

# Developing Prediction-Based Portfolio Optimization Framework Using LSTM-Autoencoder and Worst-Case Omega Model

Kittanai Yamkleeb<sup>1</sup>, Prompong Sungunnasil<sup>2</sup>, and Sumalee Sangamuang<sup>\*2</sup>

<sup>1</sup> Master's Degree Program in Data Science, Chiang Mai University, Chiang Mai, Thailand

kittanai\_yam@cmu.ac.th

<sup>2</sup> College of Art, Media and Technology, Chiang Mai University, Chiang Mai, Thailand  
{prompong.sugunnasil,sumalee.sa}@cmu.ac.th

**Abstract.** Translating financial return predictions into effective portfolio decisions remains challenging due to the gap between predictive accuracy and investment performance. This study presents a prediction-to-decision framework that integrates deep learning-based return forecasting with Omega-based portfolio optimization. Using daily OHLCV data from 2018-2024, an autoencoder-long short-term memory (AE+LSTM) model is used to generate return forecasts, which are incorporated into a worst-case Omega allocation scheme to account for asymmetric return preferences. Forecasting performance is evaluated against autoregressive and neural baselines using both numerical error metrics (MAE, MSE) and directional measures (Hit Rate), while portfolio performance is assessed under consistent rebalancing rules with transaction costs and compared with equally weighted and mean-variance benchmarks. Out-of-sample backtesting across different market regimes examines annualized return, volatility, sharpe ratio, and maximum drawdown. The results suggest that differences in directional prediction behavior are associated with variations in portfolio-level outcomes under Omega-based allocation. In particular, models with more consistent directional patterns tend to provide a more balanced trade-off between return, risk, and turnover. Overall, the framework offers a systematic approach for examining how predictive signals translate into portfolio decisions across varying market conditions.

**Keywords:** Portfolio Optimization, Deep Learning for Finance, Directional Prediction, Omega Ratio, Financial Time Series.

## 1 Introduction

Global financial markets continue to exhibit persistent volatility driven by financial crises, geopolitical tensions, and shifts in monetary policy [3]. Traditional portfolio optimization methods, such as mean-variance and conditional value-at-risk frameworks, rely on assumptions of linear relationships and stable return

---

\* corresponding author

distributions. While these approaches form the foundation of modern portfolio theory, their performance often deteriorates under regime shifts and volatility clustering, where assumptions of normality and stationarity are violated, leading to unstable allocations and underestimated downside risk [2].

Recent advances in deep learning have introduced powerful tools for modeling financial time series. Architectures such as autoencoder and long short-term memory (LSTM) networks can capture nonlinear structures and temporal dependencies beyond the scope of traditional statistical models [1,4]. These models have demonstrated strong predictive performance in return-forecasting tasks. However, existing studies typically evaluate predictive models in isolation, without explicitly examining how prediction outputs translate into downstream portfolio decisions. As a result, improvements in predictive accuracy do not directly lead to improved investment performance [8,6].

This gap highlights a fundamental challenge: predictive accuracy alone is insufficient for decision-making. In particular, directional prediction behavior and error asymmetry may play a critical role in shaping portfolio outcomes, especially under volatile and asymmetric market conditions. Despite this, the relationship between predictive signals and allocation performance remains underexplored.

To address this issue, this study is guided by three research questions. First, it examines whether autoencoder-based feature extraction improves forecasting performance compared to autoregressive and standard LSTM models (RQ1). Second, it investigates whether incorporating predicted returns into a worst-case Omega (WCO) allocation framework [5] improves portfolio performance under realistic trading conditions (RQ2). Third, it evaluates whether the proposed approach maintains stable performance across different market regimes and transaction-cost settings, with particular attention to turnover and allocation behavior (RQ3).

To support this investigation, the proposed framework integrates autoencoder-based representation learning with LSTM-based return forecasting, followed by portfolio construction using the WCO model. This design enables the extraction of latent market structure, the generation of predictive signals, and the incorporation of asymmetric risk preferences within a unified pipeline.

The empirical evaluation is structured to directly address the research questions. Forecasting performance is assessed using standard errors and directional metrics, while portfolio performance is evaluated based on realized returns, volatility, and maximum drawdown under consistent rebalancing rules and transaction costs. Additional analysis across pre-crisis, crisis, and post-crisis periods provides insights into model behavior under varying market conditions. The results offer a systematic perspective on how predictive modeling interacts with portfolio decision-making in practice.

## 2 Literature Review

Portfolio optimization has developed from classical risk-return models toward approaches that incorporate prediction, nonlinear modeling, and alternative risk measures.

### 2.1 Traditional Portfolio Optimization

Classical portfolio models, such as the mean-variance (MV) framework, conditional value-at-risk (CVaR), and the Sharpe ratio, have been widely used for asset allocation. These methods rely on assumptions of linear relationships among assets and stable return distributions. While they work reasonably well in stable markets, their performance often deteriorates under regime changes and volatility clustering. In such cases, return distributions become heavy-tailed and time-varying, leading to unstable portfolio weights and an underestimation of downside risk [2, 3].

CVaR-based methods were introduced to improve downside risk control, but their performance still depends on distributional assumptions and model specification. The Sharpe ratio evaluates risk and return symmetrically and does not distinguish between upside and downside variation. To address this limitation, Kapsos et al. [5] proposed the Worst-Case Omega ratio, which accounts for unfavorable distribution scenarios during optimization. Although these extensions improve robustness, traditional models remain limited when market dynamics are nonlinear and rapidly changing.

### 2.2 Forecasting-Integrated Approaches

To improve adaptability, recent studies incorporate return forecasts into portfolio optimization. Instead of relying only on historical averages, these approaches use predicted returns to adjust portfolio weights over time. Ma et al. [6] proposed a framework that combines Autoencoders (AEs) and LSTM models for return prediction with Worst-Case Omega allocation, showing improved performance over mean-variance portfolios. Similarly, Martinez Barbero et al. [7] used LSTM-based forecasts within MV and CVaR optimization and reported better results in the S&P 500 dataset.

These studies show that prediction-based allocation can improve performance under changing market conditions. However, many approaches still use linear allocators, such as MV, and some do not explicitly account for asymmetric risk preferences. As a result, the potential benefit of combining prediction with more robust allocation models has not been fully explored.

### 2.3 Deep Learning for Return Forecasting

Deep learning models have been widely applied to financial time series due to their ability to capture nonlinear relationships and temporal patterns. Bao et al. [1] combined stacked Autoencoders (SAEs) with LSTM networks to denoise time series data and improve prediction accuracy, but performing ARIMA and shallow neural networks. Fischer and Krauss [4] showed that LSTM models outperform Random Forests and feedforward networks in predicting stock returns, particularly because they can model sequential dependencies.

Some studies extend deep learning models to portfolio-related tasks. Zhang et al. [8] used LSTM models to directly generate portfolio weights by optimizing the Sharpe ratio. Their results suggest that deep learning can improve both prediction and allocation performance. Overall, these studies confirm that deep learning is effective for extracting useful signals from financial data, although the connection between prediction and allocation is not always fully developed.

#### 2.4 Omega-Based Portfolio Optimization

The Omega ratio has been proposed as an alternative performance measure that separate upside gains from downside losses. Unlike the Sharpe ratio, which treats all volatility equally, Omega focuses on returns relative to a threshold and better reflects asymmetric investor preferences. Kapsos et al. [5] extended this idea to the Worst-Case Omega model, which optimizes portfolio weights under adverse distribution scenarios.

Recent work has combined prediction models with Omega-based allocation. Ma et al. [6] integrated Autoencoder and LSTM forecasts with Worst-Case Omega optimization and reported improved results compared to mean-variance and Sharpe-based portfolios. These results indicate that Omega-based allocation can be combined with predictive models to improve portfolio performance under uncertainty.

#### Summary of Related Work

Previous studies show three main directions. First, traditional portfolio models are limited when market conditions are unstable and returns are not normally distributed [2, 3]. Second, deep learning models such as Autoencoders and LSTM networks improve return prediction by capturing nonlinear patterns [1, 4]. Third, Omega-based optimization provides a more flexible way to handle asymmetric risk preferences [5].

However, most studies focus on only part of the pipeline. Some works improve prediction but do not consider allocation in detail, while others use prediction within allocation but rely on simple optimization models. In addition, evaluation across different market regimes and under transaction costs is not always consistent. This study builds on these directions by combining representation learning, prediction, and Omega-based allocation in a single framework and evaluating the results under multiple market conditions.

### 3 Research Contributions and Scope

This study presents a compact prediction-to-decision pipeline that connects deep forecasting with portfolio allocation and evaluates its performance under real market conditions. The analysis is structured around three research questions, covering prediction quality, decision-level impact, and model stability.

**RQ1: Does the proposed model improve forecasting performance?**

We develop an autoencoder-long short-term memory model and compare it with autoregressive baselines and a standard long short-term memory model without autoencoder features. The evaluation includes mean absolute error, mean square error, and directional hit rate. Results are reported with repeated runs, 95% confidence intervals, and statistical significance testing (see Sec.5.1).

**RQ2: Does improved forecasting translate into better portfolio performance?**

The forecasting outputs are integrated into a worst-case Omega portfolio strategy and compared with equally weighted under consistent rebalancing rules and transaction costs. The evaluation considers annualized return, volatility, Sharpe ratio, Omega ratio, and maximum drawdown (see Sec.5.2).

**RQ3: Is the proposed approach stable across different market conditions?**

We examine the behavior of the proposed framework across pre-crisis, crisis, and post-crisis regimes. Additional sensitivity and ablation studies are conducted on rebalancing frequency, window length, autoencoder representation size, and Omega parameters. Portfolio turnover and concentration are also analyzed (see Sec.5.3).

## 4 Research Framework and Methodology

This study proposes a prediction-based portfolio optimization framework that integrates representation learning, time-series forecasting, and portfolio allocation under uncertainty. The objective is to examine whether improvements in feature representation and prediction quality can be translated into consistent portfolio performance under realistic market conditions.

The framework is organized into three components: Representation Learning (Sec. 4.2), Return Prediction (Sec. 4.3), and Portfolio Optimization (Sec. 4.4). These components form a sequential pipeline, with each stage providing input to the next.

To enable structured analysis, the framework supports both component-level and system-level evaluation. This allows the study to examine how each component contributes to the final portfolio outcome and whether improvements in prediction performance are reflected in allocation decisions.

### 4.1 Overall Framework Design

This study proposes a prediction-based portfolio optimization framework illustrated in Figure 1. The framework follows a three-stage pipeline that transforms raw financial data into portfolio allocation decisions.

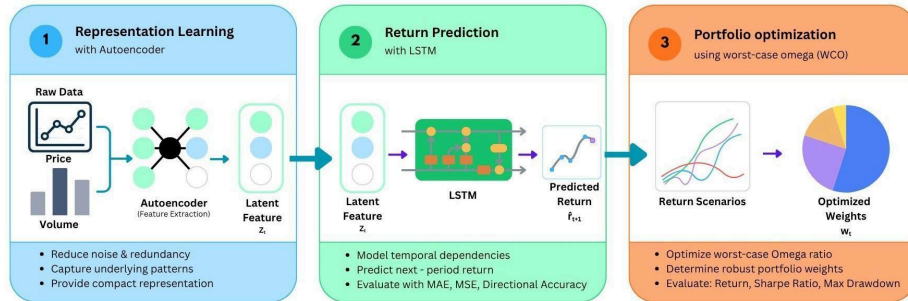


Fig. 1: Overall framework design of the proposed prediction-based portfolio optimization pipeline.

In the first stage, raw input data, including price and volume information are processed using Autoencoder to generate latent feature representations. This step reduces noise and extracts underlying structures from the data.

In the second stage, latent features are used as inputs to a Long Short-Term Memory (LSTM) model for return prediction. The model captures temporal dependencies and estimates the next-period return.

In the final stage, predicted returns are incorporated into a portfolio optimization model based on Worst-Case Omega (WCO) framework. This stage determines portfolio weights under uncertainty by considering adverse return scenarios.

This design allows both component-level and system-level analysis, enabling the study to examine how improvements in prediction quality propagate to portfolio decisions.

#### 4.2 Representation Learning via Autoencoder

In the first stage, Autoencoder is used to learn latent representations from historical price and volume data. These inputs may contain noise and redundancy, which can reduce the effectiveness of downstream prediction models.

Autoencoder is trained to reconstruct input data, where encoder maps input into a lower-dimensional latent space and decoder reconstructs the original signals. By minimizing reconstruction error, the model is expected to capture meaningful underlying patterns.

In the first stage, Autoencoder is used to learn latent representations from historical price and volume data. These inputs may contain noise and redundancy, which can reduce the effectiveness of downstream prediction models.

Autoencoder is trained to reconstruct input data, where encoder maps input into a lower-dimensional latent space and decoder reconstructs the original signals. By

minimizing reconstruction error, the model is expected to capture meaningful underlying patterns.

The results from latent features are used as inputs to the prediction model in the next stage. The effectiveness of this representation is evaluated by comparing prediction performance with and without the use of latent features.

#### 4.3 Return Prediction using LSTM

In the second stage, the Long Short-Term Memory (LSTM) model is used to predict future returns based on latent features. The LSTM is selected due to its ability to model sequential dependencies in financial time series.

The model takes the latent representation at time  $t$  as input and estimates the return at time  $t + 1$ . Prediction performance is evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and Directional accuracy (Hit Rate) is also considered to assess whether the model correctly predicts the direction of price movement. These metrics are particularly relevant for portfolio decision-making.

#### 4.4 Portfolio Optimization using Worst-Case Omega

In the final stage, predicted returns are used as inputs to a portfolio optimization model based on the Worst-Case Omega framework. The objective is to determine portfolio weights that remain effective under uncertain return distributions.

The WCO model evaluates a set of return scenarios and optimizes the worst-case Omega ratio. This approach emphasizes robustness and accounts for asymmetric return behavior.

Portfolio performance is evaluated using annualized return, Sharpe ratio, Omega ratio, and maximum drawdown. This stage examines whether predictive signals can be translated into practical investment decisions.

#### 4.5 Evaluation Strategy

The evaluation is designed to directly address the three research questions of this study. At the component level (RQ1), the focus is on the effectiveness of representation learning and prediction performance. Models using latent features are compared with those using original input features. Prediction performance is evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and directional metrics which are Hit Rate (HR), positive Hit Rate (HR+), and negative Hit Rate (HR-). These methods help in assessing the ability of a model to correctly predict the direction of return value. Directional metrics are included because correctly predicting return direction is often more relevant for portfolio decisions than minimizing numerical error alone.

At the system level (RQ2), the evaluation focuses on portfolio performance. The constructed portfolios are assessed using annualized return, volatility, Sharpe ratio, Omega ratio, turnover, and maximum drawdown. Baseline strategies include LSTM without representation learning, equally weighted portfolios, and buy and hold.

To address RQ3, the framework is evaluated across different market regimes: pre-crisis, crisis, and post-crisis. A sensitivity analysis is conducted to examine the impact of key parameters, such as the input window length and model configuration.

Overall, this evaluation design links prediction performance to portfolio outcomes in a consistent and empirically grounded manner.

#### 4.6 Experimental Design

The experiments are conducted using historical daily data from S&P 500 sector indices, including price and volume information. The raw data are preprocessed to construct rolling input sequences, where each input window represents the market state at time  $t$ , and the corresponding next-period return is used as the prediction target. All input features are normalized prior to model training.

A rolling-window evaluation scheme is adopted to simulate realistic investment conditions and to avoid look-ahead bias. In each evaluation round, models are trained on a fixed-length historical window and evaluated on the subsequent period. This process is repeated across the dataset to ensure consistent out-of-sample testing.

The proposed framework is implemented as a sequential pipeline. First, Autoencoder is trained to learn latent feature representations from the input data. These latent features are then used as inputs to the LSTM model for return prediction. Predicted returns are subsequently passed to the portfolio optimization stage, where portfolio weights are determined using the corresponding allocation strategy.

At each rebalancing point, portfolios are constructed using predicted returns from each model. Transaction costs are incorporated during rebalancing to reflect realistic trading conditions. All strategies, including the proposed method and baselines, are evaluated under identical data splits, rebalancing frequency, and portfolio constraints to ensure fair comparison.

A sensitivity analysis is conducted to examine the effects of key parameters, including the input window length and model configuration. In addition, the evaluation is performed consistently across all experimental settings to ensure comparability of results.

This experimental setup provides a controlled, reproducible, and practically relevant environment for evaluating the proposed framework.

## 5 Results and Experimental Evaluation

This section reports the experimental results of the proposed prediction-to-decision framework. The evaluation is organized into three components: (i) predictive performance, (ii) portfolio-level performance under an Omega-based allocation strategy, and (iii) risk characteristics and robustness, including drawdown behavior.

All experiments are conducted under a rolling-window setting to approximate realistic deployment conditions. Results are presented in a structured manner corresponding to the defined evaluation metrics and experimental setup.

Interpretation of the observed patterns and their implications for decision making are provided in Sec.6.

### 5.1 Experimental Design

Table 1: Predictive performance comparison across models (2018–2024) in terms of numerical accuracy (MAE, MSE) and directional behavior (HR, HR+, HR-), revealing asymmetric prediction characteristics across models.

Year	Model	MAE	MSE	HR (%)	HR- (%)	HR+ (%)
2018	AE+LSTM	<b>0.0133</b>	<b>0.0004</b>	<b>52.30</b>	16.58	<b>84.08</b>
	LSTM	0.0134	0.0004	50.80	<b>46.81</b>	54.35
	1D-CNN	0.0134	0.0004	50.70	28.87	70.12
	ARIMA	0.0134	0.0004	51.86	28.65	72.51
2019	AE+LSTM	<b>0.0115</b>	<b>0.0003</b>	<b>53.71</b>	22.29	<b>78.64</b>
	LSTM	0.0116	0.0003	51.66	36.28	63.86
	1D-CNN	0.0117	0.0003	51.68	<b>46.18</b>	54.04
	ARIMA	0.0116	0.0003	51.47	36.82	63.11
2020	AE+LSTM	<b>0.0203</b>	<b>0.0010</b>	<b>51.68</b>	22.50	<b>78.63</b>
	LSTM	0.0204	0.0010	50.87	<b>41.13</b>	59.86
	1D-CNN	0.0205	0.0010	51.01	31.81	68.76
	ARIMA	0.0204	0.0010	51.37	33.50	67.88
2021	AE+LSTM	<b>0.0128</b>	<b>0.0003</b>	<b>52.07</b>	20.72	<b>79.68</b>
	LSTM	0.0130	0.0003	51.79	28.26	72.51
	1D-CNN	0.0131	0.0003	50.07	<b>51.57</b>	48.75
	ARIMA	0.0129	0.0003	51.88	37.51	64.54
2022	AE+LSTM	<b>0.0175</b>	<b>0.0006</b>	<b>49.48</b>	16.24	<b>83.54</b>
	LSTM	0.0176	0.0006	49.72	41.95	57.68
	1D-CNN	0.0177	0.0006	49.75	<b>42.81</b>	56.86
	ARIMA	0.0178	0.0006	49.30	36.59	62.31
2023	AE+LSTM	<b>0.0121</b>	<b>0.0003</b>	<b>51.29</b>	28.73	<b>71.29</b>
	LSTM	0.0123	0.0003	51.01	49.30	52.52
	1D-CNN	0.0122	0.0003	49.59	<b>51.74</b>	47.68
	ARIMA	0.0123	0.0003	50.53	36.21	63.23
2024	AE+LSTM	<b>0.0129</b>	<b>0.0004</b>	<b>51.16</b>	28.07	<b>71.66</b>
	LSTM	0.0131	0.0004	50.38	40.92	58.78
	1D-CNN	0.0132	0.0004	49.88	<b>49.71</b>	50.04
	ARIMA	0.0130	0.0004	50.88	38.63	61.76

This subsection reports the predictive performance of all models on unseen data, focusing on both numerical accuracy and directional behavior.

**Point Forecast Accuracy** from Table 1 summarizes the forecasting results across models from 2018 to 2024 in terms of MAE and MSE.

Across all evaluation periods, the AE+LSTM model yields the lowest MAE even though MSE values are similar to each other. During 2020, The error value was increased for all models, reflecting the higher market volatility during that time. Despite this increase, the relative ranking of models remains consistent across years.

The reported results provide a comparative view of numerical prediction performance across models. Further interpretation of these patterns is provided in Sec. 6.

**Directional Accuracy** from Table 1 also reports directional performance in terms of HR, HR<sup>+</sup>, and HR<sup>-</sup>. The AE+LSTM model achieves the highest overall hit rate (HR) in most evaluation years. In addition, the model exhibits higher values in HR<sup>+</sup> compared to other models across all periods. In contrast, HR<sup>-</sup> values are generally lower for AE+LSTM than for baseline models such as LSTM and 1D-CNN.

These results indicate differences in directional prediction behavior across models, particularly between positive and negative return prediction. A detailed analysis of these patterns and their implications for portfolio construction is presented in Sec. 6.

## 5.2 Portfolio Performance Evaluation

Table 2: Portfolio performance comparison across allocation strategies and fore-casting models.

Strategy	Return (%)	Volatility (%)	Max DD (%)	Turnover(%)
AE + LSTM + WCO	<b>455.44</b>	28.21	-36.02	348.85
AE + LSTM + EW	452.25	27.75	-34.96	354.06
LSTM + WCO	253.13	27.52	-40.07	545.12
LSTM + EW	256.35	27.29	-39.17	544.4
1D-CNN + WCO	330.69	28.00	-36.85	813.74
1D-CNN + EW	299.43	27.87	-36.53	796.13
ARIMA + WCO	261.58	29.64	-43.23	1653.26
ARIMA + EW	259.81	29.48	-43.79	1639.34
Buy & Hold (Benchmark)	272.32	<b>21.11</b>	<b>-33.64</b>	0.00

This subsection reports portfolio-level results based on each model's predicted returns. All strategies are evaluated under identical conditions, including the same dataset (2018–2024), the same rebalancing frequency, and the same transaction cost assumptions.

**Portfolio Performance Comparison** from Table 2 summarizes the portfolio performance across models and allocation strategies.

Across all configurations, the AE+LSTM+WCO strategy achieves the highest cumulative return (455.44%), indicating its effectiveness in translating

predictive signals into allocation decisions. This performance is slightly higher than AE+LSTM+EW (452.25%), suggesting that the WCO framework provides a marginal improvement over equal weighting when combined with strong predictive features.

In comparison, baseline models yield substantially lower returns. LSTM based strategies produce cumulative returns around 253 to 256%, while 1D-CNN and ARIMA strategies exhibit mixed performance, with returns ranging from approximately 259% to 330%. These results indicate that the choice of predictive model significantly affects portfolio-level outcomes.

Turnover varies considerably across strategies. AE+LSTM-based portfolios maintain moderate turnover (approximately 350%), whereas ARIMA-based strategies exhibit extremely high turnover (above 1600%), reflecting unstable allocation behavior and frequent rebalancing.

Buy & Hold benchmark achieves a cumulative return of 272.32 %, which is lower than most learning-based strategies despite this method having the lowest risk level.

The bolded values in the table highlight the best-performing metric in each column. In particular, The highest return achieved by AE+LSTM+WCO reflects its superior ability to capture profitable signals, while the lowest volatility and drawdown observed in the Buy & Hold strategy indicate its inherently conservative and stable behavior. This contrast illustrates the trade-off between maximizing returns and controlling risk across different strategies. A detailed interpretation of these results is provided in Sec. 6.

**Risk Metrics and Drawdown** from Table 2 also reports volatility and maximum drawdown across strategies.

AE+LSTM-based portfolios exhibit volatility around 28%, which is higher than the Buy & Hold benchmark (21.11%) but comparable to other learning-based strategies. This indicates that the improved return performance is associated with increased exposure to market fluctuations.

Maximum drawdown varies across models. Buy & Hold strategy exhibits the lowest drawdown (-33.64%), reflecting its passive and stable nature. In contrast, AE+LSTM+WCO strategy experiences a larger drawdown (-36.02%), suggesting higher downside risk despite its superior returns.

ARIMA-based portfolios show the highest drawdowns (around -43%), indicating weaker risk control under volatile conditions. Similarly, high turnover strategies tend to have larger drawdowns, suggesting that frequent rebalancing may amplify downside exposure.

Overall, these results highlight a trade-off between return performance and risk stability. While AE+LSTM+WCO archives the highest return, it does so with moderately higher volatility and drawdown than the benchmark, suggesting a more aggressive allocation strategy.

### 5.3 RegimeBased Evaluation

Table 3-6 represent the performance of all strategies across three distinct market regimes: pre-crisis, crisis, and post-crisis, enabling a regime-aware comparison of return, risk-adjusted performance, tail behavior, and downside risk,

In terms of total return (Table 3), ARIMA+EW has the highest return in the pre-crisis period (87.25 %), indicating that simpler statistical models can perform strongly under stable market conditions. During the crisis regime, 1D-CNN+WCO delivered the highest return (152.08%), suggesting that convolutional models are more responsive to abrupt market movements and local temporal patterns. In the post-crisis period, LSTM+EW models are more responsive to abrupt market movements and local temporal patterns. In the post-crisis period, LSTM+EW has the highest return (61.67%), demonstrating its ability to capture longer-term temporal dependencies during market recovery. In contrast, AE+LSTM-based strategies exhibit relatively stable returns across all regimes, without dominating in any single regime, indicating consistent but less specialized behavior.

Risk-adjusted performance, measured by the Sharpe ratio (Table 4), reveals a similar regime dependent pattern. ARIMA+EW has the highest Sharpe ratio in the pre-crisis period (1.4987), This shows efficient returns during low volatility. During the crisis regime, 1D-CNN+EW achieves the highest Sharpe ratio (1.5708), indicating a strong return relative to risk during turbulent conditions. In the post-crisis regime, LSTM+EW has the highest Sharpe ratio (0.7394), suggesting improved efficiency during market stabilization. AE+LSTM-based strategies maintain moderate but consistent Sharpe ratios across all regimes, indicating balanced performance without extreme sensitivity to specific market conditions.

The Omega ratio results (Table 5) provide further insight into asymmetric return behavior. ARIMA+EW achieves the highest Omega ratio in the pre-crisis regime (1.2923), indicating favorable gain-to-loss characteristics and understandable conditions. During the crisis regime, 1D-CNN+WCO achieves the highest Omega ratio (1.3554), suggesting stronger upside potential relative to downside risk during extreme market fluctuations. In the post-crisis regime, LSTM+WCO has the highest Omega ratio (1.1275), reflecting improved tail performance during recovery. AE+LSTM-based strategies again exhibit stable and moderate Omega values across regimes, indicating robustness without extreme tail sensitivity.

In terms of downside risk, measured by maximum drawdown (Table 6), Buy & Hold has the lowest drawdown in both pre-crisis (-17.61%) and post-crisis (-23.35%) periods, This shows its passive and conservative nature. During the crisis regime, 1D-CNN+WCO revealed the lowest drawdown (-30.73%), suggesting greater resilience to extreme volatility than other learning-based models. In contrast, ARIMA-based strategies experience the largest drawdowns during crisis periods (around -42%), indicating vulnerability to sudden market shocks.

Overall, the results show a clear regime-dependent trade-off. Simpler models, such as ARIMA, perform well under stable conditions but deteriorate under market stress. Deep learning models such as 1D-CNN and LSTM exhibit stronger performance

during volatile or transitional regimes, where nonlinear patterns are stronger. The AE+LSTM-based hybrid approach does not dominate any single regime but provides consistently stable performance across all regimes, suggesting robustness to both structural market changes and varying volatility levels.

Differences in performance across regimes and their implications are further discussed in Sec. 6.

Table 3: Total return (%) across different market regimes.

Strategy	Pre-Crisis	Crisis	Post-Crisis
AE+LSTM+WCO	25.00	144.29	55.73
AE+LSTM+EW	27.39	138.07	55.82
LSTM+WCO	37.00	84.95	57.46
LSTM+EW	38.62	85.68	<b>61.67</b>
CNN+WCO	35.44	<b>152.08</b>	34.90
CNN+EW	30.89	142.46	32.62
ARIMA+WCO	84.45	78.92	50.89
ARIMA+EW	<b>87.25</b>	81.64	51.18
Buy&Hold	47.49	81.57	39.03

## 6 Discussion and Analysis

This section interprets the empirical findings by examining how predictive signals are translated into portfolio decisions under different modeling and market conditions. Rather than evaluating forecasting and allocation independently, the analysis focuses on their interaction, with particular emphasis on directions consistency, allocation sensitivity, and regime-dependent behavior.

### 6.1 Directional Asymmetry as a Driver of Portfolio Outcomes (RQ1 → RQ2)

The results demonstrate that predictive performance differs not only in numerical accuracy but also in directional structure. While AE+LSTM achieves competitive MAE and MSE, its primary advantage lies in producing more consistent directional signals, particularly in identifying positive return movements.

This finding establishes that directional consistency, rather than numerical accuracy alone, is a primary determinant of portfolio performance under asymmetric allocation mechanism. In the WCO framework, which emphasized downside protection and asymmetric return preferences, correctly identifying upward movements contributes disproportionately to cumulative gains.

As a result, models that generate stable directional patterns can outperform alternatives even when differences in numerical error are marginal. This highlights a

limitation of traditional evaluation metrics such as MAE and MSE, which do not capture the decision relevance of predictive outputs. Instead, predictive models should be evaluated based on how effectively their directional signals align with the allocation objective.

## 6.2 Allocation as a Nonlinear Amplifier of Predictive Signals

The results in Sec.5.2 show that portfolio performance is jointly determined by both the forecasting model and the allocation mechanism. In particular, the WCO-based allocation consistently enhances performance when combined with directionally consistent predictive signals, as observed in the AE+LSTM configurations.

Table 4: Sharpe ratio across different market regimes.

Strategy	Pre-Crisis	Crisis	Post-Crisis
AE+LSTM+WCO	0.6486	1.5020	0.6488
AE+LSTM+EW	0.6994	1.4671	0.6556
LSTM+WCO	0.9649	1.0408	0.7026
LSTM+EW	1.0075	1.0537	<b>0.7394</b>
CNN+WCO	0.7146	1.6323	0.5256
CNN+EW	0.6539	<b>1.5708</b>	0.5044
ARIMA+WCO	1.4468	0.9974	0.6842
ARIMA+EW	<b>1.4987</b>	1.0138	0.6867
Buy&Hold	1.2196	1.2231	0.6911

Note: The Sharpe ratio measures risk-adjusted return by normalizing excess returns with respect to their standard deviation. Higher values indicate more efficient return generation per unit of risk. Bold values denote the best-performing strategy within each market regime.

This behavior indicates that allocation strategies act as nonlinear amplifiers of predictive signals. When predictions are directionally aligned with market movements, the allocation mechanism magnifies gains. However, when directional predictions are inconsistent, the same mechanism can amplify noise, leading to increased volatility and unstable portfolio behavior.

Therefore, the effectiveness of an allocation strategy cannot be evaluated independently of the predictive model. Portfolio outcomes emerge from the interaction between signal quality and allocation sensitivity, emphasizing the importance of jointly designing prediction and decision components.

## 6.3 Trade-off between Return, Risk, and Turnover

The empirical results reveal a clear and systematic trade-off between return risk, and turnover across strategies, as summarized in Table 2 and Table 7. High-return

strategies, particularly those driven by adaptive models, often incur higher turnover or increased downside exposure.

As shown in Table 7, ARIMA-based strategies exhibit extremely high turnover across all regimes, reaching approximately 900-1000% annually. This reflects unstable allocation behavior and frequent rebalancing. Although these strategies achieve moderate returns (Table 2), their excessive turnover implies substantial transaction costs and sensitivity to noise, which limits their practical applicability.

Table 5: Omega ratio across different market regimes.

Strategy	Pre-Crisis	Crisis	Post-Crisis
AE+LSTM+WCO	1.1178	1.3137	1.1123
AE+LSTM+EW	1.1278	1.3077	1.1135
LSTM+WCO	1.1793	1.2163	<b>1.1275</b>
LSTM+EW	1.1871	1.2204	1.1350
CNN+WCO	1.1334	<b>1.3554</b>	1.0931
CNN+EW	1.1214	1.3424	1.0891
ARIMA+WCO	1.2808	1.2133	1.1227
ARIMA+EW	<b>1.2923</b>	1.2179	1.1231
Buy&Hold	1.2388	1.2875	1.1250

Note: The Omega ratio evaluates performance by considering the entire return distribution, capturing both upside potential and downside risk. Higher values indicate a more favorable balance between gains and losses under asymmetric return conditions. Bold values denote the best-performing strategy within each market regime.

In contrast, AE+LSTM-based strategies maintain significantly lower turnover, with values ranging from approximately 165% to 312% across regimes (Table 7). These values are consistently among the lowest non-trivial turnover level (excluding the Buy & Hold benchmark), as highlighted in bold. This indicates more stable allocation adjustments while still achieving the highest cumulative return (455.44%) in Table 2. This combination suggests a more efficient balance between return generation and implementation cost.

From a risk perspective, Table 2 shows that the Buy & Hold strategy achieves the lowest volatility (21.11%) and lowest maximum drawdown (-33.64%), as indicated by the bold values. This reflects its passive and highly stable nature. However, this stability comes at the cost of lower return performance compared to learning-based strategies.

Among adaptive models, AE+LSTM+WCO has the highest return but with moderately higher volatility (28.21%) and moderately high drawdown (-36.02%). Importantly, its drawdown remains significantly lower than ARIMA-based (around -43%), indicating better downside control despite higher return targeting.

The bold values in Table 7 further highlight that AE+LSTM-based strategies consistently get lower turnover than other learning-based models, while Buy & Hold trivially has zero turnover due to the absence of rebalancing. This contrast emphasizes that the proposed method reaches a practical middle ground between passive stability and highly reactive strategies.

Overall, these results highlight a fundamental trade-off: strategies that aggressively pursue return tend to increase turnover and risk exposure, whereas conservative strategies reduce risk at the expense of return. The proposed AE+LSTM+WCO framework occupies an intermediate position, offering a balanced profile with strong return, controlled risk, and moderate turnover.

Table 6: Max DD (%) across different market regimes.

Strategy	Pre-Crisis	Crisis	Post-Crisis
AE+LSTM+WCO	-22.24	-33.71	-37.03
AE+LSTM+EW	-22.26	-33.19	-35.39
LSTM+WCO	-21.00	-36.81	-38.26
LSTM+EW	-20.49	-37.16	-37.68
CNN+WCO	-35.32	<b>-30.73</b>	-30.71
CNN+EW	-35.08	-31.46	-31.22
ARIMA+WCO	-22.16	-42.92	-29.10
ARIMA+EW	-21.65	-42.89	-28.95
Buy&Hold	<b>-17.61</b>	-33.64	<b>-23.35</b>

This trade-off also reflects a broader limitation of highly adaptive strategies, where increased responsiveness often comes at the cost of higher transaction overhead and instability. Therefore, performance evaluation should consider not only return but also risk exposure and operational feasibility.

#### 6.4 Regime-Dependent Behavior and Robustness (RQ3)

The regime-based analysis reveals that model performance is strongly dependent on market conditions. ARIMA-based strategies perform well in stable, pre-crisis environments, where linear dependencies dominate. In contrast, 1D-CNN-based models exhibit superior performance during crisis periods, suggesting sensitivity to abrupt market changes. LSTM-based strategies perform more effectively during post-crisis recovery, capturing longer-term temporal dependencies.

Table 7: Annual turnover (%) across different market regimes.

Strategy	Pre-Crisis	Crisis	Post-Crisis
AE+LSTM+WCO	<b>165.66</b>	<b>275.00</b>	<b>312.06</b>
AE+LSTM+EW	<b>172.11</b>	<b>271.43</b>	<b>306.79</b>
LSTM+WCO	389.04	292.50	367.43

LSTM+EW	387.25	300.00	364.31
CNN+WCO	512.03	455.00	543.60
CNN+EW	494.82	457.14	536.88
ARIMA+WCO	996.45	920.00	983.17
ARIMA+EW	996.81	900.00	977.90
Buy&Hold	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>

The AE+LSTM hybrid approach shows a distinct and consistent behavior. Rather than achieving peak performance in any single regime, it maintains competitive results across all regimes, including stable, volatile, and recovery periods.

This behavior can be interpreted as regime-agnostic stability, where performance consistency is preserved despite structural changes in market conditions. Such robustness is particularly valuable in real-world settings, where future regimes are uncertain and model adaptability must be balanced with stability.

These findings suggest that while regime specialization exists across models, robustness across regimes may be more desirable than peak performance under a single condition.

### 6.5 Summary of Findings

Overall, the results demonstrate that predictive quality should be evaluated beyond traditional accuracy metrics. Directional consistency plays a central role in determining portfolio outcomes, particularly when combined with allocation strategies that emphasize asymmetric return preferences.

Portfolio performance is not determined solely by the predictive model or the allocation method, but by their interaction. Allocation mechanisms can amplify both useful signals and prediction errors, underscoring the importance of alignment between prediction outputs and decision objectives.

A clear trade-off exists between return, risk, and turnover. While some strategies achieve high returns under specific conditions, they often incur higher transaction costs or greater downside risk. In contrast, the AE+LSTM+WCO framework provides a more balanced profile, achieving competitive returns with moderate turnover and stable risk characteristics.

Finally, the regime-based analysis highlights that no model is universally optimal. However, the proposed framework demonstrates consistent behavior across different market conditions and reduced sensitivity to stochastic variation.

These findings establish that, in financial decision-making under uncertainty, stability and directional consistency are more critical than marginal improvements in predictive accuracy or regime-specific performance.

## 7 Conclusion and Future Work

This study proposed a prediction-to-decision framework that integrates representation learning, sequence modeling, and Omega-based portfolio optimization for financial time series. The results demonstrate that predictive performance should not be evaluated solely based on numerical accuracy, but also on directional consistency, as these characteristics directly influence portfolio-level outcomes under asymmetric allocation mechanisms.

The empirical findings show that no single model consistently dominates across all market regimes. Instead, model effectiveness is strongly regime-dependent. While ARIMA-based strategies perform well under stable conditions, learning-based models such as 1D-CNN and LSTM exhibit advantages during volatile or transitional periods. The proposed AE+LSTM+WCO framework does not achieve the highest performance in any single regime, but consistently maintains competitive results across all regimes. This behavior indicates improved robustness to both changes in market conditions and stochastic variation during model training.

A clear trade-off is observed between return, risk, and turnover. Strategies that achieve high returns often incur higher turnover or greater downside risk, which may limit their practical applicability. In contrast, the proposed framework offers a more balanced performance profile, delivering competitive returns with moderate turnover and relatively stable risk.

Overall, this study highlights that effective financial modeling should consider not only predictive accuracy, but also decision-level properties such as directional alignment, stability across regimes, and implementation constraints. These factors are essential for translating predictive signals into reliable investment decisions under non-stationary market conditions.

Future work will focus on improving downside risk control within the proposed framework. While the current approach delivers strong returns, its ability to mitigate drawdowns remains limited under extreme market conditions. One promising direction is to incorporate downside-aware signals, such as negative return prediction or tail-risk indicators, into both the forecasting and allocation stages. In addition, extending the framework with risk-aware learning objectives, adaptive allocation mechanisms, or regime-sensitive strategies may further enhance robustness and practical applicability. These extensions aim to achieve a more effective balance between return generation and risk management in real world portfolio applications.

## References

1. Bao, W., Yue, J., Rao, Y.: A deep learning framework for financial time series using stacked autoencoders and long short-term memory. *PLOS ONE* **12**(7), e0180944 (2017). <https://doi.org/10.1371/journal.pone.0180944>

2. Cheng, P., Young, S., Stefan, M.: Portfolio optimization on multivariate regime switching garch model with normal tempered stable innovation. *Journal of Risk and Financial Management* **15**(5), 230 (2022). <https://doi.org/10.3390/jrfm15050230>
3. European Central Bank: Ecb economic bulletin, box article april 2025. *ECB Economic Bulletin* (2025), [https://www.ecb.europa.eu/press/economic-bulletin/focus/2025/html/ecb.ebbox202504\\_05~2dc91bb9e3.en.html](https://www.ecb.europa.eu/press/economic-bulletin/focus/2025/html/ecb.ebbox202504_05~2dc91bb9e3.en.html)
4. Fischer, T., Krauss, C.: Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research* **270**(2), 654–669 (2018). <https://doi.org/10.1016/j.ejor.2017.11.054>
5. Kapsos, M., Christofides, N., Rustem, B.: Worst-case robust omega ratio. *European Journal of Operational Research* **234**(2), 499–507 (2014). <https://doi.org/10.1016/j.ejor.2013.10.030>
6. Ma, Y., Wang, W., Ma, Q.: A novel prediction based portfolio optimization model using deep learning. *Computers & Industrial Engineering* **177**, 109023 (2023). <https://doi.org/10.1016/j.cie.2023.109023>
7. Martínez Barbero, X., Calvo, R., Ramos, J.: Portfolio optimization with prediction based return using long short term memory neural networks. *Computational Economics* **65**(4), 1479–1504 (2025). <https://doi.org/10.1007/s10614-024-10592-9>
8. Zhang, Z., Zohren, S., Roberts, S.: Deep learning for portfolio optimization. *Machine Learning* **110**(8), 2113–2145 (2021). <https://doi.org/10.1007/s10994-021-05946-y>