NYU

Bridging Dimensions in Portfolio Optimization: A Comparative Study of High-Dimensional Shrinkage and LSTM-Based Investor Views

Supervisor: Dr. James Adams

Presented by:
Ishit Bajpai
Aryan Singal
Arnav Sinha

High-Dimensional Data: Statistics or Deep Learning?

- **Problem Statement**: To Evaluate the trade-offs between high-dimensional portfolio optimization using shrinkage methods and reduced-dimensional approaches incorporating LSTM-based investor views.
- **Data Source**: Bloomberg. NSE200 time series data, January 2015 – March 2025.

Overview

- **Two-Phase Framework:** We compare high-dimensional shrinkage (Ledoit-Wolf) with reduced-dimensional optimization (using LSTM-generated investor views).
- **Investor Views:** Predictive signals (via Linear Regression or LSTM) feed into the Black-Litterman model to refine expected returns.
- **Portfolio Strategies:** BL is benchmarked against classic MVO and Equal-Weighted portfolios across both phases.
- **Evaluation Metrics:** Performance assessed using Sharpe Ratio, Drawdown, Volatility, and Rolling Indicators over 2018–2025.

Results

- Phase 1 (Linear Regression) models offered **quick and interpretable predictions** but suffered from limited complexity, often **underperforming in volatile regimes due to linear assumptions**.
- Phase 2 (LSTM) models captured **non-linear temporal dynamics**, leading to **better signal quality, improved Sortino ratios and Drawdown**, and more **stable performance** in BL portfolios.
- Black–Litterman portfolios benefited more in Phase 2, where **LSTM's adaptive views** blended well with market priors, especially in MS and MVO strategies, reducing overfitting risks seen in Phase 1.
- Equal-weighted baselines **remained stable across both phases** but **lacked responsiveness,** especially during market rebounds, a gap better closed in Phase 2.
- Phase 2 required **significantly more computational resources** due to model training complexity and larger data dependencies compared to the faster, **lightweight setup of Phase** 1.

NYU

Scope of the Project

- **Dataset:** Utilizes Close Price and MarketCap data from Bloomberg, *limited to static constituents* (153 Stocks) of the Nifty 200 over the study period.
- **Portfolio Strategy:** Adheres to a long-only approach - no short selling permitted.
- **Performance Metrics:** Focuses on annualized return, volatility, Sharpe ratio, Sortino ratio and maximum drawdown
- **Methodological Scope:** Limited to LSTM & Linear Regression for view generation and Ledoit-Wolf shrinkage for optimization, excluding alternative methods.
- **Dataset Scope**: Confined to a specific time frame (January 1, 2015, to March 18, 2025) and a single market (India), potentially limiting broader applicability.

NYU

Methodology Workflow

Step 1

Data: Closing Price and Market Cap for Nifty 200 constituents
Data Source: Bloomberg
Time Period: Jan 2015 – Mar 2025

Step 2

High-Dimensional Portfolio Optimization

- Ledoit-Wolf shrinkage for covariance.
- MVO, BL (linear regression views), Equal-Weighted; evaluate Sharpe, returns, volatility, drawdown.

Step 3

Reduced-Dimensional Optimization:

- LSTM for BL views, standard covariance.
- MVO, BL, Equal-Weighted; evaluate Sharpe, returns, volatility, drawdown.

Step 4

Comparative Analysis:

- Compare high- vs. reduced-dimensional portfolios.
- Analyze computational cost vs. performance trade-offs.

NYU

Originality Compared to Reference Papers

Aspect	Punyaleadtip et al. (2024)	Antonov (2016)	Our Approach
Core Focus	LSTM + SVR for BL view generation	ML classifiers (LogReg, SVM, NB, RF) for BL views	LSTM + clustering for reduced-dimensional BL views
Covariance Handling	Not emphasized	Not emphasized	Explicit comparison: Ledoit-Wolf shrinkage vs. LSTM BL
Dataset	Thailand SET, Dow Jones (2016–2022)	14 global futures indices, reduced to 8 (2014–2018)	Nifty 200 (India), decade-long data (2015–2025)
Market Scope	Multi-market, mid-size dataset	Global, broad but short-term dataset	Single-market, large-scale, long-term dataset
Novelty Contribution	Combines LSTM and SVR for BL view generation	Benchmarks multiple ML classifiers for BL view generation	First to compare Ledoit-Wolf (high-dimensional) vs. clustered LSTM (reduced-dimensional) in BL context
Innovation Angle	Investor view generation using RNN-SVR fusion	Exploratory classifier use in BL	Integrated dimensionality reduction + BL; single-market focus with in-depth temporal coverage

NYU

Portfolio Optimization techniques

- **Maximizing Sharpe** - Allocates more weight to assets with the highest return-to-risk ratio. In our project we applied this using both traditional and Black–Litterman predicted returns

$$\max_{w} \frac{w^\top \mu - r_f}{\sqrt{w^\top \Sigma w}} \quad \text{s.t.} \sum w_i = 1,\ w_i \in [0, 0.5]$$

- **Minimum Volatility** - Focuses purely on reducing portfolio risk, regardless of return. In our setup it is especially useful for investors prioritizing downside protection

$$\min_{w} \sqrt{w^\top \Sigma w} \quad \text{s.t.} \sum w_i = 1,\ w_i \in [0, 0.5]$$

- **Mean-Variance Optimization** - Balances expected return against portfolio risk using a risk-aversion parameter

$$\max_{w} \left(w^\top \mu - \lambda w^\top \Sigma w\right) \quad \text{s.t.} \sum w_i = 1,\ w_i \in [0.01, 0.4]$$

- **Equal Weighted Portfolio** - Allocates the same weight to all assets, ignoring return/risk estimates. It acts as a benchmark in our performance analysis

$$w_i = \frac{1}{N} \quad \text{for all } i$$

NYU

Black-Litterman

Core Idea

The Black-Litterman model blends:

- **Market Equilibrium Returns** (implied from market cap weights)
- **Investor Views** (our informed predictions for asset returns)

1. Mathematical Framework

$$\boxed{\pi = \delta \cdot \Sigma \cdot w_{\text{mkt}}}$$

→ What returns the market implies, based on current cap weights

2. Combined Posterior (BL) Returns

$$\boxed{\mu_{BL} = \left[(\tau\Sigma)^{-1} + P^{\top}\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\pi + P^{\top}\Omega^{-1}Q\right]}$$

→ Weighted average of **market beliefs** and **our views**

3. Optional Adjusted Covariance

$$\boxed{\Sigma_{BL} = \Sigma + \text{uncertainty adjustment}}$$

Market's Guess (Π)	Implied returns from market capitalization
Risk (Σ)	Asset variances and covariances from historical data
Our Guess (Q)	Expected returns for specific assets
Risk Comfort (δ)	Quantified risk aversion (trade-off between risk and return)
Trust Levels (τ, Ω)	Confidence in views (diagonal matrix)
View Matrix (P)	Identifies which assets are in views

NYU

Linear Regression Training

- **Training the model** - For our analysis, we first trained our model using data from 2015 to 2018, then updated its predictions daily by relearning from all past data up to each new day. This kept our model fresh and adaptable as we moved forward with the 153 Indian stocks (Nifty 200)

- **PCA** and **Feature Selection** – We started with 15 features, a mix of **Technical Indicators** and **Statistical Measures** derived from price data. Using **Principal Component Analysis,** we reduced the number of features to 6 and these explained 95% of all Variance in the model

- **Final Features** – Our final features were **Returns** (Daily percentage change), **5-day Moving Average**, **10-day Moving Average**, **5-day Volatility**, **3-day Momentum,** and **Ratio of Daily Close to its 5-day Moving Average**

- **Interpretability Advantage** – Unlike black-box models, Linear Regression offers full transparency: we can inspect the coefficients to understand which features (like momentum or volatility) are driving predicted price changes at any point in time.

NYU

Ledoit-Wolf Shrinkage

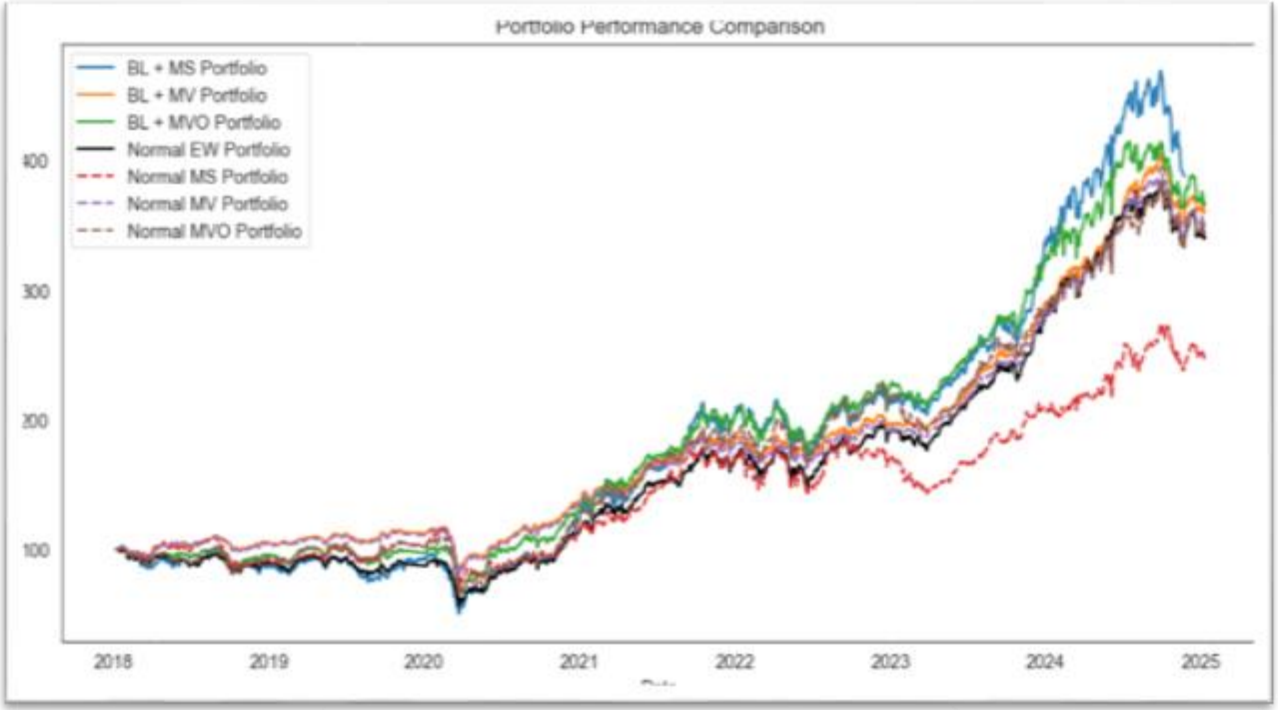
- **What is Ledoit-Wolf Shrinkage?**
  It's a regularization technique that **shrinks the noisy sample covariance matrix (S)** toward a more stable **structured estimator (F)** (usually the identity or constant correlation matrix):

$$\boxed{\hat{\Sigma}_{\text{shrink}} = \delta \cdot F + (1 - \delta) \cdot S}$$

  Where:
  - S = sample covariance matrix
  - F= shrinkage target
  - δ = shrinkage intensity, calculated to **minimize the mean squared error**

- **Why we used it:**
  Linear regression requires an accurate estimation of the **covariance matrix** when selecting features and calculating portfolio risk. But when working with 153 Indian stocks (Nifty 200 subset), the **sample covariance matrix becomes unstable** — especially for High-Dimensional data like ours. This can distort portfolio weights drastically**.**

Linear Regression Results



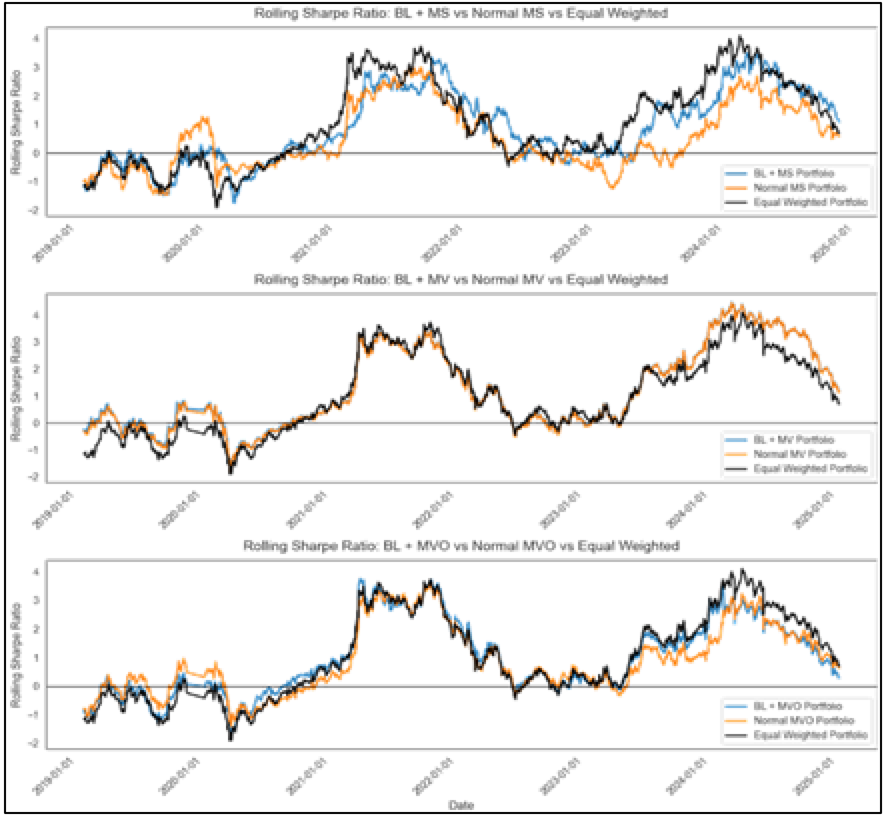
Linear Regression – Rolling Sharpe Comparison

- **Panel 1 (Top): BL + MS vs Normal MS vs Equal Weighted**
- BL + MS outperforms Normal MS, especially in the 2021–2024 rally, due to stable signal integration. Normal MS dips sharply in mid-2022, suggesting noisy or overconfident predictions. Equal Weighted stays competitive and leads in early recoveries.

- **Panel 2 (Middle): BL + MV vs Normal MV vs Equal Weighted**
- Both models perform well in bull phases, but BL + MV offers smoother Sharpe trends post-2022, reflecting better risk-adjusted blending. Equal Weighted lags during rallies, lacking alpha-driven responsiveness.

- **Panel 3 (Bottom): BL + MVO vs Normal MVO vs Equal Weighted**
- Early trends are similar, but BL + MVO outperforms post-2021, peaking higher in 2022–23 due to incorporation of informed priors. Equal Weighted is stable but underwhelms during sharp uptrends.

- **Cross-Observation:**
- BL-enhanced models maintain higher Sharpe ratios across most timeframes as compared to Non-BL enhanced models. Equal Weighted provides a stable baseline



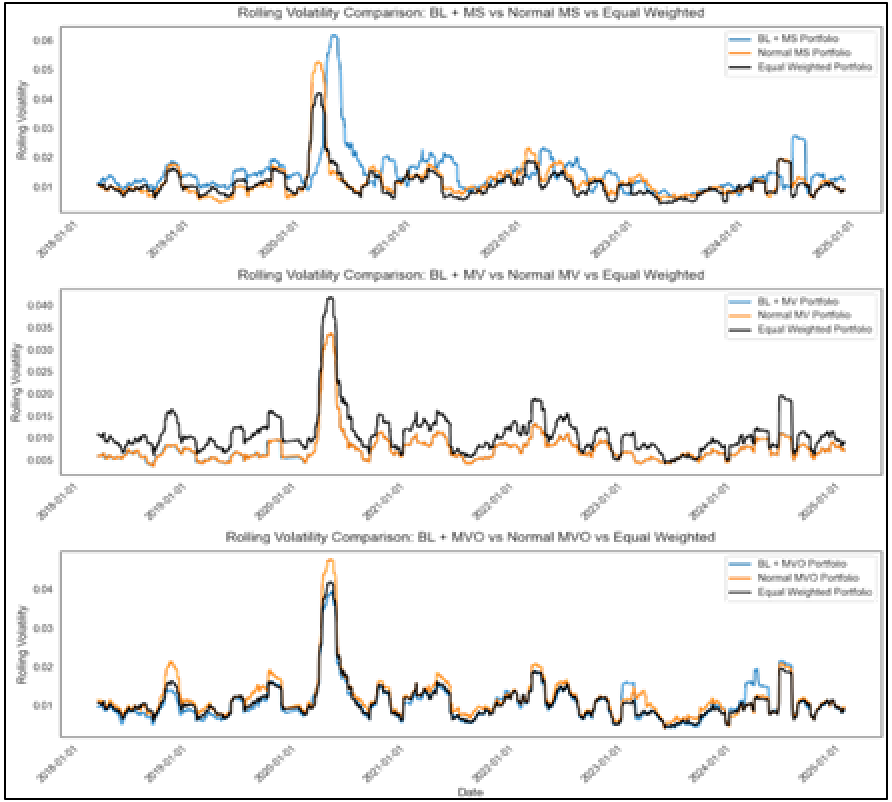
Linear Regression – Rolling Volatility Comparison

- **Panel 1 (Top): BL + MS vs Normal MS vs Equal Weighted**
- Normal MS and BL + MS exhibit sharper volatility spikes, especially during COVID-19 (2020) and late-2024. Equal Weighted volatility remains consistently less, revealing smoother sailing.

- **Panel 2 (Middle): BL + MV vs Normal MV vs Equal Weighted**
- BL + MV and Normal MV both show low volatility, fulfilling their objective by design. Equal Weighted portfolio experiences noticeable surges in 2020 and 2023, due to unoptimized asset exposure

- **Panel 3 (Bottom): BL + MVO vs Normal MVO vs Equal Weighted**
- All three strategies track closely during calm periods. BL + MVO outperforms in volatility control post-2020, especially during market rallies. However it shows slight spikes after 2023

- **Cross-Observation**
- Black–Litterman models exhibit smoother volatility trends, especially in MVO setups. Equal Weighted portfolios are consistently more volatile in MV and MVO. BL strategies show better shock-resilience by blending market consensus with model views except for MS



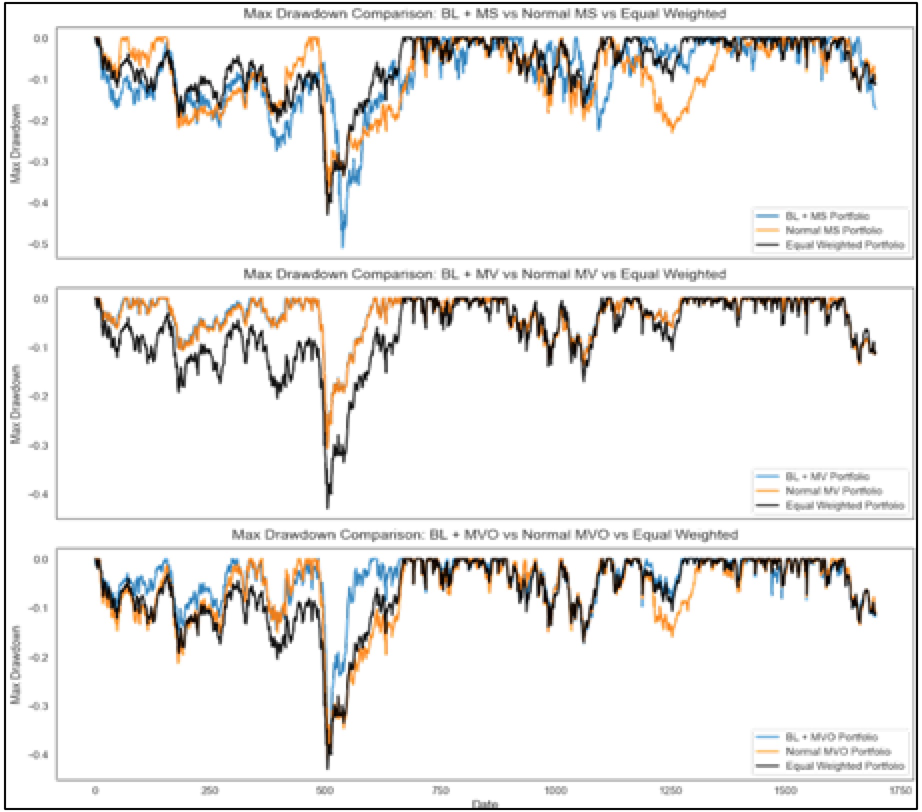
Linear Regression – Max Drawdown Comparison

- **Panel 1 (Top): BL + MS vs Normal MS vs Equal Weighted**
- BL + MS shows deeper drawdown in 2020, possibly due to aggressive allocation to predicted winners pre-crash. Post-2021, BL + MS recovers quickly and remains resilient, avoiding steep falls. Normal MS was inconsistent. Equal Weighted acts as a shock absorber in early periods.

- **Panel 2 (Middle): BL + MV vs Normal MV vs Equal Weighted**
- BL + MV and MV had consistently lower drawdowns, reflecting better risk control. EW dipped deeper during crises, reflecting adverse risk control

- **Panel 3 (Bottom): BL + MVO vs Normal MVO vs Equal Weighted**
- EW showed early volatility but stabilized strongly in late 2023–24. BL + MVO has slightly smoother recovery but suffers more frequent dips in volatile cycles. Normal MVO held mid-ground performance

- **Cross-Observation:**
- BL-enhanced strategies reduce extreme tail-risk except for MS, especially post-2021. BL models balance model confidence with market consensus, preventing aggressive overweights that hurt during crises. EW portfolios consistently fail to reduce losses and frequently underperform



Linear Regression Results

	Sharpe Ratio	Sortino Ratio	Volatility	Annualized Returns	Max Drawdown
BL + MS Portfolio	0.615067	0.762613	0.256708	0.222893	-0.511423
BL + MV Portfolio	1.068083	1.279617	0.135946	0.210201	-0.306586
BL + MVO Portfolio	0.781773	0.901885	0.188354	0.212250	-0.373999
Normal EW Portfolio	0.70544	0.814696	0.190345	0.199277	-0.430767
Normal MS Portfolio	0.366414	0.433833	0.21159	0.142529	-0.422044
Normal MV Portfolio	1.030201	1.233211	0.13596	0.205066	-0.307884
Normal MVO Portfolio	0.644446	0.738551	0.209716	0.200150	-0.410258

LSTM

Objective

- Predicted short-term (3-day) stock returns using LSTM models

$$\text{Investor's views} = \frac{(\text{Current value} - \text{Initial value})}{\text{Initial value}}$$

Training the model

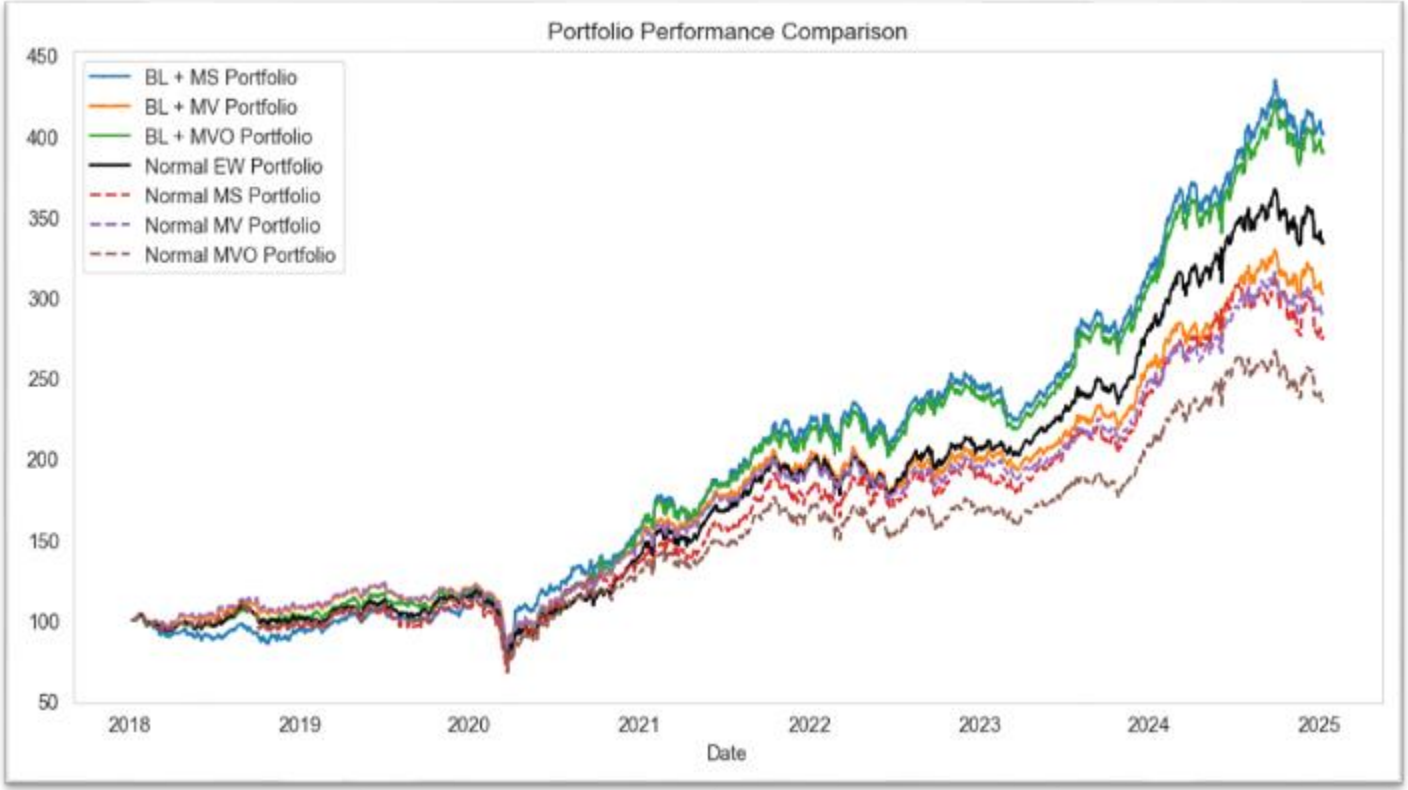
- Trained LSTM on subset of stocks from different sectors (based on MarketCap)
- Rolling forward in time with multiple splits (Train: 2.5 yrs, Validation: 0.5 years, Test: 1 year)
- Technical Indicators used :- SMA, EMA, RSI, Momentum

Model Parameters

Model Architecture & Hyperparameters

- 2 LSTM layers: 100 and 50 units
- Loss: Mean Squared Error
- Optimizer: Adam
- Epochs 100 learning rate 5e-3

LSTM Results



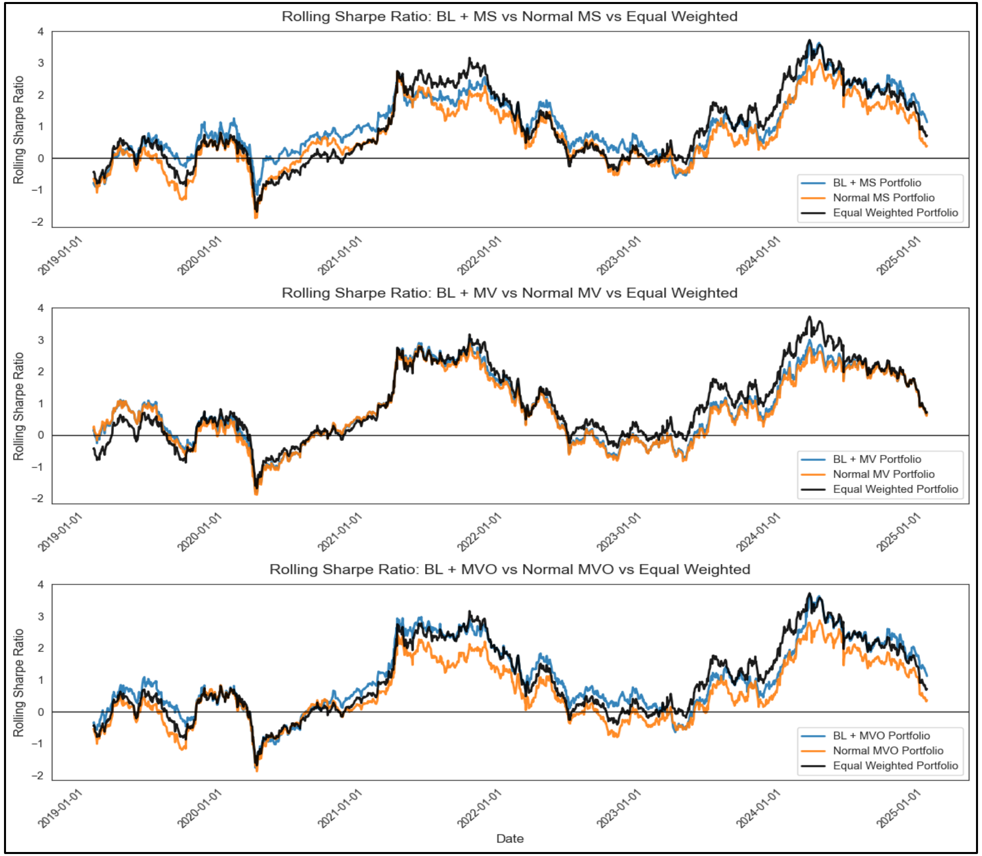
LSTM – Rolling Sharpe Comparison

- **Panel 1 (Top): BL + MS vs Normal MS vs Equal Weighted**
- BL + MS consistently outperforms Normal MS, especially post-2020 recovery. Equal Weighted performs competitively, showing resilience in volatile periods.

- **Panel 2 (Middle): BL + MV vs Normal MV vs Equal Weighted**
- BL + MV portfolio shows higher Sharpe stability, especially in mid-cycle phases. Normal MV lags slightly but follows similar trendline. Equal Weighted spikes noticeably in 2023–24, indicating diversification payoff.

- **Panel 3 (Bottom): BL + MVO vs Normal MVO vs Equal Weighted**
- BL + MVO sustains advantage, consistently ahead of Normal MVO. Peak outperformance observed in late 2023; subsequent convergence visible.

- **Cross-Observation:** Across all strategies, Equal Weighted portfolio provides a surprisingly competitive baseline, peaking during market rallies. BL-enhanced strategies systematically maintain higher Sharpe Ratios than normal ones, validating incorporation of investor views and improved covariance estimation.



NYU

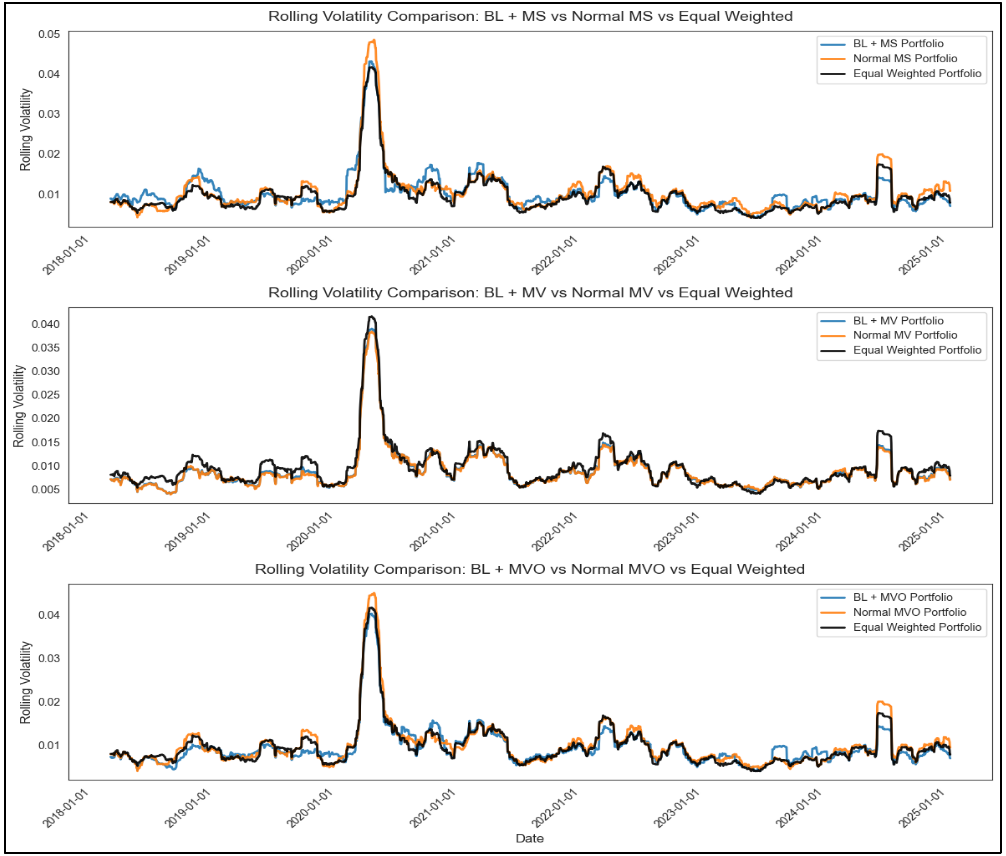
LSTM – Rolling Volatility Comparison

- **Panel 1 (Top): BL + MS vs Normal MS vs Equal Weighted**
- BL + MS portfolio exhibits slightly lower volatility vs. Normal MS. Equal Weighted Portfolio spikes most sharply during COVID crash, indicating higher exposure to market-wide stress.

- **Panel 2 (Middle): BL + MV vs Normal MV vs Equal Weighted**
- BL + MV consistently achieves similar volatility with Normal MV, showing risk efficiency. Volatility spike in March 2020 common across all, but Equal Weighted shows amplified reaction.

- **Panel 3 (Bottom): BL + MVO vs Normal MVO vs Equal Weighted**
- BL + MVO demonstrates volatility compression, especially post-2021, maintaining tighter band than peers. Equal Weighted portfolio maintains elevated volatility regime, reflecting lack of optimization.

- **Cross-Observation:** BL-enhanced portfolios consistently deliver marginal volatility reduction, confirming stabilizing impact of better covariance estimation. Equal Weighted and Non-BL portfolio exhibit higher volatility, underscoring diversification without optimization leads to higher risk.



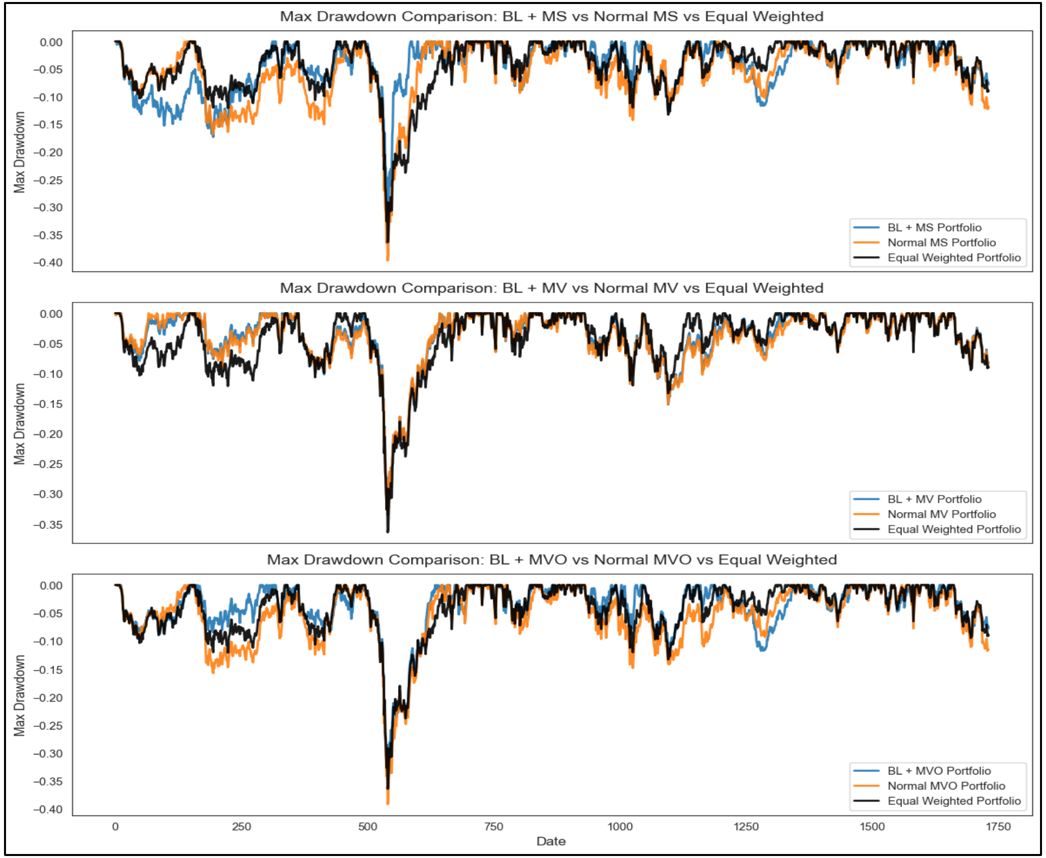
LSTM – Max Drawdown Comparison

- **Panel 1 (Top): BL + MS vs Normal MS vs Equal Weighted**
- BL + MS Portfolio demonstrates visibly shallower drawdowns compared to Normal MS and Equal Weighted, especially during COVID crash. Equal Weighted suffers the deepest drawdowns, highlighting vulnerability during crisis.

- **Panel 2 (Middle): BL + MV vs Normal MV vs Equal Weighted**
- BL + MV portfolio mitigates drawdowns similar to Normal MV, confirming resilience. During stress periods, Equal Weighted again faces steeper drawdowns, reflecting non-optimized risk exposure.

- **Panel 3 (Bottom): BL + MVO vs Normal MVO vs Equal Weighted**
- BL + MVO exhibits best relative drawdown control post-2020, maintaining tighter recovery profiles. EW follows closely, but Normal MVO shows clear vulnerability.

- **Cross-Observation:** BL-enhanced portfolios consistently perform better in terms of Maximum drawdown. Equal-Weighted and Normal both exhibit higher drawdowns, with Normal slightly worse.



NYU

LSTM Results

	Sharpe Ratio	Sortino Ratio	Volatility	Annualized Returns	Max Drawdown
BL + MS Portfolio	0.888495	1.194015	0.178708	0.223781	-0.316035
BL + MV Portfolio	0.923925	1.192465	0.166336	0.218682	-0.347433
BL + MVO Portfolio	0.70018	0.883733	0.156767	0.174765	-0.335145
Normal EW Portfolio	0.74689	0.91088	0.169599	0.191672	-0.364420
Normal MS Portfolio	0.498344	0.621483	0.187336	0.158358	-0.397528
Normal MV Portfolio	0.658758	0.840037	0.155083	0.167162	-0.334431
Normal MVO Portfolio	0.382679	0.468183	0.178323	0.133240	-0.391882

Balancing Trade-Offs in Portfolio Optimization

No One-Size-Fits-All Model

- **Risk-Adjusted Returns**: BL + MV (Phase 1) leads with Sharpe (1.068) & Sortino (1.280), ideal for steady risk-reward; BL + MS (Phase 1) also strong at Sharpe 0.615
- **Absolute Returns**: BL + MS (Phase 1) tops at 22.29% annualized returns, great for growth; BL + MVO (Phase 1) close at 21.23%
- **Downside Protection**: LSTM-based BL + MV (Phase 2) offers better drawdown control, reducing losses in tough markets

Data Coverage Insights

- **Phase 1 (Nifty 200):** Covers 153 Indian stocks (2015-2025), capturing broad market trends and diversity for stable risk estimates.
- **Phase 2 (Top 15 Stocks):** Focuses on 15 high-quality stocks, catching short-term trends and nonlinear patterns with more precision.

What We Learned:

- **Combine Strengths:** Use Phase 1’s broad stability for long-term balance, Phase 2’s focus for short-term gains in volatile periods
- **Tailor to Goals**: Choose based on priorities—risk (BL + MV), growth (BL + MS), or downside protection (LSTM portfolios)

THANK YOU

NYU