

## HYBRID PORTFOLIO OPTIMIZATION WITH PCA, CLUSTERING, AND THE BARZILAI-BORWEIN METHOD

Kai Shin Chiew<sup>1</sup>, Wei Yeing Pan<sup>2\*</sup>, Hong Seng Sim<sup>3</sup>, Jia Hou Chin<sup>4</sup> and Yap Jia Lee<sup>5</sup>

<sup>1,2\*,3,5</sup>Department of Mathematics and Actuarial Science, Lee Kong Chian  
Faculty of Science, University Tunku Abdul Rahman, Sungai Long  
Campus, Malaysia.

<sup>4</sup>School of Accounting and Finance, Faculty of Business and Law,  
Taylor's University, 47500 Subang Jaya, Selangor, Malaysia.

<sup>1</sup>yannis25@utar.my, <sup>2\*</sup>panwy@utar.edu.my, <sup>3</sup>simhs@utar.edu.my,  
<sup>4</sup>jiahou.chin@taylors.edu.my, <sup>5</sup>yjlee@utar.edu.my

(\* indicates the corresponding author)

### ABSTRACT

Portfolio optimization aims to balance risk and return by identifying an effective mix of assets. In this study, we integrate principal component analysis (PCA) and hierarchical clustering for stock selection with the Barzilai–Borwein (BB) gradient method for portfolio optimization. Forty-eight U.S. stocks from the Kaggle fundamental stock dataset were initially collected, and 42 stocks were retained after preprocessing. Financial ratios from 2006 and adjusted closing prices from 2016–2017 were analysed, with one representative stock from each cluster selected using the Sharpe ratio. The BB method was then applied to determine optimal weights, ensuring full capital allocation without short selling. Among the tested approaches, the Barzilai–Borwein gradient method 1 (BB1) step size achieved strong performance, producing an annual return of 25.6% while maintaining relatively low volatility. The portfolio also generated a Jensen's alpha of 1.55, confirming the presence of positive abnormal returns beyond market expectations. These results suggest that combining PCA-based clustering with the BB optimization method offers a practical and efficient way to construct diversified portfolios. The study highlights the BB algorithm's potential as a lightweight yet effective alternative to more complex optimization techniques in financial decision-making.

**Keywords:** Barzilai-Borwein Gradient Method, Hierarchical Clustering, Portfolio Optimization, Principal Component Analysis

Received for review: 15-12-2025; Accepted: 25-03-2026; Published: 01-04-2026

DOI: 10.24191/mjoc.vol11i1.9078

### 1. Introduction

Portfolio optimization is an approach to selecting the ideal mix of assets and proportions to reduce risk and maximize the return on investment (Kalayci *et al.*, 2019). The concepts of portfolio optimization have been critical to making financial decisions and understanding the trend of the financial market. The main purpose of investors is to maximize their expected return while facing the least risk in the investment. However, market inefficiencies such as calendar anomalies can distort return patterns and risk estimates, as observed in the Malaysian



This is an open access article under the CC BY-SA license  
(<https://creativecommons.org/licenses/by-sa/3.0/>).

stock market (Shariff & Yusof, 2021). The Markowitz model is the earliest modern portfolio theory, and it has laid down a foundation for the development of modern portfolio theory (Markowitz, 1952). Many efforts have been made since then to enhance this model by taking into account real-world circumstances and addressing model variants utilizing numerous ways based on varied data and performance indicators, in conjunction with advances in computational capacity and algorithmic improvements. Sharpe (1963), Lintner (1975), and Mossin (1966) improvised to a significant extent to overcome the Markowitz model's constraints.

In recent years, portfolio optimization models can be divided into two categories: meta-heuristic models and mathematical models (Zanjirdar, 2020). The Barzilai-Borwein (BB) gradient method is categorized as a mathematical model. It is a gradient-based approach that improves the efficiency of traditional gradient methods by using a modified step size (Barzilai & Borwein, 1988). Stock selection is a critical approach for investment management and should be used as the first phase in portfolio management to identify target stocks that warrant further investigation. The stock scoring mechanism is the core concept of stock selection, where it helps to assess the value of a stock (Yang *et al.*, 2019). Fundamental analysis and technical analysis are used in this study to determine the high-valued stock. Principal component analysis (PCA) has been widely used in portfolio management to reduce dimensionality and identify latent financial factors driving asset co-movements (Jolliffe & Cadima, 2016; Dhingra *et al.*, 2021; Tubtimkaew, 2024). By transforming correlated financial indicators into orthogonal components, PCA facilitates more stable clustering and portfolio diversification.

Recent developments in portfolio optimization highlight the importance of algorithmic choice in balancing convergence speed, computational efficiency, and portfolio performance. Traditional conjugate gradient methods, while effective for quadratic optimization, have been observed to produce conservative allocations that emphasize low-risk but correspondingly low-return portfolios, which are less suitable for aggressive growth investment strategies (Kalayci *et al.*, 2019; Costa *et al.*, 2022). Limited-memory quasi-Newton approaches, such as L-BFGS, are powerful but require greater storage and computational cost for curvature updates, which can be unnecessary when the problem structure is quadratic and convex. In contrast, the BB gradient method provides a computationally lightweight and rapidly convergent alternative. By dynamically adjusting step sizes based on two consecutive iterates (Barzilai & Borwein, 1988; Ou & Themelis, 2024), the BB method accelerates convergence compared to steepest descent while avoiding the memory and overhead requirements of quasi-Newton methods. These advantages make the BB algorithm particularly well-suited for portfolio optimization, ensuring efficient allocation of investment weights while maintaining full utilization of available funds without short selling (Crisci *et al.*, 2024; Zhou *et al.*, 2025).

Although many studies have applied PCA, clustering, or optimization techniques in portfolio management, these approaches are often studied separately. The integration of dimensionality reduction, clustering-based diversification, and efficient gradient optimization methods, such as the BB algorithm, remains limited. This study, therefore, proposes a hybrid portfolio optimization framework combining PCA, hierarchical clustering, and the BB method.

The remainder of the paper is structured as follows. In Section 2, we briefly introduce the modern portfolio theory proposed by Markowitz, the stock selection and risk management in portfolio development, and the optimization techniques implemented by researchers. Section 3 details our method, including the selection of stocks, the Barzilai-Borwein gradient method, and the evaluation of stock performance. A description of the data is presented in Section 4. The results of our study are given in Section 5. Section 6 provides the discussion and conclusion of our study.

## 2. Literature Review

Traditional portfolio theory (TPT) analyses the individual securities based on the risk and return of the investment. It was distinguished by a straightforward, non-systemic, and inadequately analytical approach to developing an ideal portfolio (Leković, 2021). According to Leković (2021), John Burr Williams, John Richard Hicks, and Dickson Hammond Leavens are the most prominent representatives in traditional portfolio theory. They have studied portfolio investment using TPT and pushed for simple diversification. However, TPT is a weak portfolio theory in performing portfolio selection.

Modern portfolio theory (MPT) contributes to performance quantification by describing the association between risk and return. MPT is a significant advancement in the field of finance and investment practice. It provides a theoretical framework for optimizing return and risk ratios and reacts one step ahead of TPT. This is because MPT transfers the focus from individual stocks to portfolio characteristics. Markowitz (1952) introduced the mean-variance portfolio optimization model, which is also known as the Markowitz model, to address the problems in traditional portfolio theory.

Stock selection plays a crucial role in portfolio construction as it enables investors to identify assets with favourable risk-return characteristics. In financial research, various quantitative techniques have been applied to improve stock selection and portfolio diversification. Among these, principal component analysis (PCA) has been widely used to reduce dimensionality and extract latent factors from large financial datasets. By transforming correlated variables into a smaller set of orthogonal components, PCA simplifies complex datasets while preserving most of the underlying information (Sáenz *et al.*, 2023; Mattera *et al.*, 2025). Recent studies have also combined PCA with machine learning models to enhance prediction accuracy and reduce data redundancy in financial markets, while evidence also shows that PCA performs effectively compared with other feature selection techniques such as RFE and LASSO in financial and machine learning applications (Wu & Zhang, 2023; Mojahid *et al.*, 2025). In addition, hybrid decision-making approaches, such as the VIKOR method with Z-number reasoning, have been applied to portfolio selection under uncertainty (Tarmudi *et al.*, 2024). Clustering techniques are frequently used alongside PCA to group similar assets and enhance diversification strategies in portfolio construction (Sáenz *et al.*, 2023; Mattera *et al.*, 2025). Furthermore, rotational techniques such as Varimax, Quartimax, and Promax are commonly applied in PCA to improve the interpretability of the extracted components (Sáenz *et al.*, 2023; Despois *et al.*, 2023). To evaluate the performance of selected stocks or portfolios, financial indicators such as the Sharpe ratio are widely used to measure risk-adjusted returns relative to a risk-free benchmark (Kan, 2024).

In terms of optimization techniques, steepest descent is a fundamental method for optimizing non-linear functions that involves iteratively updating variables in the negative gradient direction, gradually approaching the ideal value. Recent BB variants address step-size sensitivity and convergence stabilization through adaptive and proximal BB-like step rules (e.g., AdaBB and BB-like rules in proximal gradient methods) (Zhou *et al.*, 2025; Crisci *et al.*, 2024). Barzilai and Borwein (1988) proposed a two-point step size approach based on steepest descent for quicker convergence. The literature examines step length selection, including non-monotone line search, to ensure convergence in nonlinear optimization. Recent portfolio optimization strategies show that the conjugate gradient method generates low-risk, low-return investment portfolios, making it unsuitable for aggressive growth investments. Building on this, recent studies show that robust and distributionally robust formulations improve out-of-sample performance and risk control and offer a unifying theoretical foundation for ambiguity sets and tractable reformulations (Costa *et al.*, 2022; Fan *et al.*, 2024; Fransisca *et al.*, 2024; Kuhn *et al.*, 2025; Waga *et al.*, 2025).

Despite extensive research on portfolio optimization and feature selection techniques, few studies integrate PCA-based stock selection with fast gradient optimization methods such as the BB algorithm. This gap motivates the hybrid framework proposed in this study.

### 3. Methodology

The overall methodology of the study will be divided into two sections: first, the stock selection phase and second, the optimization phase. This study integrates stock selection with optimization to create an optimal portfolio. Principal component analysis (PCA) and hierarchical clustering are used to separate stocks based on financial features in the stock selection phase, while the Sharpe ratio serves as a benchmark to select the higher performance stocks. The Barzilai-Borwein (BB) method is employed for optimization due to its quicker convergence and ability to solve unconstrained optimization issues. The modified BB algorithm uses two step sizes, Barzilai–Borwein gradient method 1 (BB1) and Barzilai–Borwein gradient method 2 (BB2), to calculate projected stock weightage. BB1 approximates the inverse Hessian scaling and generally produces larger step sizes, while BB2 approximates Hessian scaling and tends to produce smaller, more conservative updates. These two step-size strategies help improve the convergence efficiency of the optimization process compared with the standard steepest descent method. Performance evaluation is conducted using expected return and investment risk compared with an equally weighted portfolio. Lastly, the abnormal return of the BB optimized portfolio is also calculated to measure the excess return of the portfolio as compared to the market return. Figure 1 summarizes the methodological framework adopted in this study. The workflow consists of four main stages: data preprocessing and feature normalization, dimensionality reduction through PCA, stock grouping using hierarchical clustering, and portfolio optimization using the BB gradient method. The optimized portfolio is subsequently evaluated using return, risk, and Jensen’s alpha metrics.

#### 3.1 Stock Selection

Principal component analysis (PCA) is applied to reduce the dimensionality of the financial ratio dataset and identify the most relevant variables contributing to stock selection. PCA identifies key financial ratios through eigenvalue decomposition of the covariance matrix, retaining components with eigenvalues greater than 1 that together explain at least 70% of the total variance, as confirmed by scree plots. To improve interpretability, Varimax rotation with Kaiser normalization is applied to maximize the variance of squared loadings until convergence (Sáenz et al., 2023; Despois et al., 2023).

The reduced dataset is then subjected to hierarchical clustering using Ward’s agglomerative method to group stocks with similar financial characteristics. The clustering results are visualized using a dendrogram and flattened using Python’s *fcluster* function to obtain cluster labels. Within each cluster, stocks with the highest Sharpe ratios are selected as representative assets for portfolio construction (Kan, 2024). The Sharpe ratio is calculated as shown in Equation (1):

$$\text{Sharpe ratio} = \frac{E(R_p - R_f)}{\sigma_p} \quad (1)$$

where  $R_p$  is the return of portfolio,  $R_f$  is the risk-free rate, and  $\sigma_p$  is the standard deviation of excess return of the portfolio.

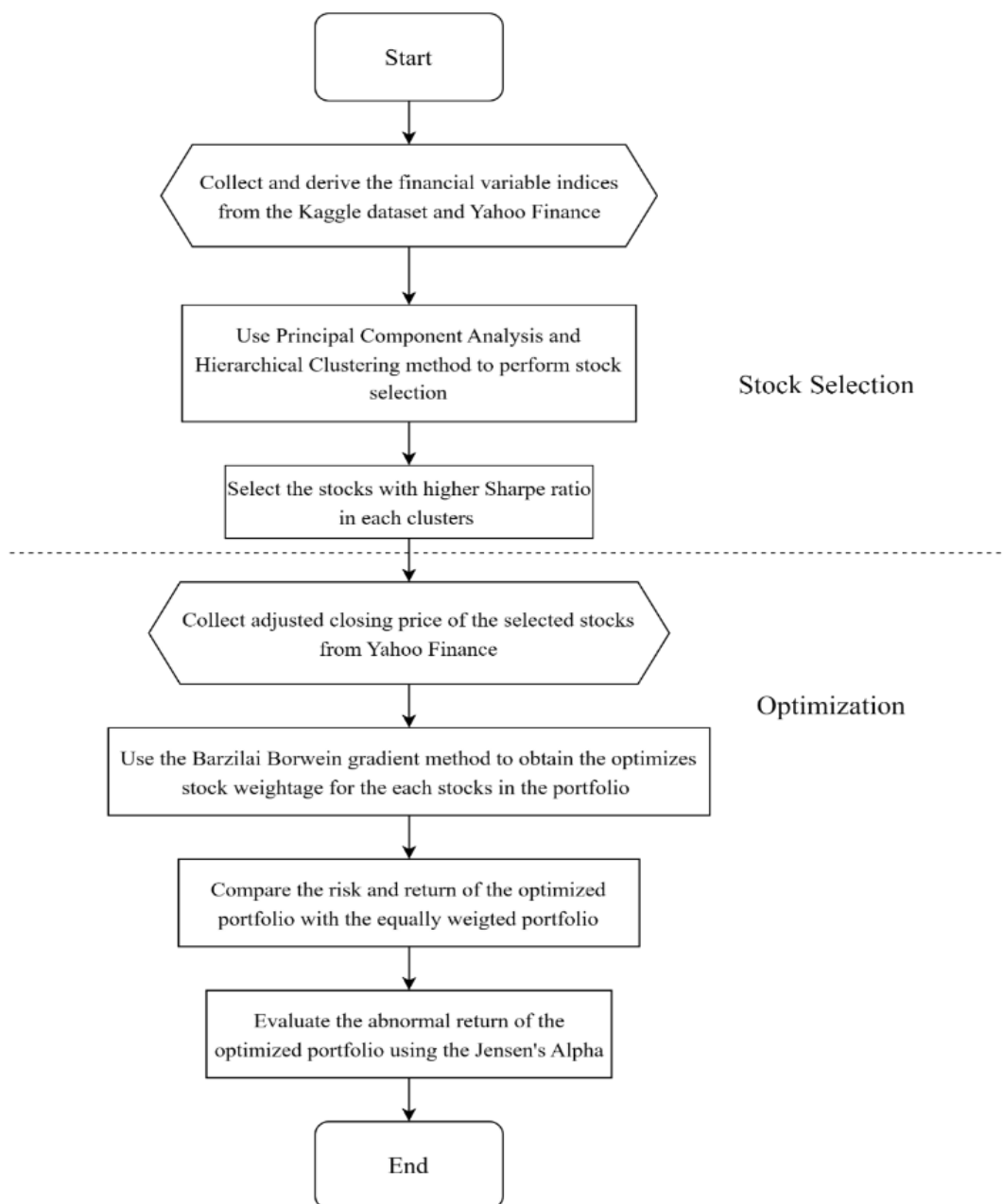


Figure 1. Hybrid Portfolio Optimization Process Using PCA and BB Method

### 3.2 Barzilai-Borwein Gradient Method

The Barzilai-Borwein (BB) method is an iterative gradient descent method using either one of the two step sizes derived from the objective function of the previous iteration. The method was introduced by Barzilai and Borwein in 1988 and significantly improves the performance of the classic gradient method. This study utilizes the BB method due to its significantly faster convergence compared to the steepest descent method, and its ability to effectively optimize large-scale unconstrained optimization problems.

In this study, an unconstrained minimization problem is considered as follows:

$$\min_{x \in R^n} f(x) \tag{2}$$

where  $f: R^n \rightarrow R$ . The updating formula to solve the problem in Equation (2) is defined by

$$x_{k+1} = x_k - \alpha_k g_k \tag{3}$$

where  $x_k$  is the weightage of stock in the current iteration,  $x_{k+1}$  is the weightage of stock in the next iteration,  $\alpha_k$  is the step size in scalar, and  $g_k$  is the gradient of  $f(x)$ .

Let  $L$  be the Lipschitz constant for the function  $f(x)$ ,  $\beta_k^{BB1}$  is the BB1 parameter, and  $\beta_k^{BB2}$  as the parameter for BB2,

$$\beta_k^{BB1} = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} \tag{4}$$

and

$$\beta_k^{BB2} = \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}} \tag{5}$$

where  $s_{k-1} = x_k - x_{k-1}$  and  $y_{k-1} = g_k - g_{k-1}$ . The BB parameters and  $L$  are incorporated into Equation (3) to form the new updating formula as shown below:

$$x_{k+1} = x_k - L\beta_k^{BB1} g_k \tag{6}$$

$$x_{k+1} = x_k - L\beta_k^{BB2} g_k. \tag{7}$$

Let  $A \in R^{n \times n}$  be the covariance return matrix, symmetric and positive definite, and let  $b \in R^n$  be the average return vector. We follow the standard quadratic programming formulation of mean–variance optimization (Griva *et al.*, 2009). The simplest form of the mean–variance optimization problem can be formulated as

$$\min_{x \in R^n} f(x) = \min_{x \in R^n} \frac{1}{2} x^T A x - b^T x \tag{8}$$

subject to  $e^T x = 1$ .

To handle the equality constraint efficiently, we adopt the Augmented Lagrangian approach (Griva *et al.*, 2009), leading to Equation (8), we obtain

$$L(x) = \frac{1}{2} x^T A x - b^T x + \lambda(e^T x - 1) + \frac{\delta}{2}(e^T x - 1)^2 \tag{9}$$

where

$x$  = the weightage of the stocks in matrix

$e$  = vector with all components equal to 1

$\lambda, \delta$  = Lagrange multipliers

The Lagrangian function was differentiated to obtain the gradient function. The general gradient function for all stocks is shown in Equation (10).

$$g_k = \frac{dL}{dx_k} = Ax_k - b + \lambda e + \delta(e^T x_k - 1)e \tag{10}$$

where  $k$  be the iteration number,  $k \geq 1$ . The initial step size was set to be 0.01, and the Lagrange multipliers  $\lambda$  and  $\delta$  were set to be 1 and 90, respectively. The penalty parameters were chosen following the standard augmented Lagrangian approach suggested by Griva *et al.* (2009). The initial Lagrange multiplier was set to  $\lambda = 1$  as a standard initialization value, while the penalty parameter  $\delta = 90$  was selected based on preliminary experiments that ensured stable convergence of the BB optimization algorithm without oscillation.

Lipschitz continuity is a feature in optimisation that is utilised to ensure the convergence of the BB optimisation method. A function is proven to be Lipschitz continuous

if there exists a real constant,  $L \geq 0$ , such that the following inequality holds for every  $x$  and  $y$  in the function's domain  $\|g(x) - g(y)\| \leq L \|x - y\|$ . The constant,  $L$  is known as the Lipschitz constant, and it is chosen in a way to provide a good estimation of step size. Lipschitz continuity carries uniform continuity that is particularly strong for functions. To identify a Lipschitz continuous function, the function is limited in the speed of change when there exists a real number.

Assume there are two points,  $x$  and  $y$  in the objective function. The derived process to determine the Lipschitz constant of the objective function, in Equation (10), is shown as below:

$$\begin{aligned} \|L'(x) - L'(y)\| &= \|Ax - b + \lambda e + \delta(e^T x - 1)e\| - \|Ay - b + \lambda e + \delta(e^T y - 1)e\| \\ &= \|A(x - y) + \delta e e^T (x - y)\| \end{aligned}$$

then

$$\begin{aligned} \|L'(x) - L'(y)\| &\leq \|A + \delta e e^T\| \|x - y\| \\ &\leq \|A\|_2 + \delta \|e e^T\|_2 \|x - y\| \end{aligned}$$

or it can be written as

$$\|L'(x) - L'(y)\| \leq \sqrt{\text{tr}(AA^T)} + \delta n \cdot \|x - y\|.$$

When compared with Lipschitz inequality,  $L = \sqrt{\text{tr}(AA^T)} + \delta n$ . The step size for Equations (6) and (7) is chosen based on  $\min(1, 1/L)$ . The multiplier  $\lambda$  was updated in each iteration using the equation below:

$$\lambda_{k+1} = \lambda_k + \delta(e^T x - 1).$$

The objective function gradually converged to obtain the minimum weight of the stock, and the iteration stops when the norm of the gradient is less than the tolerance rate of  $1 \times e^{-6}$ .

### 3.3 Performance Evaluation

The performance of an optimized portfolio is measured in terms of projected return and risk when compared to an equally weighted portfolio. The BB gradient approach predicts ideal weights for stocks in the US market based on previous year returns, and the predicted return is calculated by multiplying the optimized stock weights by the next year's annualized average returns. Risk is assessed as the standard deviation from the anticipated return, with smaller values indicating reduced portfolio risk. If the optimized portfolio fails to perform well, more algorithm limitations are applied to improve performance.

The convergence of the BB technique is determined by iteration count and gradient convergence to zero, indicating optimal weights and low risk. Jensen's alpha, derived from the CAPM framework (Jensen, 1968), measures the abnormal return of a portfolio relative to its expected return given its systematic risk. Positive alpha shows outperformance after accounting for systematic risk, negative alpha indicates underperformance due to market swings or optimization effects, and zero alpha represents performance that is consistent with market expectations. Jensen's alpha is determined by combining the portfolio's beta and market return, reflecting differences between actual and predicted returns. Positive, negative, or zero alpha indicates whether the portfolio outperforms, underperforms, or meets market expectations. The equation for Jensen's alpha computation is shown in Equation (11),

$$\text{Jensen's alpha} = R_p - [R_f + \text{Beta} \times (R_m - R_f)] \tag{11}$$

where  $R_p$  indicates the return of portfolio,  $R_f$  indicates the risk-free rate,  $R_m$  indicates the expected market return and  $\text{Beta}$  represents the risk of the portfolio.

#### 4. Data Description

This study used a two-pronged approach to stock selection, combining both fundamental and technical analysis. The initial dataset contained 48 stocks listed in the United States stock market. The dataset is archived and available at: <https://www.kaggle.com/datasets/artemburenok/fundamental-stock-data>. After preprocessing, which included removing stocks with missing or inconsistent financial records and normalizing features, 42 stocks with complete information were retained for further analysis.

Fundamental analysis determined the underlying value of each stock by investigating a company's financial condition and asset holdings. This strategy sought to find inexpensive equities with a likelihood of future price appreciation. Technical analysis focuses on historical price movements and market trends. The purpose of analyzing historical price data was to measure the opinion of investors while discovering potentially profitable points of entry and withdrawal for selected equities.

Tables 1 and 2 present the financial characteristics of these 42 stocks. Table 1 depicts six of the fundamental features, which are the price-to-sales ratio (P/S), operating margin, return on equity, sales per share trailing twelve months (ttm), debt-to-equity ratio, and price-to-earnings ratio (P/E). On the other hand, Table 2 illustrates the technical features of the 42 stocks. There are three technical characteristics considered, which are market capitalization and price evolution between the years 2006 and 2007. The price evolution is measured by the ratio between the minimum price during the previous year in 2006 and the current price in 2007, ( $P_{min}/P$ ), and the ratio between the maximum price during the previous year in 2006 and the current price in 2007, ( $P_{max}/P$ ). All of the financial ratios are based on the historical data from the year 2006.

Tables 1 and 2 show substantial variability in the financial indicators. For instance, the P/E ratio ranges from  $-43.39$  (CVX) to  $291.45$  (AMZN), while return on equity varies from  $-57.74$  (PM) to  $214.49$  (UPS). In addition, the P/S ratio ranges from  $0.39$  (WMT) to  $13.86$  (V). Such differences in scale may bias statistical analysis; therefore, the financial variables are standardized prior to applying PCA. Scaling normalizes the range of characteristics, positioning them on a comparable scale. The process involves converting the characteristics to unit variance and zero mean. This guarantees that all characteristics contribute equally to distance computations, resulting in fairer comparisons and potentially boosting model accuracy and efficiency. Data preparation includes methods for dealing with missing values, inconsistencies, and irrelevant characteristics. This procedure ensures that the data is of high quality and appropriate for machine learning models such as PCA and clustering techniques by cleaning and converting the dataset. Missing data values indicate an omission of information, which might lead to biased or erroneous conclusions. The dataset is checked for missing values before proceeding to modeling algorithms. The Kaiser-Meyer-Olkin (KMO) test measures data sampling adequacy, ensuring an adequate sample size for PCA. A KMO value greater than  $0.5$  indicates sufficient sampling for future analysis. A lower value of the KMO measure indicates the need for additional data or more thorough preprocessing to enhance the dataset. The Bartlett's Sphericity test evaluates the correlation between financial characteristics, assessing if the correlation matrix has an identity matrix or if features have no correlation. The null hypothesis must be rejected by having a  $p$ -value lower than  $0.05$  to show the results are significant and prove the dataset is correlated.

Table 1. Fundamental Financial Indices for All 42 Stocks

Ticker	P/S ratio	Operating Margin	Return on Equity	Sales per share (ttm)	Debt-to-Equity ratio	P/E ratio
AAPL	2.45	0.28	39.10	9.95	1.39	11.09
ABBV	3.18	0.36	120.91	14.22	10.80	14.54
ABT	2.72	0.15	12.14	13.71	0.91	25.35
ACN	1.97	0.14	53.42	50.89	1.54	17.13
AMZN	2.86	0.03	10.26	12.36	3.22	291.45
AVGO	4.20	-0.03	6.86	29.79	1.21	23.47
AZN	2.62	0.14	15.12	9.62	2.89	24.65
BMJ	5.20	0.24	15.72	10.34	1.18	42.26
CMCSA	1.78	0.21	15.31	15.30	2.13	16.40
COP	1.66	-0.37	-14.97	21.48	1.52	-7.43
CSCO	2.37	0.26	17.08	9.67	0.90	11.19
CVX	1.14	-0.04	0.55	65.45	0.75	-43.39
DHR	3.73	0.15	13.47	19.32	0.96	16.01
DIS	2.83	0.26	18.96	33.02	0.90	16.92
GOOGL	7.11	0.26	14.42	5.38	0.21	30.59
HD	1.57	0.14	102.13	71.36	5.90	19.42
JNJ	3.78	0.27	21.63	25.20	0.91	17.38
KO	3.62	0.21	27.85	9.89	2.55	21.30
LIN	2.73	0.22	29.97	37.84	2.63	18.70
LLY	3.60	0.14	16.16	19.17	1.41	30.49
MA	10.27	0.52	63.58	8.97	1.78	26.84
MCD	3.70	0.31	99.54	27.51	10.67	19.61
MRK	3.42	0.21	10.97	13.52	1.24	22.84
MSFT	4.49	0.29	21.12	11.15	1.62	30.85
NEE	2.76	0.28	11.47	9.27	2.57	18.02
NVDA	4.38	0.22	18.28	2.59	0.69	28.02
ORCL	4.07	0.34	18.92	8.74	1.39	16.91
PEP	2.04	0.16	43.38	42.29	4.93	22.76
PFE	2.95	0.23	10.70	8.26	1.67	19.68
PG	3.00	0.21	15.47	23.23	1.15	20.63
PM	1.43	0.14	-57.74	47.59	-4.20	16.06
RTX	1.32	0.12	24.20	65.21	1.96	9.98
SCHW	5.34	0.44	10.97	5.13	12.40	24.08
TMO	3.32	0.13	9.47	43.69	1.04	29.18
TMUS	1.02	0.12	6.44	41.19	2.68	34.06
TTE	0.58	0.01	4.12	55.46	1.28	19.78
TXN	4.13	0.36	31.45	12.55	0.60	17.53
UPS	1.30	0.13	214.49	65.77	13.87	15.56
V	13.86	0.52	21.02	5.37	0.91	29.02
VZ	1.21	0.22	96.25	31.89	10.97	9.95
XOM	1.17	0.05	6.83	53.42	0.92	22.79
WMT	0.39	0.05	17.98	151.35	1.50	12.98

Table 2. Technical Financial Indices for All 42 Stocks

Ticker	Technical Features			Ticker	Technical Features		
	Market Capitalization	Price Evolution			Market Capitalization	Price Evolution	
		$\frac{P_{min}}{P}$	$\frac{P_{max}}{P}$			$\frac{P_{min}}{P}$	$\frac{P_{max}}{P}$
AAPL	24.36	0.83	1.10	MCD	84.89	1.10	1.31
ABBV	46.31	1.12	1.50	MRK	46.04	0.99	1.35
ABT	37.27	1.05	1.34	MSFT	49.91	0.84	1.12
ACN	100.64	0.88	1.21	NEE	25.55	1.03	1.33
AMZN	35.23	0.63	1.12	NVDA	10.88	0.25	1.19
AVGO	152.11	0.81	1.30	ORCL	35.56	0.96	1.22
AZN	25.33	1.17	1.60	PEP	86.25	1.10	1.31
BMJ	53.89	1.03	1.63	PFE	24.34	1.13	1.50
CMCSA	27.18	0.89	1.21	PG	69.83	1.08	1.31
COP	35.68	0.77	1.32	PM	68.38	1.34	1.65
CSCO	22.86	0.92	1.31	RTX	85.94	0.90	1.20
CVX	74.17	0.87	1.37	SCHW	27.33	0.59	1.11
DHR	72.07	0.81	1.08	TMO	144.88	0.85	1.14
DIS	93.38	0.85	1.05	TMUS	41.82	0.57	1.02
GOOGL	38.14	0.83	1.04	TTE	31.95	1.17	1.54
HD	111.85	0.97	1.23	TXN	51.79	0.77	1.25
JNJ	95.28	0.99	1.32	UPS	85.77	0.96	1.32
KO	35.82	1.19	1.41	V	74.68	0.87	1.11
LIN	103.1	0.93	1.21	VZ	38.50	1.15	1.50
LLY	69.08	0.98	1.31	XOM	62.16	1.11	1.48
MA	91.94	0.78	1.08	WMT	59.61	1.00	1.25

## 5. Results and Discussion

### 5.1 Results of the Principal Component Analysis

From Table 3, the KMO has a measure of 0.515, which is higher than the expected value of 0.5. Furthermore, Bartlett's test for Sphericity has a chi-squared value of 138.934. A large chi-squared value represents a high tendency for the null hypothesis to be rejected. With the p-value approximating zero and lower than the significant value of 0.05, the null hypothesis is rejected. The Bartlett's test is significant and indicates that the financial features matrix is correlated. The financial ratio data can proceed to PCA without further data cleaning. Data analysis techniques are utilized to identify similarities and differences across assets in the United States stock market. First, PCA is employed to minimize the number of variables describing the assets. Table 4 shows the findings achieved by PCA. It includes the eigenvalues of financial variables in the correlation matrix, the percentage of variance explained by each principal component (PC), and the respective total variance.

There are two categories of sums of squared loadings in Table 4. The initial eigenvalues category represents the variance explained by each PC before applying the rotation method. Meanwhile, the varimax rotation category is the variance explained for each PC after the varimax rotation. The percentage of variance explained rapidly declines after the fourth

principal component (PC4), with a cumulative variance of 79.58% when four PCs are used. This result exceeds the acceptable criterion of 70% total variance explained, demonstrating a strong representation of the original characteristics. Additionally, utilizing rotation sum of squared loadings improves interpretability greatly, highlighting the relevance of the first four PCs in capturing the majority of feature variation. The variance and proportion of total variance explained have both increased on average. The cumulative proportion of total variance explained by the top four PCs using rotation sum of squared loadings is similarly 79.58%. Nine principal components can theoretically represent original characteristics, but they introduce noise, potentially leading to erroneous predictions.

Table 3. KMO and Bartlett’s Test

Test		Value
Kaiser Meyer Olkin (KMO)	-	0.515
Bartlett’s Test for Sphericity	chi-squared	138.934
	p-value	< 0.001

Table 4. Principal Component Analysis Proportion of Variance Explained by Components

Component	Initial Eigenvalues			Varimax Rotation		
	Variance	% of total variance	Cumulative % of total variance	Variance	% of total variance	Cumulative % of total variance
1	2.4729	27.48%	27.48%	2.0409	22.68%	22.68%
2	1.9262	21.40%	48.88%	1.8572	20.64%	43.31%
3	1.5505	17.23%	66.11%	1.8629	20.70%	64.01%
4	1.2124	13.47%	79.58%	1.4009	15.57%	79.58%
5	0.7992	8.88%	88.46%			
6	0.4356	4.84%	93.30%			
7	0.2602	2.89%	96.19%			
8	0.1799	2.00%	98.19%			
9	0.1631	1.81%	100.00%			

The rotated component matrix in Table 5 displays the loadings of variables onto each of the components, with loadings greater than 0.8 deemed significant and loadings less than this level ignored. Seven of the first nine characteristics were chosen based on their high loadings, including the P/S ratio, operating margin, return on equity, debt-to-equity ratio, market capitalization, and the minimum and maximum price ratios to the current price. Values approaching 1 or -1 indicate a considerable influence on the component, while values around zero indicate a minor impact. Furthermore, negative loadings represent that the characteristics impact the component in an adverse manner. For example, the P/S ratio (0.8715) and operating margin (0.8715) show strong positive loadings on Component 1, indicating that firms with higher profitability and valuation ratios contribute strongly to this component. In contrast, sales per share (-0.6599) has a negative loading on Component 1, suggesting that higher sales per share move in the opposite direction of the component structure.

Based on the rotated loadings, Component 1 primarily reflects profitability and valuation characteristics, as indicated by the high positive loadings of the P/S ratio and operating margin. Component 2 captures financial leverage and profitability, represented by the strong loadings of return on equity and the debt-to-equity ratio. Component 3 reflects price dynamics relative to historical price ranges, indicated by the high loadings of  $P_{max}/P$  and

$P_{min}/P$ , while Component 4 is mainly associated with firm size, represented by the high loading of market capitalization.

Table 5. Rotated Component Matrix

Features	Component			
	1	2	3	4
P/S ratio	0.8715	-0.1034	-0.2311	0.0748
Operating Margin	0.8715	0.2371	0.0029	-0.0196
Return on Equity	0.0595	0.9325	0.03	0.1761
Sales per Share (ttm)	-0.6599	0.1221	0.0716	0.5636
Debt-to-Equity ratio	0.0003	0.9471	0.0117	-0.0829
P/E ratio	-0.1226	0.0425	-0.5058	-0.4869
Market Capitalization	-0.0497	0.0583	-0.1129	0.8617
$\frac{P_{max}}{P}$	-0.0907	0.0605	0.89	0.0793
$\frac{P_{min}}{P}$	-0.2392	0.0011	0.8617	-0.2312

### 5.2 Results of Hierarchical Clustering

Next, clustering algorithms are utilized to create classes of comparable assets. Figure 2 presents the dendrogram obtained using agglomerative hierarchical clustering with Ward’s linkage method. The dendrogram illustrates the similarity relationships among the 42 stocks based on the selected financial characteristics. Stocks that merge at lower linkage distances exhibit more similar financial profiles, while those that merge at higher distances are more dissimilar. By examining the dendrogram structure and selecting an appropriate cut-off level, 28 clusters were identified, ensuring that stocks within each cluster share similar financial attributes while maintaining diversification across clusters. The Sharpe ratio is then used to identify the best-performing stock within each cluster. Consequently, 28 representative stocks are selected and used as inputs for the BB portfolio optimization stage.

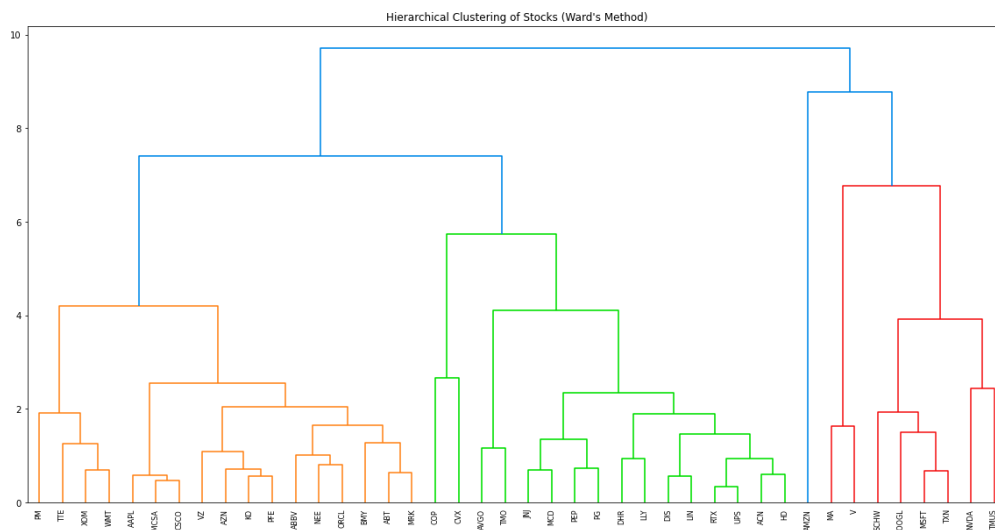


Figure 2. Dendrogram of Hierarchical Clustering

### 5.3 Results of Barzilai-Borwein Optimization

The BB algorithm optimized the stock weightage of 28 selected stocks based on the Sharpe ratio using historical adjusted closing prices from 2016. These optimal weightages were then used to calculate expected returns and risk for 2017. The BB1 method required 97 iterations to reach a near-zero minimum gradient, after which the gradient increased linearly. In contrast, the BB2 method converged after 11 iterations but failed to reach a near-zero minimum gradient before stopping, indicating that the solution had not fully stabilized. Consequently, the BB1 step size was selected for portfolio optimization as it demonstrated better convergence behaviour and produced a more reliable solution. Based on the stock weights optimized using BB1, all stocks contribute nearly equal weightage to the portfolio. The range of weightage among all the stocks is 0.00058%. Additionally, the total stock contributions in the portfolio sum up to 1 without the presence of short selling. This proves that the BB optimization method satisfies the study’s objectives of avoiding short selling while fully utilizing the available investment funds.

Table 6: Comparison of Portfolio Performance and Convergence Behaviour of BB1 and BB2 Methods

Method	Iterations	Convergence	Return	Risk
Equal-weight portfolio	-	-	25.85%	16.60%
BB1	97	Reached near-zero gradient	25.58%	0.58%
BB2	11	Did not reach near-zero gradient	-	-

Table 6 shows the comparison of portfolio performance and convergence behaviour between the equal-weight portfolio and the BB optimization methods. Although the BB2 method converged in fewer iterations, it did not reach a near-zero gradient and therefore did not produce stable portfolio weights. In contrast, BB1 achieved convergence and was therefore selected for portfolio optimization.

The BB1-optimized portfolio offers a return of 25.58% in a year, which is comparable to the 25.85% return of the equally weighted portfolio. However, as shown in Table 6, the volatility of the BB1-optimized portfolio is significantly lower at 0.58%, compared to 16.60% for the equally weighted portfolio. This reduction in volatility is largely attributable to the diversification effect across the selected stocks and the optimized allocation weights derived from the covariance structure of asset returns. This indicates that the BB optimization method is able to substantially reduce portfolio risk while maintaining competitive returns.

On the other hand, Jensen's alpha coefficient of 1.5571 indicates that the optimized portfolio has produced an anomalous return of 1.56% each period after adjusting for market risk. The calculation uses a risk-free rate of 2.33%, corresponding to the US 10-year Treasury yield in 2017, which is commonly used as a proxy for the risk-free rate in financial studies. The positive alpha indicates that the optimized portfolio outperformed the expected return predicted by the Capital Asset Pricing Model (CAPM), suggesting superior risk-adjusted performance.

In our experiments, the conjugate gradient approach was also tested as a baseline. While it successfully converged, the allocations were conservative, leading to portfolios characterized by relatively low risk but also limited returns. This indicates that conjugate gradient methods tend to produce stable but less aggressive solutions, which may not align with growth-oriented investment objectives. By contrast, the Barzilai-Borwein method demonstrated faster convergence and produced a more competitive return–risk profile, highlighting its practical advantage for portfolio optimization in this study.

## 6. Conclusion and Recommendation

The Barzilai-Borwein (BB) portfolio optimization approach in the US stock market has shown success in achieving higher returns and low risk. The optimization process starts with a diversified portfolio of high-performance stocks, resulting in significant risk reduction. The BB-optimized portfolio outperformed the market, demonstrating its effectiveness. Stock selection strategies, such as principal component analysis with varimax rotation and hierarchical clustering, were used to select high-performance companies. The BB approach is effective in fast convergence, distributing stocks without short selling, and fully utilizing investment funds. The study's current work has limitations in data collection for financial ratios and not incorporating financial expenses in the objective function. The BB2 optimization approach was unable to converge, and the optimized portfolio offered similar expected earnings to the equally weighted portfolio. Additionally, the study was limited to a single period in the US stock market, limiting the depth of examination. Future research should expand portfolio optimization strategies by using risk estimating tools like Value-at-Risk measure (VaR) for comprehensive risk assessment and selecting assets with the lowest VaR. Extending the research period and considering economic situations like downturns or volatility can provide insights into the effectiveness of the BB algorithm. Comparing BB with alternative approaches and considering real-world limitations like transaction costs and regulatory requirements can enhance its practicality and appeal in real-world investment scenarios.

## Acknowledgement

The authors thank the handling editor and anonymous reviewers for their constructive feedback.

## Funding

The author(s) received no specific funding for this work.

## Author Contribution

Author 1 conceived the study, conducted the literature review, developed the methodology, performed the analysis, and drafted the manuscript; Author 2 contributed to study design and critical review; Author 3 refined the optimization methodology and reviewed results; Author 4 curated data, computed indicators, and implemented PCA and hierarchical clustering; Author 5 implemented the optimization code, ran experiments, and prepared figures/tables. All authors read and approved the final manuscript.

## Conflict of Interest

The authors have no conflicts of interest to declare.

## References

- Barzilai, J., & Borwein, J. M. (1988). Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1), 141–148.
- Costa, G., De Angelis, V., & Ievlev, A. (2022). Data-driven distributionally robust risk-parity portfolio optimization. *Optimization Methods and Software*, 37(6), 1905–1938.
- Crisci, S., Rebegoldi, S., Toraldo, G., & Viola, M. (2024). Barzilai–Borwein-like rules in proximal gradient schemes for  $\ell_1$ -regularized problems. *Optimization Methods and Software*, 1–33.
- Despois, T., & Doz, C. (2023). Identifying and interpreting the factors in factor models via sparsity: Different approaches. *Journal of Applied Econometrics*, 38(4), 533–555. <https://doi.org/10.1002/jae.2967>
- Dhingra, V., Sharma, A., & Gupta, S. K. (2021). Sectoral portfolio optimization by judicious selection of financial ratios via PCA. *arXiv preprint arXiv:2106.11484*.
- Fan, Z., Ji, R., & Lejeune, M. A. (2024). Distributionally robust portfolio optimization under marginal and copula ambiguity. *Journal of Optimization Theory and Applications*, 203(2), 2870–2907. <https://doi.org/10.1007/s10957-024-02550-y>
- Fransisca, D. C., Sukono, Chaerani, D., & Halim, N. A. (2024). Robust mean-variance optimization for capital allocation: A systematic review. *Computation*, 12(8), 166. <https://doi.org/10.3390/computation12080166>
- Griva, I., Nash, S. G., & Sofer, A. (2009). *Linear and nonlinear optimization* (2nd ed.). Society for Industrial and Applied Mathematics (SIAM).
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, 23(2), 389–416.
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Kalayci, C. B., Ertenlice, O., & Akbay, M. A. (2019). A comprehensive review of deterministic models and applications for mean-variance portfolio optimization. *Expert Systems with Applications*, 125, 345–368.
- Kan, R. (2024). In-sample and out-of-sample Sharpe ratios of multi-factor strategies. *Journal of Financial Economics*, 155, 103837.
- Kuhn, D., Esfahani, P. M., Nguyen, V. A., & Shafieezadeh-Abadeh, S. (2025). Distributionally robust optimization. *Acta Numerica*, 34, 1–212.
- Leković, M. M. (2021). Historical development of portfolio theory. *Tehnika*, 76(2), 220–227.
- Lintner, J. (1975). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. In *Stochastic Optimization Models in Finance* (pp. 131–155). Academic Press.
- Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.

- Mattera, R., Di Nardo, A., & Sessa, S. (2025). Time-series clustering for portfolio selection: A comparative study. *Soft Computing*, 29, 4219–4231.
- Mojahid, H. Z., Zain, J. M., Yusoff, M., Basit, A., Jumaat, A. K., & Ali, M. (2025). Examining the impact of feature selection techniques on machine and deep learning models. *Malaysian Journal of Computing*, 10(1), 2135-2158. <https://doi.org/10.24191/mjoc.v10i1>
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 34(4), 768–783. <https://doi.org/10.2307/1910098>
- Ou, H., & Themelis, A. (2024). Safeguarding adaptive methods: global convergence of Barzilai-Borwein and other stepsize choices. *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2802–2807. <https://doi.org/10.1109/codit62066.2024.10708282>
- Sáenz, J. V., López-García, M., & Zuluaga, J. (2023). Using clustering models of stocks to improve price prediction and trading returns. *International Review of Financial Analysis*, 88, 102758.
- Shariff, N. S. M., & Yusof, N. A. (2021). Stock market anomalies: Calendar effects on the Malaysian stock market. *Malaysian Journal of Computing*, 6(1), 772–777. <https://doi.org/10.24191/mjoc.v6i1.11212>
- Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science*, 9(2), 277–293.
- Tarmudi, Z., Ahmad, S. N., Mohammad, S. A., Ghazali, A. F., Abd Rahman, M., & Triana, Y. S. (2024). VIKOR method with Z-number approach for portfolio selection decision. *Malaysian Journal of Computing*, 9(1), 1759–1767. <https://doi.org/10.24191/mjoc.v9i1.26052>
- Tubtimkaew, K. (2024). The application of principal component analysis for portfolio investment selection. *NIDA Business Journal*.
- Waga, M., Valladão, D., & Street, A. (2025). Robust portfolio optimization meets Arbitrage Pricing Theory. *European Journal of Operational Research*. <https://doi.org/10.1016/j.ejor.2025.04.004>
- Wu, S., & Zhang, L. (2023). Research on the impact of PCA-LSTM on stock price forecast. In *Proceedings of the 2023 7th International Conference on Deep Learning Technologies* (pp. 98–102).
- Yang, F., Chen, Z., Li, J., & Tang, L. (2019). A novel hybrid stock selection method with stock prediction. *Applied Soft Computing*, 80, 820–831.
- Zanjirdar, M. (2020). Overview of portfolio optimization models. *Advances in Mathematical Finance and Applications*, 5(4), 419–435.
- Zhou, D., Ma, S., & Yang, J. (2025). ADABB: Adaptive Barzilai-Borwein Method for Convex optimization. *Mathematics of Operations Research*. <https://doi.org/10.1287/moor.2024.0510>