Enhanced Mean-Variance Optimization Using Multi-Branch LSTM

Chaitanya Anil Palghadmal

work.chaitanyap@gmail.com

October 16, 2025

Table of contents

Introduction

This report outlines the development and implementation of a stock price prediction model using a Multi-Branch Long Short-Term Memory (LSTM) network, integrated with a mean-variance optimization framework for portfolio management. The objective of this work was not to outperform a tuned baseline, but to evaluate whether architectural inductive bias (via asset-specific branches) meaningfully improves robustness and scalability under realistic constraints.

Methodology

Research and Model Selection:

After reviewing various stock price prediction models from research papers, including Recurrent Neural Networks (RNNs), Random Forests, and XGBoost, LSTM was chosen due to its ability to handle univariate time series data and capture long-term dependencies in stock price movements.

Multi-Branch LSTM Architecture:

A multi-branch LSTM architecture was implemented to capture interdependencies between ETFs in the portfolio, reducing complexity and computational costs compared to developing separate models for each ETF. This approach allowed for a more holistic view of the market and improved scalability.

Data Preprocessing and Model Training:

Challenges and Solutions:

- Data Preprocessing: Non-stationary raw price data was stabilized using differencing techniques.
- Model Training: High initial training loss was mitigated by adjusting weight initialization and implementing learning rate scheduling.
- Hyperparameter Tuning: Manual fine-tuning was performed to balance training efficiency and model performance, avoiding time-consuming grid or random search methods.

Failure Modes & Design Decisions

Loss Functions

Explored L2 and BCE losses (after briefly reframing the problem as classification), but observed poorer out-of-sample behavior. Ultimately settled on L1 (MAE), which provided more stable convergence under highly noisy financial targets. The simplicity and convexity of MAE appeared to improve generalization relative to variance-sensitive alternatives.

Optimization & Scheduling

Used Adam throughout due to its robustness across experiments. Tested multiple learning-rate schedules (stepwise, cyclical), but found best stability using a threshold-based early-stopping rule that halts training after sustained non-improvement relative to the best observed loss. This reduced overfitting without introducing additional tuning complexity.

Normalization & Data Handling

Evaluated several normalization strategies. Standard techniques were rejected due to loss of autocorrelation structure. Final preprocessing used log transformations followed by normalization relative to each asset's sample-space mean, preserving temporal dependencies critical for forecasting.

Throughout experimentation, priority was given to identifying fragility and regime sensitivity rather than maximizing backtest metrics. Several additional variants were explored but are omitted here for brevity, as they exhibited similar failure characteristics.

Scalability and Liquidity

The model utilizes sector ETFs, which offer high liquidity and diversification, ensuring scalability while effectively managing risk.

Results and Performance

The integration of the Multi-Branch LSTM predictions with a mean-variance optimization framework yielded mixed results. Despite the added complexity of the predictive Multi-Branch LSTM, the model only marginally surpasses market average returns. Its primary advantage lies in the efficient portfolio weight calculation for a large number of assets when the covariance matrix is precomputed.

Standard MVO Performance:

- Annualized Log Return: 0.345
- Annualized Actual Return: 0.412
- Annualized Actual Return (after 0.1% transaction costs): 0.407
- Expected Sharpe Ratio: 2.90

Enhanced MVO Performance:

- Annualized Log Return: 0.133
- Annualized Actual Return: 0.142
- Annualized Actual Return (after 0.1% transaction costs): 0.140
- Expected Sharpe Ratio: 1.10

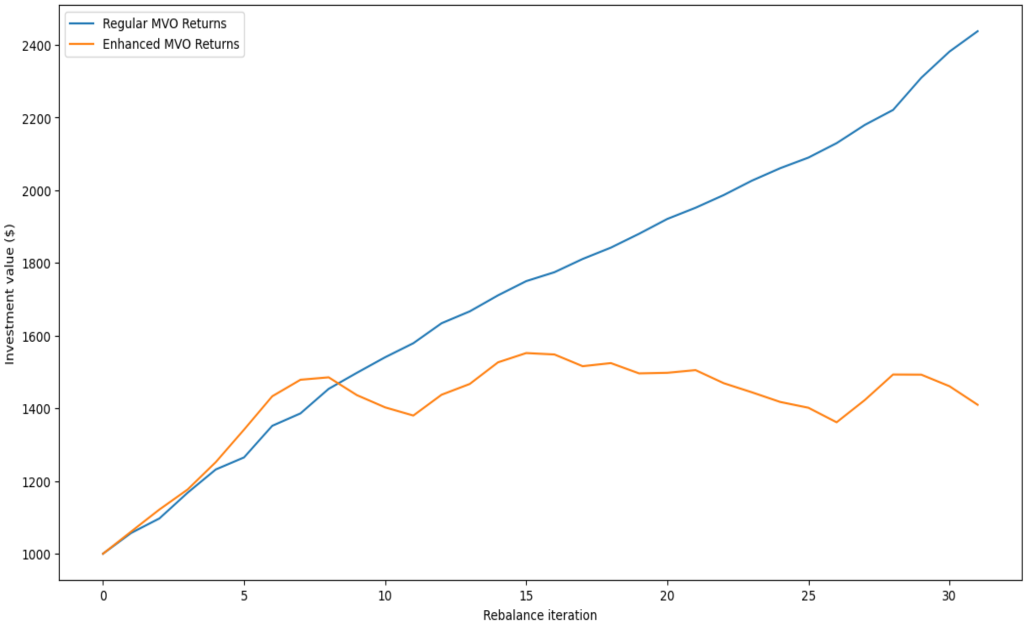

Figure 2: Investment Growth Comparison

Key Takeaways:

- Standard MVO with a shorter rebalance frequency (e.g., monthly) significantly outperforms the Enhanced MVO, achieving higher returns and a superior Sharpe ratio.
- The Enhanced MVO only slightly exceeds market average returns, raising questions about the value of its added complexity.
- The primary benefit of the Multi-Branch LSTM approach may be its computational efficiency in large-scale portfolio optimization.
- Robust performance across various market conditions is observed with Standard MVO under optimized rebalancing.
- The primary contribution of this work lies in architectural design and failure analysis rather than performance gains.

Future Work

While the current results highlight the limitations of the Enhanced MVO approach, future work will focus on:

- Exploring more robust risk calculation methods.

- Testing on live data to obtain a more accurate risk profile.
- Further optimizing the Multi-Branch LSTM architecture.
- Investigating the integration of additional technical indicators and fundamental data to enhance prediction accuracy.
- Comparing against simpler, more frequent rebalancing strategies to better understand trade-offs between complexity and performance.

Conclusion

The Multi-Branch LSTM approach, combined with mean-variance optimization, demonstrates that advanced machine learning techniques do not always translate to superior investment outcomes. In this case, a simpler Standard MVO strategy with a shorter rebalance frequency achieved far better results both in terms of annualized returns and risk-adjusted performance. This project underscores the importance of critically evaluating the practical benefits of model complexity and computational efficiency in portfolio management.