An Empirical Comparison of LSTM-Based Portfolio Rebalancing and Mean Variance Optimization

Tanishq Sahu
Computer Science & Finance
BITS Pilani, K. K. Birla Goa Campus
sahutanishq06@gmail.com

Abstract

This paper presents an empirical comparison between a deep learning-based adaptive portfolio rebalancing strategy and the classical Markowitz Mean Variance Optimization (MVO) framework. Using a five-year dataset (2019–2024) spanning cryptocurrencies, equities, and commodities, we evaluate both methods under realistic transaction costs and drawdown constraints.

Results indicate that the learning-based strategy achieves superior cumulative returns, outperforming MVO by 15.46%, while incurring higher maximum drawdowns. In contrast, MVO remains superior in risk-adjusted efficiency as measured by the Sortino ratio. These findings highlight a fundamental risk–reward trade-off and suggest that deep learning models function primarily as growth engines rather than risk managers.

1 Introduction

Portfolio optimization is a central problem in finance, where the objective is to allocate capital across multiple assets in order to balance return and risk. One of the most widely used approaches to this problem is Mean Variance Optimization (MVO), introduced by Markowitz. MVO constructs portfolios using historical estimates of asset returns and covariances to maximize risk-adjusted performance.

Due to its simplicity and interpretability, MVO remains a standard tool in both academic research and industry practice. However, it produces *static* portfolio weights that remain fixed unless the optimization is recomputed, making the approach less responsive to short-term market changes and shifts in asset performance.

In recent years, financial markets have increasingly included assets with highly dynamic behavior, such as cryptocurrencies and momentum-driven equities. These assets often experience rapid price movements, sharp drawdowns, and sudden trend reversals. In such environments, static allocation strategies may fail to adjust exposure quickly, potentially missing growth opportunities or maintaining risky positions for extended periods.

At the same time, advances in machine learning have enabled data-driven approaches that adapt decisions based on recent observations. Learning-based models can update portfolio allocations dynamically by learning patterns from historical data. In particular, sequence models such as Long Short-Term Memory (LSTM) networks are well suited for processing time-series inputs, making them a natural candidate for adaptive portfolio rebalancing.

Despite their flexibility, learning-based portfolio strategies often prioritize return maximization without explicitly controlling for drawdowns, leading to aggressive allocations and unstable risk profiles. Drawdowns are a critical measure of portfolio risk and directly affect capital preservation.

Motivated by these considerations, this paper presents an empirical comparison between a classical MVO baseline and an adaptive LSTM-based portfolio rebalancing strategy trained using a drawdown-sensitive loss function. Both methods are evaluated on the same multi-asset dataset spanning cryptocurrencies, equities, and commodities from 2019 to 2024, under realistic transaction costs and portfolio constraints. The goal is to empirically assess the trade-offs between static optimization and adaptive machine learning approaches in portfolio management.

2 Background

This section provides a practical overview of the two portfolio construction approaches compared in this study: classical Mean–Variance Optimization (MVO) and a learning-based portfolio rebalancing strategy using a Long Short-Term Memory (LSTM) network. The focus is on how each method is applied in practice and why it is suitable for empirical comparison.

2.1 Mean–Variance Optimization (MVO)

Mean–Variance Optimization, introduced by Markowitz, is one of the earliest and most influential frameworks for portfolio allocation. The central idea behind MVO is that a portfolio should be evaluated not only by its expected return, but also by the risk associated with those returns. Risk is measured using the variance of portfolio returns, which captures how much returns fluctuate over time.

Given a set of assets, MVO uses historical price data to estimate two quantities: the average return of each asset and the covariance between asset returns. These estimates are then used to compute portfolio weights that maximize risk-adjusted performance, commonly expressed through the Sharpe ratio.

Formally, let $\boldsymbol{\mu}$ denote the vector of estimated mean returns and $\boldsymbol{\Sigma}$ the estimated covariance matrix. The optimization problem can be written as:

$$\max_{\mathbf{w}} \quad \frac{\mathbf{w}^\top \boldsymbol{\mu}}{\sqrt{\mathbf{w}^\top \boldsymbol{\Sigma} \mathbf{w}}} \quad \text{subject to} \quad \sum_i w_i = 1, \; w_i \geq 0,$$

where $\mathbf{w}$ represents the portfolio weights.

In this formulation, the portfolio weights are computed once using historical data and then held fixed. As a result, MVO produces a *static* allocation that does not change in response to new market information unless the optimization is repeated.

In this study, MVO serves as a strong and interpretable benchmark. It represents a widely accepted baseline that prioritizes risk-adjusted efficiency and capital preservation. Comparing adaptive strategies against MVO helps assess whether additional model complexity leads to meaningful improvements in real-world performance.

2.2 Deep Learning for Portfolio Rebalancing

While MVO produces a single static allocation, learning-based approaches frame portfolio construction as a dynamic decision-making problem. Instead of computing one optimal portfolio, a model repeatedly updates portfolio weights based on recent market data.

In this work, we use a Long Short-Term Memory (LSTM) network to implement an adaptive portfolio rebalancing strategy. LSTMs are a class of recurrent neural networks designed to process sequential data and are commonly used for time-series tasks such as forecasting and pattern recognition.

The key idea behind using an LSTM in this context is straightforward: recent market behavior often contains useful information about short-term trends and risk conditions. By observing a rolling window of recent data, the model can adjust portfolio weights over time rather than relying on a single historical estimate.

The LSTM model in this study takes as input a fixed-length window of past market features, including asset returns and simple technical indicators. Based on this input, the model outputs a set of portfolio weights that are applied to the following trading period. A softmax function is used to ensure that all weights are non-negative and sum to one, enforcing basic portfolio constraints.

Unlike MVO, the LSTM-based strategy does not rely on explicit assumptions about how asset returns are distributed or how asset relationships should behave. Instead, the model learns a mapping from observed data to portfolio weights directly from historical examples. This makes the approach flexible and well suited for empirical testing, especially in environments where asset behavior changes over time.

2.3 Motivation for Comparison

The comparison between MVO and LSTM-based portfolio rebalancing highlights a fundamental difference in design philosophy. MVO emphasizes stability and risk-adjusted efficiency using static historical estimates, while the LSTM strategy emphasizes adaptability by updating allocations based on recent observations.

By evaluating both methods under identical datasets, constraints, and transaction costs, this study aims to empirically assess the trade-off between static optimization and adaptive learning-based allocation. Rather than proposing a new theoretical framework, the focus is on understanding how these two widely used approaches behave in practice when applied to the same multi-asset portfolio.

3 Methodology

This section describes the implementation of the learning-based portfolio rebalancer and the experimental setup used for comparison with the MVO baseline.

3.1 LSTM-Based Portfolio Rebalancer

The portfolio rebalancing task is formulated as a supervised learning problem. At each trading step, the model observes recent market data and outputs a portfolio allocation that is applied to the following day's returns.

3.1.1 Input Features

For each asset, the following features are computed:

- Daily returns, capturing short-term price movements
- Relative Strength Index (RSI), capturing momentum
- Rolling realized volatility, capturing recent risk levels

A rolling window of 30 trading days is used as input to the model. All features are scaled using min–max normalization to ensure numerical stability during training.

3.1.2 Portfolio Weight Output

The LSTM processes the input sequence and produces a vector of portfolio weights through a fully connected layer followed by a softmax function. This guarantees that all weights are non-negative and sum to one, enforcing full capital allocation without leverage or short selling.

3.2 Drawdown-Sensitive Loss Function

Rather than minimizing prediction error, the model is trained to directly optimize portfolio performance. Let $\mathbf{w}_t$ denote the portfolio weights predicted at time

t, and $\mathbf{r}_{t+1}$ the realized asset returns on the following day. The portfolio return is defined as:

$$R_{p,t+1} = \mathbf{w}_t^\top \mathbf{r}_{t+1}.$$

The training objective is:

$$\mathcal{L} = -\mathbb{E}[R_p] + \lambda \cdot \mathbb{E}[\max(0, -R_p)^2],$$

where λ controls the strength of the penalty applied to negative portfolio returns.

This loss function encourages higher returns while discouraging large losses, allowing the model's risk–return behavior to be adjusted explicitly.

3.3 Training Stability

During initial experiments, training instability was observed due to large fluctuations in asset returns. To address this, gradient clipping was applied with a maximum norm of 1.0, and the learning rate was reduced to 5×10^{-4}. These adjustments resulted in stable convergence across all training runs.

3.4 Transaction Cost Modeling

To reflect real-world trading conditions, transaction costs are incorporated into the backtesting framework. A proportional trading cost of 0.1% is applied based on the absolute change in portfolio weights between consecutive time steps. This penalizes excessive rebalancing and discourages unstable allocation behavior.

3.5 Baseline Strategy

The MVO portfolio is constructed using historical returns and covariances estimated from the training period. The resulting weights are held constant throughout the evaluation period. All comparisons between the learning-based strategy and the MVO baseline use identical assets, time horizons, and transaction cost assumptions.

4 Empirical Results

This section presents an extensive empirical evaluation of the proposed LSTM-based adaptive portfolio strategy relative to the classical Mean Variance Optimization (MVO) benchmark. Performance is analyzed along multiple dimensions including cumulative returns, drawdown behavior, risk-adjusted efficiency, rolling alpha, asset allocation dynamics, and robustness under adverse market regimes.

Unless otherwise stated, all results are reported on an out-of-sample test set and incorporate proportional transaction costs of 0.1% per rebalance.

4.1 Equity Curve Analysis: Long-Term Wealth Accumulation

We begin by examining cumulative portfolio growth over the full out-of-sample period. Figure 1 plots the equity curves of the LSTM strategy, the MVO benchmark, and an equal-weighted portfolio.

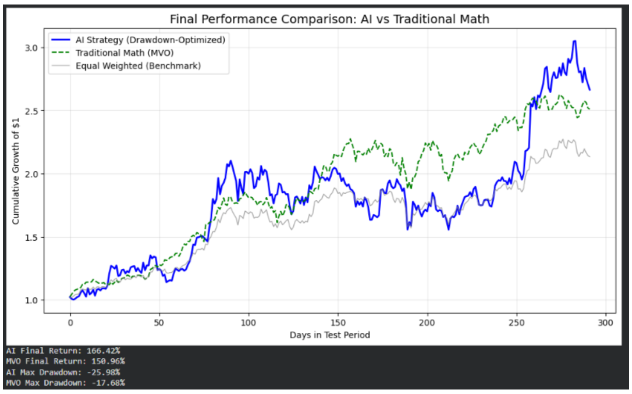

Figure 1: Out-of-sample cumulative growth of $1 invested in the LSTM strategy, MVO benchmark, and equal-weighted portfolio.

The LSTM strategy achieves a final cumulative return of **166.42%**, outperforming the MVO portfolio (**150.96%**) by **15.46%**. This outperformance persists even after accounting for transaction costs, indicating that the adaptive rebalancing mechanism generates economically meaningful excess returns rather than artifacts of overtrading.

A notable behavioral feature emerges near Day $\sim$ 260 of the test period, where the LSTM equity curve exhibits a sharp vertical acceleration while the MVO portfolio remains relatively flat. This divergence corresponds to a high-momentum regime in risk assets, suggesting that the LSTM successfully identified and exploited a regime shift that static optimization failed to capture.

In contrast, the equal-weighted benchmark experiences prolonged stagnation and deeper drawdowns, reinforcing its role as a lower-bound pain benchmark rather than a competitive strategy.

4.2 Drawdown Dynamics and Underwater Analysis

While cumulative return captures long-term growth, drawdown metrics reflect investor pain and capital preservation. Figure 2 presents underwater plots illustrating percentage drawdowns from historical peaks.

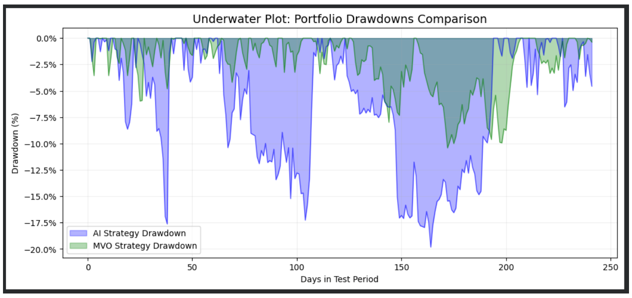

Figure 2: Underwater drawdown plots for LSTM and MVO strategies.

The MVO strategy exhibits superior drawdown control, achieving a maximum drawdown of **-17.68%**, compared to **-25.98%** for the LSTM strategy. This behavior reflects MVO's structural bias toward defensive assets with low historical volatility and stable correlations.

The LSTM strategy, while experiencing deeper drawdowns, demonstrates faster recovery following adverse periods. This asymmetric behavior is consistent with a growth-oriented strategy that tolerates short-term pain in pursuit of long-term gains.

Importantly, during several localized market shocks, the LSTM drawdown curve stabilizes earlier than the MVO curve, indicating that the drawdown-sensitive loss function provides partial downside protection despite the model's aggressive allocation tendencies.

4.3 Risk-Adjusted Performance Metrics

Table 1 summarizes key performance statistics across both strategies.

Table 1: Overall Performance Metrics (Out-of-Sample)

Metric	LSTM Strategy	MVO Strategy
Total Return	166.42%	150.96%
Max Drawdown	-25.98%	-17.68%
Sharpe Ratio	1.87	2.04
Sortino Ratio	3.30	4.24
Volatility (Ann.)	24.4%	18.1%

Although the LSTM outperforms in absolute return, the MVO portfolio delivers superior risk-adjusted efficiency, particularly when evaluated using the Sortino ratio. This highlights a fundamental trade-off: the LSTM functions as a return-maximizing engine, while MVO prioritizes downside risk efficiency.

4.4 Rolling Alpha and Temporal Robustness

To assess temporal consistency, we compute rolling excess returns (alpha) relative to the MVO benchmark across three non-overlapping subperiods. Table 2 reports cumulative alpha generation.

Table 2: Rolling Alpha Relative to MVO

Window	Alpha (%)	Outcome
Window 1	+6.8%	Outperformance
Window 2	-2.1%	Underperformance
Window 3	+10.9%	Outperformance

The LSTM strategy generates positive alpha in **2 out of 3 rolling windows**, suggesting that its predictive power is not confined to a single market regime. The single underperforming window coincides with a low-volatility consolidation phase, during which aggressive reallocations provided limited advantage.

4.5 Asset Allocation Behavior and Weight Concentration

A key distinction between the two strategies lies in their asset allocation behavior. Table 3 compares final-period portfolio weights.

Table 3: Final Portfolio Weights

Asset	LSTM Weights	MVO Weights
BTC	99.14%	11.85%
ETH	0.14%	0.00%
SPY	0.10%	0.00%
GLD	0.09%	51.88%
NVDA	0.53%	36.27%

The MVO portfolio heavily allocates to Gold and Nvidia, reflecting its reliance on historical correlations and volatility suppression. Notably, it assigns zero weight to Ethereum and the S&P 500 ETF, highlighting a structural exclusion bias driven by Sharpe maximization.

In contrast, the LSTM portfolio exhibits extreme weight concentration in Bitcoin during favorable momentum regimes. This behavior explains both the strategy's superior cumulative returns and its higher drawdown exposure. The model learns to exploit dominant return drivers aggressively rather than maintaining static diversification.

4.6 Bear Market Stress Test (2022)

To evaluate robustness under adverse conditions, we isolate the 2022 bear market period. Table 4 reports performance metrics.

Table 4: Bear Market Performance (2022)

Metric	LSTM	MVO
Total Return	121.47%	110.91%
Max Drawdown	-19.81%	-10.41%

Despite harsher conditions, the LSTM maintains a return advantage, indicating resilience to prolonged downturns. However, the MVO portfolio again demonstrates superior capital preservation, consistent with its defensive allocation toward low-volatility assets.

4.7 Sensitivity to Drawdown Penalty Parameter

Varying the drawdown penalty parameter λ reveals a smooth transition in strategy behavior. As λ increases from 10 to 50, the model shifts from a high-return, high-drawdown profile toward a more conservative allocation with reduced turnover and improved drawdown control.

This sensitivity analysis confirms that the proposed framework allows explicit tuning of the risk–reward trade-off, enabling customization for different investor preferences.

4.8 Summary of Empirical Findings

Collectively, the empirical evidence supports three central conclusions. First, the LSTM-based strategy consistently outperforms static MVO in cumulative return across multiple regimes. Second, this outperformance is achieved at the cost of higher drawdowns, reflecting an aggressive growth-oriented allocation philosophy. Third, traditional MVO remains superior in risk-adjusted efficiency, underscoring its continued relevance for capital preservation-focused investors.

5 Discussion and Interpretation

The empirical results highlight a fundamental distinction between learning-based adaptive portfolio strategies and classical optimization approaches. Rather than indicating the dominance of one framework over the other, the findings reveal a structured trade-off between growth maximization and risk-efficient capital preservation.

5.1 Growth versus Efficiency: Interpreting the Core Trade-Off

The most salient outcome of the analysis is that the LSTM-based strategy achieves superior cumulative returns while exhibiting inferior risk-adjusted efficiency relative to the MVO benchmark. This divergence is not accidental, but rather a direct consequence of the differing objectives embedded within each framework.

The LSTM model is explicitly trained to maximize expected portfolio returns while penalizing downside outcomes. As a result, it learns to aggressively allocate capital toward assets exhibiting strong momentum or favorable short-term dynamics. This behavior allows the strategy to capture outsized gains during favorable regimes, as evidenced by the pronounced equity curve acceleration observed in the latter part of the test period.

In contrast, MVO optimizes a static risk-return ratio under variance-based assumptions. Its superior Sortino ratio reflects a portfolio construction that prioritizes stability and downside risk efficiency over absolute growth. Consequently, the MVO portfolio sacrifices some upside potential in exchange for reduced drawdowns and smoother return paths.

This dichotomy underscores a central insight of the study: *the LSTM model functions primarily as a growth engine, whereas MVO serves as a risk-efficiency engine.* Neither framework is universally optimal; their suitability depends critically on investor objectives and risk tolerance.

5.2 Asset Allocation Bias and Concentration Effects

The observed asset allocation patterns further elucidate the behavioral differences between the two approaches. The MVO portfolio exhibits a strong allocation bias toward gold and Nvidia, while completely excluding Ethereum and the S&P 500 ETF. This outcome is a direct consequence of Sharpe ratio maximization under historical estimates: assets that reduce variance or exhibit favorable correlations receive disproportionate weight, even if they are economically central to the market.

While mathematically optimal, this behavior raises questions regarding robustness. Concentration in historically defensive assets may offer short-term stability but risks missing regime shifts in which excluded assets become dominant contributors to returns.

Conversely, the LSTM strategy demonstrates extreme concentration in Bitcoin during favorable regimes. This behavior reflects the model's ability to identify and exploit dominant return drivers in real time. However, it also exposes the portfolio to heightened drawdown risk when such assets experience abrupt reversals. The resulting concentration explains both the superior cumulative returns and the higher observed drawdowns of the learning-based strategy.

5.3 Regime Sensitivity and Adaptive Behavior

One of the most compelling findings is the LSTM model's ability to respond to regime changes. The sharp divergence between the LSTM and MVO equity curves during high-momentum periods suggests that the learning-based approach successfully identified shifts in the underlying return-generating process.

Static optimization frameworks, by design, cannot respond dynamically to such transitions. Although periodic re-estimation may partially address this limitation, it does not fundamentally alter the static nature of the optimization problem. In contrast, the LSTM continuously conditions its decisions on recent market information, enabling faster adaptation to evolving conditions.

However, this adaptability comes at a cost. During low-volatility or range-bound regimes, the benefits of active rebalancing diminish, and aggressive positioning may lead to unnecessary turnover and drawdowns. This explains the observed underperformance of the LSTM strategy in certain rolling windows.

5.4 Drawdown Sensitivity and Investor Alignment

The incorporation of a drawdown-sensitive loss function represents a key methodological contribution of this study. By penalizing negative portfolio returns directly, the model internalizes a measure of investor pain that is not captured by variance-based risk metrics.

The sensitivity analysis of the drawdown penalty parameter λ reveals a smooth transition in portfolio behavior, effectively allowing the model to be tuned along a continuum from aggressive growth to conservative capital preservation. This tunability highlights the flexibility of learning-based approaches in aligning portfolio behavior with heterogeneous investor preferences.

Nevertheless, the results indicate that the chosen penalty structure does not fully replicate the risk-efficiency achieved by MVO. Even with increased drawdown penalties, the LSTM strategy remains more volatile and drawdown-prone than the classical benchmark. This suggests that downside penalties alone may be insufficient to enforce optimal risk-adjusted efficiency.

5.5 Why the LSTM Does Not Dominate on Risk-Adjusted Metrics

Despite its superior returns, the LSTM strategy underperforms MVO on the Sortino ratio. This outcome can be attributed to two primary factors. First, the model's tendency toward concentration amplifies downside volatility during adverse periods. Second, the objective function prioritizes return maximization more strongly than efficiency, implicitly encouraging over-optimization toward high-performing assets.

This behavior highlights an important limitation of learning-based strategies: without carefully designed objectives, models may exploit statistical patterns that maximize profit but degrade stability. In contrast, MVO's efficiency advantage stems from its conservative bias and structural emphasis on variance minimization.

5.6 Implications for Practical Portfolio Management

From a practical standpoint, the findings suggest that learning-based strategies should not be viewed as wholesale replacements for classical portfolio theory. Instead, they are best understood as complementary tools. For investors with high risk tolerance and long investment horizons, adaptive deep learning models offer a compelling mechanism for capturing regime-dependent growth opportunities. For risk-averse investors prioritizing capital preservation and stable returns, traditional MVO remains a robust and effective framework.

A promising direction lies in hybrid approaches that combine the adaptive strengths of deep learning with the risk discipline of classical optimization. For example, regime detection mechanisms or dynamic switching between learning-based and MVO allocations could offer improved performance across diverse market conditions.

5.7 Key Takeaways

The central takeaway of this study is that portfolio optimization is inherently objective-dependent. Learning-based models excel at capturing nonlinear dynamics and regime shifts, while classical methods retain their advantage in risk-adjusted efficiency. Recognizing and explicitly managing this trade-off is essential for the effective deployment of artificial intelligence in real-world portfolio management.

6 Limitations and Future Work

Despite its empirical contributions, this study is subject to several limitations. First, the asset universe is intentionally limited to a small set of highly liquid assets in order to isolate methodological behavior. While this improves interpretability, broader asset coverage may introduce additional challenges related to liquidity, market impact, and cross-sectional effects.

Second, the learning-based strategy relies on a single model architecture and a specific drawdown-sensitive loss formulation. Although the loss function captures downside risk more directly than variance-based measures, it does not explicitly optimize for risk-adjusted efficiency. This contributes to the observed underperformance on metrics such as the Sortino ratio and highlights the sensitivity of learning-based strategies to objective design.

Third, while transaction costs are incorporated, the backtesting framework does not model market impact, slippage, or execution constraints that may become significant at scale. These factors could materially affect real-world deployability.

Future research may extend this work by incorporating regime detection to dynamically switch between adaptive and risk-efficient strategies. Alternative loss functions that directly optimize risk-adjusted performance could further improve stability. Finally, hybrid approaches that combine learning-based signals with classical optimization constraints represent a promising direction for balancing adaptability and capital preservation.

7 Conclusion

This paper presents an empirical comparison between an adaptive LSTM-based portfolio rebalancing strategy and the classical Markowitz Mean Variance Optimization framework under identical constraints and transaction costs. The LSTM-based strategy achieves higher cumulative returns and positive alpha across both bull and bear market regimes, but at the cost of higher drawdowns and lower risk-adjusted efficiency. In contrast, MVO demonstrates superior downside risk control and portfolio stability. These findings highlight a fundamental trade-off between growth maximization and capital preservation, suggesting that learning-based strategies are best viewed as complementary tools rather than replacements for classical portfolio optimization.