investor's criteria.

Preliminary Portfolio Risk Assessment

For each possible stock combination, we begin by assigning equal weights to each constituent stock. The portfolio variance is then computed as:

$$\sigma_p^2 = \mathbf{w}^\top \cdot \mathbf{\Sigma} \cdot \mathbf{w}$$

Where:

- $\mathbf{w}$ represents the initial weight vector, with each element set to $\frac{1}{n}$ (where n denotes the number of stocks).
- $\mathbf{\Sigma}$ represents the covariance matrix of the stocks' returns, quantifying the interdependencies between their price movements.

The portfolio risk is subsequently calculated as the standard deviation:

$$\sigma_p = \sqrt{\sigma_p^2}$$

We then filter out any combination where the portfolio risk, σ_p, exceeds the investor's maximum acceptable risk. The investor's maximum acceptable risk is defined as the sum of the investor's inherent risk appetite and a predetermined threshold. In our model, we employ a threshold of 5%.
This filtering step effectively reduces the number of combinations to a feasible set, which is then passed on to the subsequent stage of optimization.

5.2 Sharpe Ratio

To determine the optimal portfolio from the filtered combinations, we employ the Sharpe Ratio, a widely recognized metric for evaluating a portfolio's risk-adjusted performance.

Understanding the Sharpe Ratio

The Sharpe Ratio is formally defined as:

$$\text{Sharpe Ratio} = \frac{R_P(\mathbf{w}) - R_f}{\sigma_P(\mathbf{w})}$$

Where:

- $R_P(\mathbf{w})$: The annualized return of the portfolio, which is directly dependent on the portfolio weights $\mathbf{w}$.
- R_f: The annualized risk-free rate of return.
- $\sigma_P(\mathbf{w})$: The portfolio volatility (risk), which is also a function of the weights $\mathbf{w}$.

1. Risk-Free Rate: The risk-free rate (R_f) represents the theoretical return of an investment with zero risk. In practice, it is often approximated using the return on government-backed securities, which are considered to have minimal credit risk. For our analysis, we utilize a risk-free rate of 6.7% , derived from a credible government source to ensure the robustness of our calculations:

$$R_f = 0.067$$

2. Annualized Portfolio Return ($R_P(\mathbf{w})$): The portfolio return is calculated as the weighted average of the expected annualized returns of the individual stocks within the portfolio:

$$R_P(\mathbf{w}) = \mathbf{w}^\top \mathbf{r}$$

Where:

- $\mathbf{r} = [r_1, r_2, \ldots, r_n]^\top$ is the vector of annualized returns for each stock in the portfolio.
- $\mathbf{w} = [w_1, w_2, \ldots, w_n]^\top$ is the vector of portfolio weights, representing the proportion of the investment allocated to each stock.

Each return r_i in $\mathbf{r}$ is derived from historical stock price data.

3. Portfolio Volatility ($\sigma_P(\mathbf{w})$): Portfolio volatility, a measure of the dispersion of returns, reflects the overall risk associated with the portfolio's returns. It is calculated from the covariance matrix Σ of the stock returns and the portfolio weights:

$$\sigma_P(\mathbf{w}) = \sqrt{\mathbf{w}^\top \Sigma \mathbf{w}}$$

Where:

- Σ is the covariance matrix of the stock returns, capturing the relationships between their fluctuations.
- $\mathbf{w}$ is the vector of portfolio weights.

Intuition Behind the Sharpe Ratio

The Sharpe Ratio effectively quantifies the excess return (above the risk-free rate) achieved per unit of risk undertaken. A higher Sharpe Ratio signifies superior risk-adjusted performance, and our primary objective is to identify the portfolio that exhibits the highest Sharpe Ratio.

5.3 Portfolio Weight Optimization

We then proceed to optimize the portfolio weights for each of the filtered stock combinations to maximize the Sharpe Ratio, thereby achieving the best possible balance between risk and return.

Optimization Objective

The optimization aims to determine the weight vector $\mathbf{w} = [w_1, w_2, \dots, w_n]^\top$ that maximizes the Sharpe Ratio:

$$\max_{\mathbf{w}} \quad \frac{R_P(\mathbf{w}) - R_f}{\sigma_P(\mathbf{w})}$$

Constraints

The optimization process is subject to the following constraints, which ensure the practicality and feasibility of the resulting portfolio:

1. **Budget Constraint:**
$$\sum_{i=1}^{n} w_i = 1$$
This constraint ensures that the entire investment budget is allocated across the selected stocks, with no funds left uninvested.

2. **No Short-Selling:**
$$w_i \geq 0.01 \quad \forall i \in \{1, 2, \dots, n\}$$
This constraint prevents short-selling. We set the lower bound to 0.01 (1%) to avoid any stock receiving a zero allocation.

Nature of the Optimization Problem

This optimization problem is classified as a constrained, nonlinear problem due to the following characteristics:

- The numerator $\left(\mathbf{w}^\top \mathbf{r} - R_f\right)$, representing the excess return, is linear in $\mathbf{w}$.

- The denominator $\left(\sqrt{\mathbf{w}^\top \Sigma \mathbf{w}}\right)$, representing the portfolio volatility, is nonlinear in $\mathbf{w}$.

Therefore, given the nonlinear nature of the problem, linear programming techniques are not suitable for this optimization.

5.4 Sequential Least Squares Programming (SLSQP)

To address this constrained nonlinear optimization problem, we employ the Sequential Least Squares Programming (SLSQP) algorithm. SLSQP is a robust and efficient algorithm, particularly well-suited for problems with smooth (differentiable) objective functions and constraints.

SLSQP Problem Formulation

SLSQP is designed to solve problems expressed in the following general form:

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\
\text{subject to:} \quad & c_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m \\
& d_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, p \\
& l_k \leq x_k \leq u_k, \quad k = 1, \ldots, n
\end{aligned}
$$

Where:

- $f(\mathbf{x})$ is the nonlinear objective function. In our case, it represents the negative of the Sharpe Ratio, as SLSQP is a minimization algorithm.
- $c_i(\mathbf{x})$ represents equality constraints. In our context, this corresponds to the budget constraint, which requires the sum of portfolio weights to equal 1.
- $d_j(\mathbf{x})$ represents inequality constraints, such as the non-negativity constraint (with a small lower bound) imposed on the portfolio weights.
- $l_k \leq x_k \leq u_k$ represents variable bounds. In our case, the portfolio weights are bounded between 0.01 and 1.

SLSQP Working Mechanism

SLSQP operates by iteratively approximating the nonlinear problem with a sequence of simpler quadratic programming (QP) subproblems. This iterative process refines the solution until convergence is achieved.
Step 1 - Quadratic Programming (QP) Approximation:

- At each iteration, SLSQP approximates the original nonlinear objective function (the Sharpe Ratio) with a quadratic function. This quadratic function is easier to optimize.

- A second-order Taylor series expansion is used to create this approximation around the current solution $\mathbf{x}$ (which represents our current best guess for the portfolio weights). The Taylor series expansion allows us to express the function in terms of its value, gradient, and Hessian at the current point.

- The Taylor expansion is:

$$f(\mathbf{x} + \Delta\mathbf{x}) \approx f(\mathbf{x}) + \nabla f(\mathbf{x})^\top \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^\top B \Delta\mathbf{x}$$

  Where:

  - $f(\mathbf{x} + \Delta\mathbf{x})$ is the value of the objective function at a new point near $\mathbf{x}$.
  - $f(\mathbf{x})$ is the value of the objective function at the current point $\mathbf{x}$.
  - $\nabla f(\mathbf{x})$ is the gradient of the objective function at $\mathbf{x}$. The gradient is a vector that points in the direction of the steepest increase of the function.
  - $\Delta\mathbf{x}$ is the step vector, representing the change in the portfolio weights that we are trying to find.
  - B is an approximation of the Hessian matrix. The Hessian is the matrix of all second-order partial derivatives of the objective function, and it describes the curvature of the function.

Step 2 - Linearization of Constraints:

- Similarly, the nonlinear constraints are approximated using first-order Taylor series expansions. This means we approximate the constraints with linear functions, which are also easier to handle.

- The Taylor expansion for the constraints is:

$$c_i(\mathbf{x} + \Delta\mathbf{x}) \approx c_i(\mathbf{x}) + \nabla c_i(\mathbf{x})^\top \Delta\mathbf{x}$$

$$d_j(\mathbf{x} + \Delta\mathbf{x}) \approx d_j(\mathbf{x}) + \nabla d_j(\mathbf{x})^\top \Delta\mathbf{x}$$

Step 3 - Solving the QP Subproblem and Updating the Solution:

- Now we have a Quadratic Programming (QP) subproblem: we are trying to optimize a quadratic objective function (from step 1) subject to linear constraints (from step 2). QP problems are a standard type of optimization problem that can be solved efficiently.

- The solution to this QP subproblem gives us the step vector $\Delta\mathbf{x}$, which tells us how to change our current portfolio weights $\mathbf{x}$.

- We then update our portfolio weights:

$$\mathbf{x}_{\text{new}} = \mathbf{x}_{\text{current}} + \Delta\mathbf{x}$$

- This process is repeated iteratively. In each iteration, SLSQP forms a new QP subproblem based on the current solution, solves it, and updates the solution. This continues until we reach a point where further changes in the weights no longer significantly improve the Sharpe Ratio, or until some other convergence criteria are met.

The optimization was performed using the `scipy.optimize.minimize(...)` function from the SciPy library, with the `method= 'SLSQP'` specified to employ the Sequential Least Squares Programming algorithm. Gradients were computed internally using finite differences. Finite differences are a way to approximate the derivative of a function using the function's values at nearby points. For example, a simple approximation of the derivative of a function f at a point x is given by:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

Where h is a small step size. The final solution of this optimization process is the weight vector $\mathbf{w}^*$ that maximizes the portfolio's Sharpe Ratio, representing the optimal allocation of investments across the selected stocks.

6 Results

We then evaluated the effectiveness of this optimal allocation by testing it on unseen data. Specifically, we implemented a two-phase approach: a 3-year optimization period followed by a 2-year evaluation period. Our hypothetical initial investment amount was set at Rs.10,000, and we assumed a risk appetite threshold of 15%.

6.1 Implementation

Our approach involved the following key steps, inspired in part by techniques detailed in Chapter 13 of Hilpisch's work on portfolio optimization [4]:

- **Optimization Period (3-Year Data):** We used three years of historical stock data to train our optimization model. This dataset allowed us to estimate expected returns, volatility, and correlations among the selected stocks. We then employed the SLSQP algorithm to determine the portfolio weights that would maximize the Sharpe Ratio, defined as:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

  where R_p is the portfolio return, R_f is the risk-free rate, and σ_p is the portfolio volatility.

- **Optimal Portfolio:** The optimization process yielded a specific set of portfolio weights that represented our ideal allocation based on the 3-year dataset. These weights were expected to offer the best risk-adjusted return for that period.

- **Evaluation Period (2-Year Data):** To test the robustness and out-of-sample performance of our optimized portfolio, we evaluated it over a subsequent 2-year period. This period was entirely separate from the optimization phase and served as an independent validation set.

- **Equal-Weighted Portfolio:** For benchmarking, we constructed an equal-weighted portfolio using the same set of stocks. This naïve allocation allowed us to compare the performance of our optimized portfolio against a basic strategy with no optimization.

6.2 Performance Metrics

To assess performance comprehensively, we used the following additional metrics:

Sortino Ratio: This metric focuses on downside risk, which refers to the risk of negative returns. It modifies the Sharpe Ratio by distinguishing harmful volatility (losses) from overall volatility. A higher Sortino Ratio indicates better risk-adjusted performance, particularly in terms of loss management. It is given by:

$$\text{Sortino Ratio} = \frac{R_p - R_f}{\sigma_d}$$

Where:

- σ_d is the annualized standard deviation of negative returns.

The value of σ_d is computed through the following steps:

1. Calculate the daily (or periodic) returns of the portfolio.
2. Identify returns that are negative (i.e., less than zero).
3. Compute the standard deviation of these negative returns.
4. Annualize this value by multiplying it with the square root of the number of trading periods in a year (e.g., $\sqrt{252}$ for daily returns).

The Sortino Ratio is particularly useful for investors focused on minimizing losses. By targeting downside risk, it provides a clearer view of how the portfolio performs during unfavorable market conditions.

Maximum Drawdown (MDD): MDD measures the largest peak-to-trough decline in portfolio value over a specified period. It represents the worst single loss sustained during the investment horizon. A lower MDD implies reduced exposure to significant losses and improved portfolio stability.
To compute MDD:

1. Calculate the cumulative return of the portfolio over time.
2. Identify the peak value of cumulative return up to each point.
3. Compute the drawdown at each point using:

$$\text{Drawdown} = \frac{\text{Peak Value} - \text{Current Value}}{\text{Peak Value}}$$

4. The Maximum Drawdown is the most negative value among all computed drawdowns.

MDD is critical for understanding potential downside and evaluating the portfolio's resilience. It is especially relevant for risk-averse investors concerned with the extent of potential losses in adverse market scenarios.

6.3 Optimal Portfolio Allocation

The following pie chart presents the final optimized portfolio allocation based on the 3-year training dataset. The total investment capital is assumed to be Rs. 10,000.

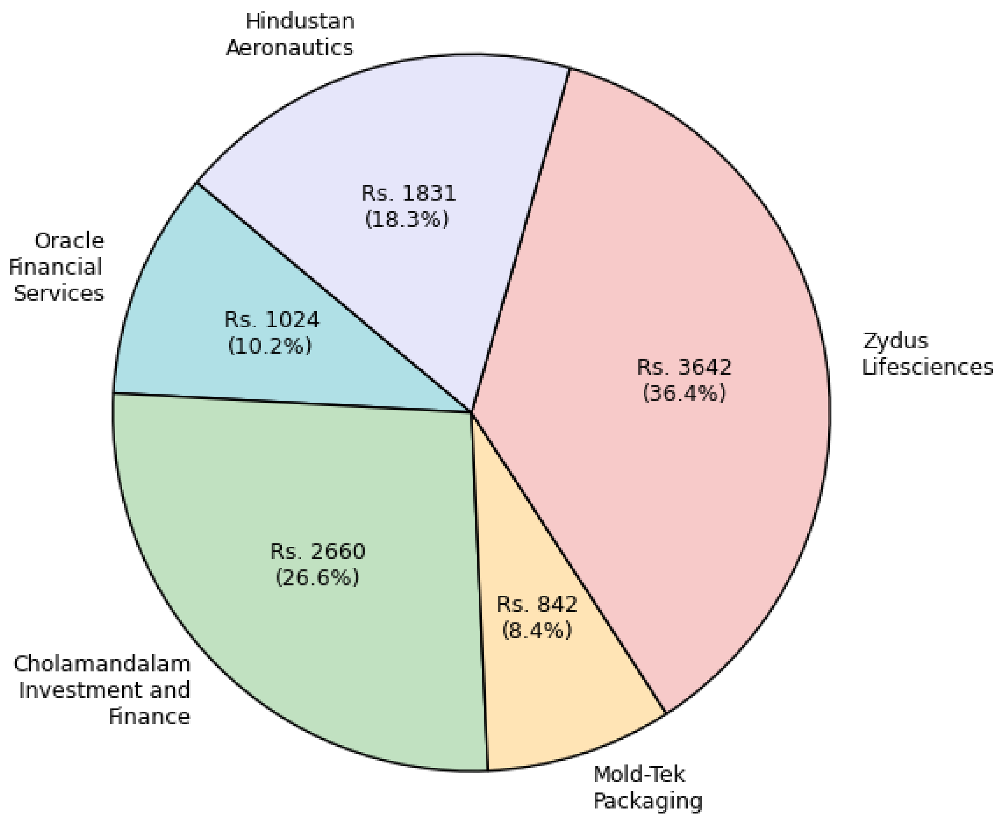

Figure 3: Optimized Portfolio Allocation

The capital distribution across the five selected stocks in the pie chart aligns well with the earlier risk-return scatter plot analysis. In that plot, we observed that:

- **Zydus Lifesciences** had one of the highest annualized returns with a relatively low volatility, making it an ideal candidate for a large allocation — which is exactly what we see with the 36.4% weight. Its strong position in the return spectrum made it a key driver for maximizing the Sharpe Ratio.
- **Cholamandalam Investment and Finance** was also placed favorably, offering a solid return while keeping volatility in check. Its 26.6% allocation reflects its role as a robust, stable performer, contributing significantly to the risk-return efficiency of the portfolio.

- **Hindustan Aeronautics**, though slightly higher in volatility, exhibited consistent returns and a relatively low correlation with other assets, justifying its 18.3% weight as a diversifier with good standalone performance.

- **Oracle Financial Services** and **Mold-Tek Packaging**, while on the more conservative end of the spectrum, showed lower volatility and marginal correlations with others, making them valuable for risk balancing. Their smaller allocations (10.2% and 8.4%, respectively) reflect a strategic inclusion aimed at diversification rather than return maximization.

This allocation, therefore, is not arbitrary - it stems from a balance between maximizing returns, minimizing risk, and reducing inter-asset correlation. The optimizer has effectively translated the insights from the risk-return landscape into a concrete, diversified investment strategy.

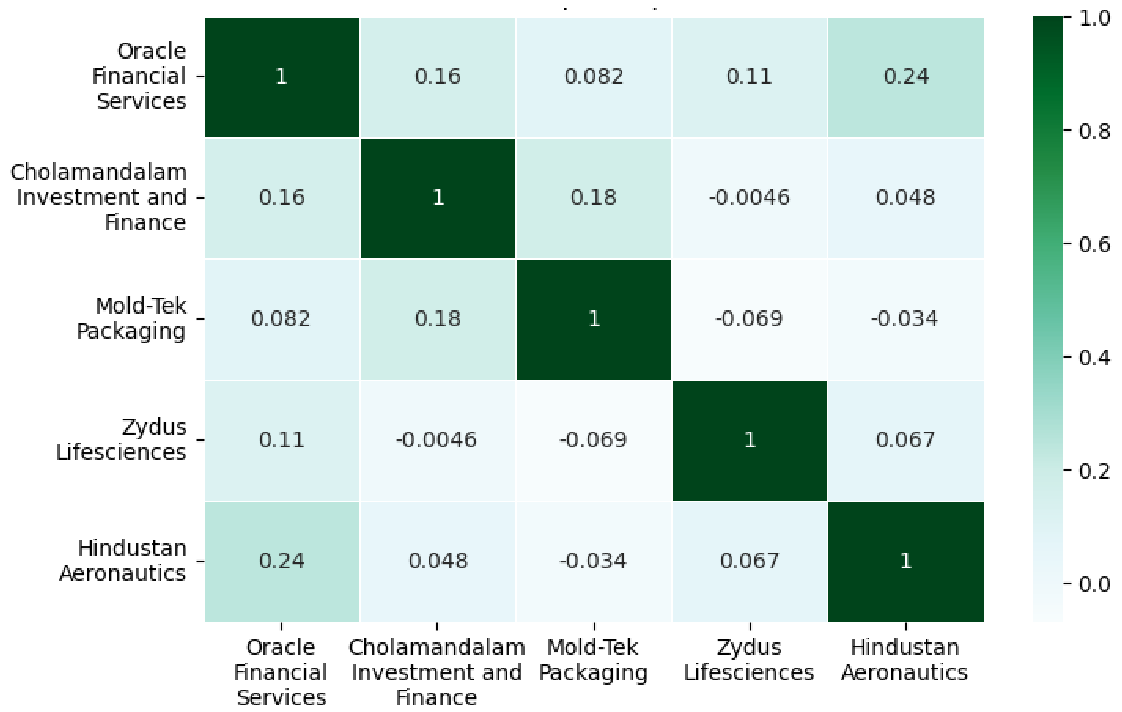

Figure 4: Correlation Matrix Heatmap of Selected Stocks

The heatmap displays the pairwise correlation between the returns of the five selected stocks in the optimized portfolio.

- **Low Correlations:** Most of the correlation values are relatively low, with a few even slightly negative (e.g., Mold-Tek Packaging and Zydus Lifesciences). This is ideal for diversification, as lower correlations help in reducing overall portfolio risk.

- **Highest Correlation:** The strongest positive relationship is between Oracle Financial Services and Hindustan Aeronautics (0.24), indicating some degree of co-movement.
- **Weak Relationships:** Several pairs (e.g., Cholamandalam Investment and Zydus Lifesciences) show correlations close to zero, suggesting that their price movements are largely independent.
- **Diversification Benefit:** This spread of correlations reinforces the strength of the optimization strategy, which aimed to build a portfolio with minimally correlated assets to maximize the Sharpe Ratio.

6.4 Portfolio Performance & Analysis

The table below summarizes the performance metrics of our optimized portfolio across both the training and evaluation periods, along with a comparison to an equal-weighted baseline.

Metric	3-Year Data (Training)	2-Year Data (Optimized)	2-Year Data (Equal Weight)
Expected Return	35.43%	36.63%	25.12%
Volatility	13.99%	20.15%	14.24%
Sharpe Ratio	2.05	1.49	1.29
Sortino Ratio	3.9	2.53	2.05
Maximum Drawdown	-14.51%	-58.55%	-46.68%

Figure 5: Portfolio Performance Metrics

Our optimized portfolio showed strong results over the 3-year training period, achieving a Sharpe Ratio of 2.05 and a Sortino Ratio of 3.90, with a relatively low volatility of 13.99% and a maximum drawdown of -14.51%. However, during the 2-year evaluation period, performance declined: the Sharpe Ratio dropped to 1.49 and the Sortino Ratio to 2.53, alongside a significant increase in volatility to 20.15% and a much deeper drawdown of -58.55%.

Interestingly, the expected return in the evaluation period (36.63%) was slightly higher than in training (35.43%), but it came at the cost of considerably higher risk. Notably, even this higher volatility remained within the investor's acceptable risk level. This suggests that while the model identified a promising portfolio based on past data, its future performance was affected by broader market dynamics.

It's important to note that market conditions have been turbulent over the past year, with many sectors seeing heightened volatility and corrections. This environment likely contributed to the dip in performance during the evaluation phase.

Still, our optimized portfolio outperformed the basic equal-weighted alternative across all key metrics, offering higher returns and better risk-adjusted performance. This shows that our model-driven strategy is a valuable tool for portfolio construction, especially compared to simpler, unoptimized approaches.

6.5 User Interface

To enhance the accessibility and usability of the portfolio optimization tool, a user-friendly website interface was developed. This interface allows investors to input their investment budget and risk appetite, tailoring the portfolio recommendations to their specific financial situation and preferences.

Upon clicking the "Optimize Portfolio" button, the website provides the following outputs:

- **Optimal Asset Allocation:** The website displays the recommended allocation of the investment budget across the top 5 selected companies. Both the percentage allocation and the corresponding investment amount for each company are presented, enabling users to easily implement the suggested portfolio strategy. A pie chart is also included for a clear visual representation of the allocation breakdown.
- **Portfolio Performance Metrics:** To help investors understand the potential performance of the optimized portfolio, we also display key performance metrics — Expected Return, Volatility, Sharpe Ratio, and Sortino Ratio— derived from the 3-year optimization data. One-line explanations of each metric are provided on the website to help users interpret the results and understand the potential performance characteristics of the recommended portfolio.

7 Conclusion

In conclusion, this project delivers a practical and easy-to-use tool for retail investors navigating the complexities of the Indian equity market. While the framework is relatively basic—using fixed sectoral constraints and standard optimization techniques—it serves as a strong starting point for more advanced, customizable investment solutions. Despite its simplicity, the model demonstrates decent performance across both training and test periods, offering competitive risk-adjusted returns. By combining sectoral diversification, user-defined risk filtering, and Sharpe ratio optimization, the system provides a structured yet accessible approach to portfolio construction. In an age of overwhelming financial information, this project shows how even a straightforward, data-driven model can empower better decision-making for individual investors.

7.1 Limitations

Despite the effectiveness of our model, we identified several limitations and challenges while working on it and testing it.

- Requiring exactly one stock per sector is a rigid constraint; ideally, sectoral diversification should emerge naturally through the optimization process rather than being enforced.
- Reliance on historical data assumes that past performance patterns will continue, which may not always be true in dynamic or volatile market. The model does not account for sudden or unforeseen market events such as political instability, environmental disasters, or macroeconomic shocks, which can significantly affect stock performance.
- The filtering step based on a strict risk threshold may eliminate all possible combinations in certain scenarios; a more flexible or adaptive filtering mechanism would better accommodate varying user risk levels.
- The tool does not currently support rebalancing or dynamic portfolio adjustment, which are crucial in real-world investment strategies for maintaining optimal allocation over time.
- Transaction costs, taxes, and liquidity constraints are not factored into the optimization process, which may affect the practical returns an investor would actually receive.

7.2 Future Work:

While our system gave us promising results, there are several areas where we could improve and enhance its performance and user experience. Below are some potential directions for further development:

- Flexible Sector Allocation & Stock Universe Expansion: : Future versions can move beyond the rigid “one stock per sector” rule by allowing dynamic sector allocation based on market signals and user preferences. Expanding the stock pool to include a wider range of companies across sectors and market caps would further enhance portfolio diversity and real-world applicability.
- Event-Responsive & Adaptive Risk Modelling: Incorporating mechanisms to account for sudden market shifts—such as political, economic, or environmental events—would improve model robustness. Additionally, replacing the fixed volatility filter with a more adaptive, probabilistic risk filtering approach could better capture user-defined risk tolerance without eliminating too many viable portfolios.
- Real-World Constraints & Intelligent Rebalancing: Introducing elements like transaction costs, taxes, and liquidity constraints would increase practical relevance. Adding features for periodic rebalancing based on market conditions or user adjustments can help maintain portfolio efficiency over time.

References

[1] W. F. Sharpe, "A linear programming algorithm for mutual fund portfolio selection," *Management Science*, vol. 13, no. 7, pp. 499–510, 1967. doi: 10.1287/mnsc.13.7.499

[2] S. M. Lee and A. J. Lerro, "Optimizing the portfolio selection for mutual funds," *The Journal of Finance*, vol. 28, no. 5, pp. 1087–1101, 1973. doi: 10.2307/2978750

[3] F. Abdi, S. Abolmakarem, A. Karbassi Yazdi, Y. Tan, and I. A. M. Choque, "Prospective portfolio optimization with asset preselection using a combination of long and short term memory and Sharpe ratio maximization," *IEEE Access*, vol. 12, pp. 144280–144294, 2024. doi: 10.1109/ACCESS.2024.3466829

[4] Y. Hilpisch, *Python for Finance*, 2nd ed., O'Reilly Media, 2018. [Chapter 13 on portfolio optimization].