Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.LD31777

Investment Portfolio Development: From Convex Modern Portfolio Theory to Genetic Algorithm Heuristic Modelling

Abdul Qadoos[1*], Hisham AbouGrad[1], Julie Wall[2], Mhd Saeed Sharif[1]

[1]*School of Architecture, Computing and Engineering, University of East London, London, UK*
[2]*School of Computing and Engineering, University of West London, London, UK*
**Corresponding Author. Email: a.qadoos@uel.ac.uk*

Abstract. Machine learning models have significantly progressed toward automation in many financial operations, including stock market prediction and investment portfolio development. Due to the volatility of the stock market, stock selection and asset allocation remain challenging for investors. The modern portfolio theory (MPT) has faced challenges in real-world cases due to its heavy reliance on variance, which fails to account for non-normal, heavy-tailed investment losses. In this research study, three advanced models are implemented and compared with the standard MPT framework, including conditional value-at-risk (CVaR), hierarchical risk parity (HPR), and a genetic algorithm (GA) for cardinality-constrained optimisation. The dataset includes historical prices for 18 US stocks from 2016 to 2025. Strict constraints are imposed during the implementation of these models, including long-only positioning and a per-asset cap. An evaluation of risk-adjusted performance using the sortino ratio (STR) demonstrates that clustering-driven HRP and tail-consistent CVaR models outperform MPT significantly. The CVaR-based portfolio has achieved a Sharpe ratio (SR) of 1.58 and a 2.42 STR, demonstrating its ability to mitigate downside risk. GA has achieved the highest annual return of 45.8% and 2.03 SR; simultaneously, drawdown and volatility are also high, proving the inherent trade-offs between maximising empirical SR and robust risk management. The research findings indicated that GA can efficiently build a risk-adjusted optimal investment portfolio while imposing real-world constraints.

Keywords: Investment Portfolio Development Model, Stock Selection Approach, Investment Management Framework, Genetic Algorithm, Financial Technology (FinTech) Architecture

1. Introduction

Investment in the stock market includes stock selection, which is one of the most challenging tasks for many investors in such evolved FinTech applications. Investors focus on both estimated returns and risk when deciding to invest in less risky stocks to achieve higher returns. To reduce risk, a portfolio can be built by accumulating various stocks to diversify investments instead of investing in a single stock. In quantitative finance, Markowitz is credited with introducing the concept of selecting investment portfolios based on high expected returns for a given level of risk, or vice-versa [1]. The MPT has introduced the fundamental model of risk-adjusted returns through asset diversification. Mean-variance optimisation (MVO) relies on estimations of returns and risk matrices, while significant estimation errors can lead to an inefficient portfolio [2]. To address the limitations of MPT, financial markets require advanced approaches that incorporate two crucial aspects. First, non-normal risk diversification, given that investment returns tend to be asymmetric and heavy-tailed, particularly when markets are overstressed, making variance ineffective for distinguishing between unacceptable downside risk and acceptable upside risk. Second, real investment constraints, such as cardinality limits, no short selling, and upper limit constraints, make the convex mean-variance problem NP-hard, necessitating the use of heuristic and ML algorithms to find feasible solutions [3,4].

This study implements and compares four quantitative portfolio-construction methods. The MPT maximises the utility function calculated by combining estimated return, variance, and risk aversion. A coherent risk measure maximises return while minimising tail risk using CVaR. The HRP uses a clustering technique to reduce estimation error. The GA is a heuristic method for addressing non-convex problems with cardinality constraints, such as maximising SR. This study compares the SR and STR using financial market data and FinTech to analyse the performance of optimisation models, examine the stability of trade-offs, manage risk, and handle nonconvex constraints. A systematic risk-adjusted return strategy CVaR and clustering-based HRP both significantly outperform an MPT portfolio. In contrast, the heuristic method GA outperforms with high absolute risk.

2. Related work

Traditional portfolio optimisation strategies are ineffective in understanding financial market mechanisms and constructing optimal portfolios. Since the development of MVO, there has been a continuous evolution of the asset allocation model for diversified portfolios [5]. Numerous studies have used the MVO model in practice. Still, the capital asset pricing model (CAPM), Ichimoku cloud indicator, and other optimisation models have consistently been formulated to eliminate unsystematic risk from diversified investments [6,7]. HRP has progressed by concentrating on risk to address the limitations of traditional statistical models. Back testing confirms its effectiveness in building a more diversified portfolio because it can minimise drawdown risk [8].

Measuring risk is key to optimising a portfolio under uncertainty and mitigating potential losses. Value at risk (VaR) has become a popular measure owing to its simplicity and inclusion, neither convex nor sub-additive. The weighted mean of VaR is a new risk measure, and a study claims that CVaR is a coherence-based risk measure, which is mainly used to develop a robust portfolio [9].

The GA algorithm is an optimisation model based on Darwin's theory of evolution that provides solutions to nonlinear and nonconvex optimisation problems [10]. Asset weights are grouped into populations to assess how well they fit the problem. Fitness scores are evaluated using the evolutionary function. Selection, crossover, and mutation functions are iteratively applied to find the best risk-return configuration and build an optimal portfolio [11]. A study reveals the potential use of GA to assist individual investors and their wealth and investment management businesses in order to assess the potential returns and offer portfolio advice, such as Robo Advisors [3,12].

3. Methods and proposed model and framework

The proposed model (framework) describes the implementation processes for evaluating four portfolio optimisation models, as shown in Figure 1. To estimate risk and return, data is collected and split into training, validation, and test sets. Using MPT, HRP, CVaR, and GA estimates,

DOI: 10.54254/2755-2721/2025.LD31777

optimisation models are applied, and their performance is evaluated and compared to select the final portfolio.

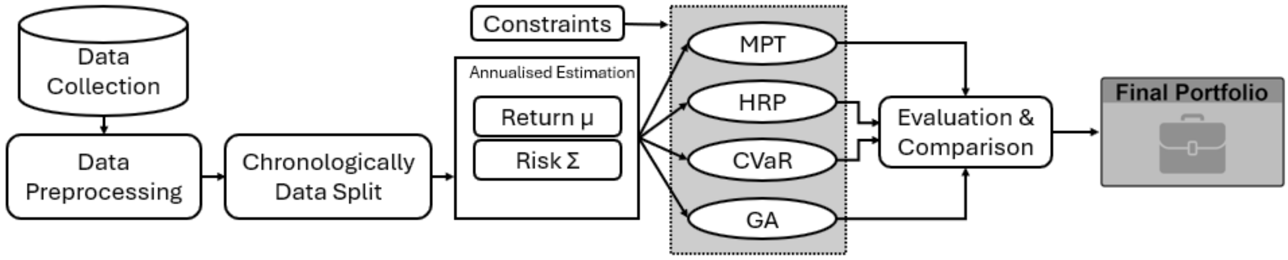

Figure 1. Proposed model architecture

3.1. Data collection and preprocessing

Stock market data collected for portfolio optimisation includes 18 distinct S&P 500 stock indexes, as shown in Figure 2. The closing prices of each index have been collected using the free, publicly available Yahoo Finance API from 2016 to 2025. Missing data has been filled using a forward-fill approach to ensure data consistency and reliability.

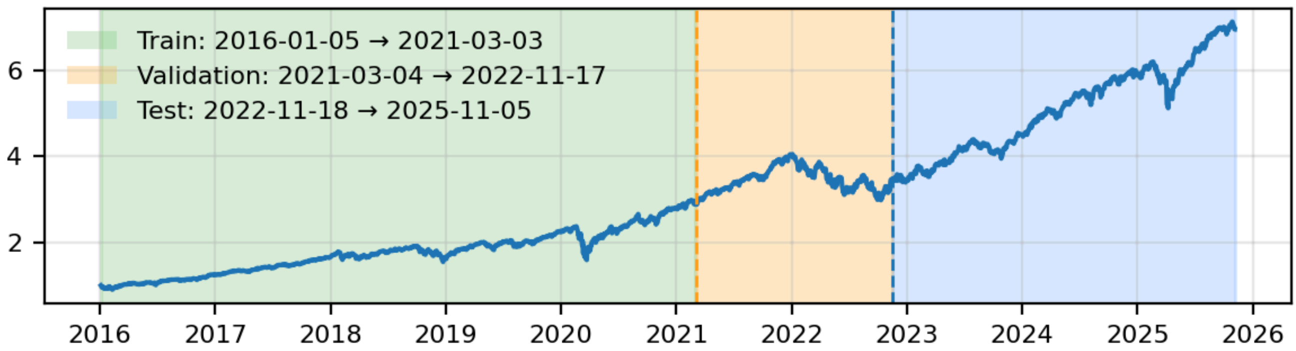

Figure 2. 10 years of data for training, validation and testing

Chronologically, data is split to prevent leakage and to adopt a realistic approach to sequential data. Initially 70% of the data is used for training and validation, while the remaining 30% is used for testing. The main objective of portfolio optimisation is to maximise return and minimise volatility. Daily price changes are calculated, and annualised returns are estimated using 252 business days.

3.2. Modern portfolio theory

The Mean-variance optimisation – MVO method based on MPT is implemented to build an optimal investment portfolio. Markwits [1,13] considers investors are risk-averse and will choose a less risky portfolio when two portfolios have the same expected return. The MVO key objective is to determine the efficient frontier of portfolios by comparing expected mean returns with expected risk. Most investors seek portfolios that offer a balance between returns and risk. The sequential least squares quadratic programming method is used to solve the MV problem efficiently [2]. The risk-reduction parameter gamma controls the amount of portfolio variation to penalise. A Ledoit-Wolf shrinkage technique can estimate covariance, as simple covariance is noisy and unstable across multiple assets.

DOI: 10.54254/2755-2721/2025.LD31777

3.3. Hierarchical risk party

The HRP model builds a portfolio using a similar type of assets in grouping instead of relying on a precise global inversion of the covariance matrix [14]. The proposed HRP model includes three steps: clustering, quasi-diagonalisation and recursive bisection. Correlation matrices are computed from asset returns and converted to correlation-distance matrices to determine the distance between them. Ward's linkage is used to build an asset hierarchy tree from a distance matrix to show correlations. The covariance matrix rows and columns are ordered according to the asset hierarchy, known as quasi-diagonalisation. It focuses on the diagonal covariance to organise risk contributions and prepare the architecture for the subsequent allocation step. The portfolio is divided into two sub-clusters, and weights are assigned inversely using a top-down approach [11,15].

3.4. Conditional value-at-risk

CVaR has been used to address MPT's inability to capture tail risk. It measures expected loss only when the loss exceeds the VaR threshold, unlike variance. Compared with MPT, CVaR optimisation provides significant downside protection during bearish market conditions, helping prevent extreme losses. Also, CVaR can be formulated as a linear program, providing a computationally efficient approach to solving CVaR objectives, even for portfolios composed of multiple assets [6].

3.5. Genetic algorithm

A genetic algorithm represents a fundamental shift towards heuristic optimisation because conventional optimisation models failed [16]. Portfolio optimisation problems that cannot be solved by convex optimisation methods, such as limiting the number of holdings, require heuristic approaches to address complex, non-convex problems. In GA, weights are optimised by maximising a fitness function that includes financial objectives and penalties for violating constraints. In L2 regularisation, weights do not become zero, but GA is prevented from finding solutions dominated by large weights [10]. By using this approach, investors can build portfolios that incorporate real-world trading constraints beyond the reach of simple quadratic or linear programming methods.

4. Experiment and results

The training dataset is chronologically split into training and validation sets to avoid overfitting. Training data is used to estimate the expected return and volatility. In contrast, validation data is used to tune the model and find the optimal hyperparameter configuration to achieve high performance with low error. The annual return (AR) measures average investment growth over a year, and the cumulative return (CR) calculates the percentage profit or loss for a specific investment period. Annual volatility (AV) quantifies the degree of fluctuation in investment return. The maximum drawdown (MDD) measures the highest loss from the peak. SR calculates investment return relative to the per-unit of risk. STR measures loss-specific efficiency using tail risk [8,16].

Each optimisation model handles diversification in starkly different ways, with constraints that prohibit short selling and limit each asset to no more than 25% of the cap weight. In the MPT portfolio, 25% of the portfolio's weight is allocated to JNJ, and approximately 21% is assigned to PG, reflecting the algorithm's tendency to place a heavy weight on assets with the most favourable return and risk characteristics, as shown in Figure 3. However, HRP's portfolio demonstrates a diversified weight assignment, diversifying risk across clusters of assets that are not even similar. All assets remain below 15% of capital, confirming the model's inherent architecture and its

DOI: 10.54254/2755-2721/2025.LD31777

objective of diversifying risk across the portfolio. CVaR has also well-diversified weights. Compared with the optimisation models, the GA portfolio has allocated 25% of its capital to NVDA and placed significant weight in high-growth technology and consumer stocks, including approximately 15% in NFLX, approximately 16% in JPM, and approximately 14% in MSFT.

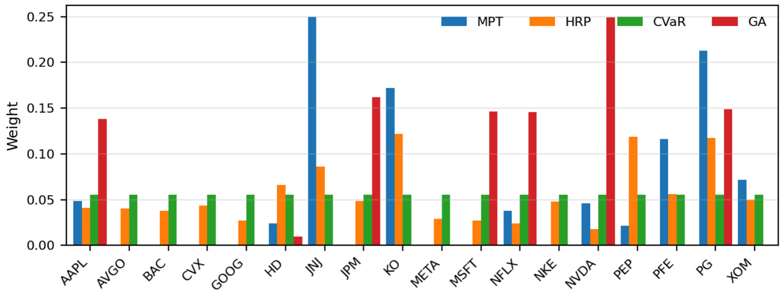

Figure 3. Capital weight assigned to the multiple indices

GA has achieved significant performance with 45.80% AR and 260.00% CR, see Table 1. The CVaR portfolio achieved 24.90% AR and 102.10% CR. A highly focused MVO solution that prioritised risk avoidance gained the lowest returns across all portfolios. In risk-adjusted performance metrics, GA and CVaR outperform HRP and MPT with high SR and STR. MPT performs very poorly compared to HRP, with 0.77 SR and 1.11 STR. A 1.09 SR confirms the stability and reliability of HRP and demonstrates its significance to portfolio performance. In volatile periods, the CVaR portfolio produced the highest 1.58 SR and 2.42 STR among all deterministic and convex methods. However, GA has outperformed all proposed models with 2.03 SR and 3.24 STR.

The portfolio developed using MPT has achieved the lowest -10.90% MDD, as Table 1 exhibits. Both low MDD and AR indicate that MPT's risk-aversion parameter and estimation error have been strictly penalised. To achieve this high AR while managing downside risk, the CVaR portfolio has incurred -17.4% MDD, which reflects CVaR's core objective of minimising tail losses.

Table 1. Performance comparison of portfolio optimisation models

	AR	AV	SR	STR	MDD	CR
MPT(Train)	0.161	0.162	0.867	1.225	-0.280	1.756
HRP (Train)	0.168	0.177	0.836	1.172	-0.323	1.844
CVaR (Train)	0.200	0.202	0.893	1.249	-0.329	2.442
GA (Train)	0.308	0.265	1.090	1.559	-0.426	5.539
MPT (Test)	0.107	0.112	0.772	1.116	-0.109	0.344
HRP (Test)	0.148	0.117	1.094	1.620	-0.136	0.515
CVaR (Test)	0.249	0.144	1.588	2.426	-0.174	1.021
GA (Test)	0.458	0.216	2.030	3.242	-0.207	2.600

Despite strong performance, the GA portfolio had the highest -20.7% MDD and the highest 21.6% AV. The result demonstrates an inherent trade-off between risk and return. To maximise SR, capital must be allocated to volatile, high-return assets, making it more vulnerable to market declines.

DOI: 10.54254/2755-2721/2025.LD31777

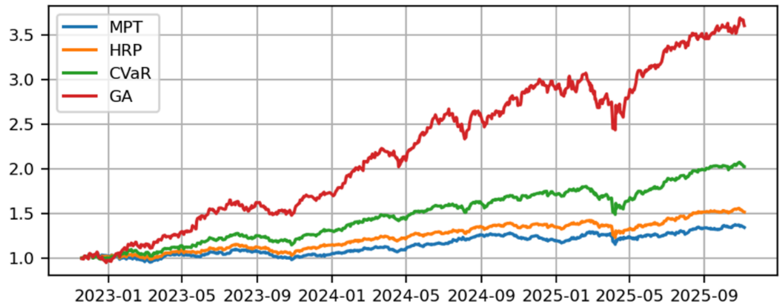

Figure 4. Cumulative return acquired using different models

The CR plot provides visual confirmation of the quantitative results acquired using four optimisation models, as shown in Figure 4. The GA portfolio outperforms the other optimisation models across the testing period, demonstrating the power of maximum SR in a sparse, highly concentrated portfolio. In contrast, CVaR shows consistent growth with very low fluctuations compared to GA, MPT, and HRP models, emphasising the importance of adopting strategies beyond MPT to achieve significant AR and manage expected risk.

5. Conclusion

The study has implemented and compared the performance of four portfolio optimisation models under real-world constraints to demonstrate the crucial progress in portfolio theory beyond a symmetrical risk estimation model. Compared with MPT, CVaR has proved to be the most compelling alternative, delivering higher risk-adjusted returns by utilising a computationally efficient tail risk measure. In addition to providing a stable, structurally diversified solution, HRP has demonstrated its capability to minimise the effect of errors associated with covariance estimation. Despite the greater AV and MDD, the ML-based GA describe its value as a heuristic for resolving non-convex problems under real-world constraints and achieving high SR. Given the complexity of optimal portfolio construction, robust portfolios must be developed using strategies that are robust to capture data-driven uncertainty and that adhere to a coherent risk framework. In further research studies, hybrid optimisation models can be implemented to build more robust solutions. The stock market needs to be predicted and incorporated into historical data to develop a more realistic, optimal portfolio that accounts for real-world constraints.

References

[1] H. Markowitz, "Portfolio Selection, " 1952. doi: https: //doi.org/10.2307/2975974.
[2] P. Xidonas, R. Steuer, and C. Hassapis, "Robust portfolio optimization: a categorized bibliographic review, " Ann Oper Res, vol. 292, no. 1, pp. 533–552, Sep. 2020, doi: https: //doi.org/10.1007/s10479-020-03630-8.
[3] S. O. Adebiyi, O. O. Ogunbiyi, and B. B. Amole, "Artificial intelligence model for building investment portfolio optimization mix using historical stock prices data, " Rajagiri Management Journal, vol. 16, no. 1, pp. 36–62, Apr. 2022, doi: https: //doi.org/10.1108/RAMJ-07-2020-0036.
[4] D. C. Fransisca, Sukono, D. Chaerani, and N. A. Halim, "Robust Portfolio Mean-Variance Optimization for Capital Allocation in Stock Investment Using the Genetic Algorithm: A Systematic Literature Review, " Aug. 01, 2024, Multidisciplinary Digital Publishing Institute (MDPI). doi: https: //doi.org/10.3390/computation12080166.
[5] Y. H. Chou, S. Y. Kuo, and Y. T. Lo, "Portfolio Optimization Based on Funds Standardization and Genetic Algorithm, " IEEE Access, vol. 5, pp. 21885–21900, Sep. 2017, doi: https: //doi.org/10.1109/ACCESS.2017.2756842.
[6] J. M. Berutich, F. López, F. Luna, and D. Quintana, "Robust technical trading strategies using GP for algorithmic portfolio selection, " Expert Syst Appl, vol. 46, pp. 307–315, Mar. 2016, doi: https:

//doi.org/10.1016/j.eswa.2015.10.040.
[7] B. Joemon, M. A. Ghazanfar, M. A. Azam, N. Z. Jhanjhi, and A. A. Khan, "Novel heuristics for Stock portfolio optimization using machine learning and Modern Portfolio Theory, " in 2nd International Conference on Business Analytics for Technology and Security, ICBATS 2023, Institute of Electrical and Electronics Engineers Inc., 2023. doi: https: //doi.org/10.1109/ICBATS57792.2023.10111321.
[8] T. Kaczmarek and K. Perez, "Building portfolios based on machine learning predictions, " Economic Research-Ekonomska Istrazivanja , vol. 35, no. 1, pp. 19–37, 2022, doi: https: //doi.org/10.1080/1331677X.2021.1875865.
[9] S. Banihashemi and S. Navidi, "Portfolio performance evaluation in Mean-CVaR framework: A comparison with non-parametric methods value at risk in Mean-VaR analysis, " Operations Research Perspectives, vol. 4, pp. 21–28, 2017, doi: https: //doi.org/10.1016/j.orp.2017.02.001.
[10] W. Ahn, H. S. Lee, H. Ryou, and K. J. Oh, "Asset allocation model for a robo-advisor using the financial market instability index and genetic algorithms, " Sustainability (Switzerland), vol. 12, no. 3, Feb. 2020, doi: https: //doi.org/10.3390/su12030849.
[11] X. Li, A. S. Uysal, and J. M. Mulvey, "Multi-period portfolio optimization using model predictive control with mean-variance and risk parity frameworks, " Eur J Oper Res, vol. 299, no. 3, pp. 1158–1176, Jun. 2022, doi: https: //doi.org/10.1016/j.ejor.2021.10.002.
[12] A. Qadoos, H. AbouGrad, J. Wall, and M. Saeed Sharif, "AI Investment Advisory: Examining Robo-Advisor Adoption Using Financial Literacy and Investment Experience Variables, " 2025, doi: https: //doi.org/10.54254/2755-2721/2025.21287.
[13] A. Chaweewanchon and R. Chaysiri, "Markowitz Mean-Variance Portfolio Optimization with Predictive Stock Selection Using Machine Learning, " International Journal of Financial Studies, vol. 10, no. 3, Sep. 2022, doi: https: //doi.org/10.3390/ijfs10030064.
[14] Tanushka M, "Customized Investment Plans Using CAPM and HRP, " International Journal of Innovative Research in Engineering & Management (IJIREM), no. 8, pp. 2350–0557, 2021, doi: https: //dx.doi.org/10.2139/ssrn.3978949.
[15] G. Costa and R. H. Kwon, "Risk parity portfolio optimization under a Markov regime-switching framework, " Quant Finance, vol. 19, no. 3, pp. 453–471, Mar. 2019, doi: https: //doi.org/10.1080/14697688.2018.1486036.
[16] S. Millán-Palacios and J. Sánchez-Soriano, "Investment Portfolios Optimization with Genetic Algorithm: An Approach Applied to the Spanish Market (IBEX 35), " Electronics (Switzerland), vol. 14, no. 13, Jul. 2025, doi: https: //doi.org/10.3390/electronics14132559.