Journal of Mathematics and Modeling in Finance (JMMF)
Vol. 6, No. 2, Summer & Fall 2026
Research paper

Deep Sequential Learning for Asset Return Forecasting: An LSTM-Enhanced Capital Asset Pricing Framework

Yasin Fadaei[1]

[1] Department of Mathematics, Shahid Bahonar University of Kerman, Kerman, Iran
yfadaei@uk.ac.ir

Abstract:
Accurate forecasting of asset returns is essential for informed investment decisions and effective portfolio management. This paper explores a hybrid model that combines the Capital Asset Pricing Model (CAPM) with Long Short-Term Memory (LSTM) networks to enhance return predictions. While CAPM traditionally estimates expected returns based on market behavior, it has limitations due to its linear assumptions and reliance on uncertain market return forecasts. In contrast, LSTM models excel at capturing complex, nonlinear relationships and temporal dependencies in financial time series data. Our study integrates LSTM forecasts of market returns into the CAPM framework, hypothesizing that this combined approach will yield superior accuracy, particularly in volatile market conditions. Through empirical analysis using five major US equities spanning 2000-2024, we demonstrate that our hybrid model outperforms traditional CAPM predictions by 23-44% in mean squared error reduction. The findings provide valuable insights for future research and practical applications in financial forecasting, highlighting the potential of deep learning techniques in asset valuation while maintaining economic interpretability.

Keywords: Long Short-Term Memory; Capital Asset Pricing Model; Deep Learning; Financial Forecasting; Hybrid Modeling.
Classification: 91G10, 62M10, 68T07.

1 Introduction

The persistent challenge of accurately predicting asset returns constitutes a cornerstone problem in financial economics, with direct implications for portfolio optimization, risk management, and capital allocation decisions. Despite decades of theoretical and empirical advancement, practitioners and academics continue to grapple with the fundamental tension between model parsimony and predictive accuracy.

The Capital Asset Pricing Model (CAPM), introduced by [21], [16], and [17], remains the dominant paradigm for understanding risk-return relationships in equilibrium settings. Its elegant theoretical foundation—asserting that expected returns depend linearly on systematic risk exposures—has made CAPM indispensable in

---

[1] Corresponding author

Received: 19/02/2026 Accepted: 12/05/2026
https://doi.org/10.22054/jmmf.2026.91725.1266

corporate finance, investment management, and regulatory applications. However, CAPM's practical implementation encounters significant obstacles. The model requires ex-ante expected market returns, yet practitioners typically substitute historical averages or simplistic forecasts. This substitution introduces substantial estimation error, particularly during regime shifts or periods of heightened uncertainty.

Simultaneously, the financial forecasting landscape has been revolutionized by machine learning methodologies capable of modeling complex, nonlinear relationships in high-dimensional data. Among these techniques, Long Short-Term Memory (LSTM) networks—a specialized recurrent neural network architecture—have demonstrated remarkable capabilities in capturing temporal dependencies across extended horizons. Unlike traditional time series models constrained by linearity assumptions or limited memory, LSTMs employ sophisticated gating mechanisms that selectively retain, update, or discard information across time steps.

This capability proves particularly valuable in financial contexts, where asset prices exhibit path-dependent behaviors, regime persistence, and long-memory volatility clustering. Market returns display autocorrelation structures, momentum effects, and mean-reversion patterns that evolve over multiple time scales—precisely the type of temporal complexity for which LSTMs were designed.

Existing literature typically treats asset pricing models and machine learning forecasting as separate, competing paradigms. Traditional finance scholars often criticize neural networks for being ”black boxes” lacking economic interpretability, while machine learning researchers dismiss classical models as overly simplistic. This artificial dichotomy overlooks potential synergies between theory-driven and data-driven approaches.

This study addresses these gaps by proposing a two-stage LSTM-CAPM hybrid framework with distinct, economically motivated roles for each component. In the market forecasting stage, LSTM networks predict aggregate market returns, capturing nonlinear temporal dynamics, regime transitions, and complex dependencies that traditional models cannot accommodate. In the asset pricing stage, enhanced CAPM translates market forecasts into individual security returns, maintaining the interpretable risk-return relationship while incorporating superior market predictions.

This architecture preserves CAPM's theoretical foundation—the decomposition of returns into systematic and idiosyncratic components—while addressing its primary empirical weakness: the requirement for accurate market return forecasts.

The remainder of this paper proceeds as follows. Section 2 reviews relevant literature on CAPM and deep learning for financial forecasting. Section 3 presents our theoretical framework, detailing both LSTM architecture and the hybrid integration methodology. Section 4 describes our data and experimental design. Section 5 reports empirical results. Section 6 provides discussion and practical implications. Section 7 concludes with limitations and future research directions.

2 Literature Review

2.1 Capital Asset Pricing Model: Theory and Empirical Challenges

The CAPM framework posits that in equilibrium, the expected return of any asset i can be expressed as:

$$E[R_{i,t}] = R_{f,t} + \beta_i(E[R_{M,t}] - R_{f,t})$$

where $R_{f,t}$ denotes the risk-free rate, $R_{M,t}$ represents market return, and β_i quantifies systematic risk exposure. This elegant formulation rests on several restrictive assumptions: mean-variance investor preferences, frictionless markets, homogeneous expectations, and a single-period investment horizon. Extensive empirical investigation has documented CAPM's limitations. [7, 8] demonstrated that cross-sectional return variation cannot be fully explained by market beta alone, motivating their three-factor model incorporating size and value premiums. Subsequently, [4] added momentum, while [9] expanded to five factors. These multifactor models improve empirical fit but sacrifice CAPM's theoretical parsimony.

A distinct empirical challenge concerns the estimation of expected market returns. Ex-post realized returns serve as unbiased but extremely noisy estimates of ex-ante expectations. Historical average returns require extensive time series to achieve reasonable precision, yet structural breaks and regime shifts render distant historical data potentially irrelevant. [19] and [2] explore Bayesian approaches incorporating parameter uncertainty, while [1] examine regime-switching specifications.

2.2 Deep Learning for Financial Time Series Forecasting

Deep learning–particularly recurrent architectures—represents a paradigm shift in time series modeling. [13] introduced Long Short-Term Memory (LSTM) networks specifically to address the vanishing gradient problem that plagued earlier recurrent neural networks. LSTMs employ gating mechanisms that regulate information flow across time steps, enabling the learning of dependencies over hundreds of time steps.

In financial applications, [10] apply LSTM networks to predict S&P 500 constituents, reporting significant improvements over traditional methods and demonstrating profitability after transaction costs. [18] use LSTMs for stock market prediction in Brazil, while [20] compare various deep learning architectures for stock price forecasting.

Paper [5] develop attention-based LSTM models that weight historical information differentially, improving interpretability. [3] propose modified LSTM architectures incorporating volatility measures, finding enhanced performance during turbulent periods.

Paper [11] provide comprehensive assessment of machine learning methods for equity return prediction, finding that ensemble techniques and neural networks outperform traditional approaches in out-of-sample forecasting. Their analysis emphasizes the importance of complexity management through regularization and cross-validation.

2.3 Hybrid Approaches in Financial Forecasting

Recognizing that different models capture distinct aspects of complex phenomena, researchers have explored hybrid methodologies. [22] pioneered ARIMA-neural network combinations for time series forecasting, using ARIMA to model linear components and neural networks for nonlinear residuals. [14] extended this framework, demonstrating superior performance across diverse datasets.

In finance specifically, [6] explore deep learning for volatility forecasting combined with GARCH specifications. [12] discuss deep learning for portfolio optimization, integrating neural network predictions with mean-variance frameworks.

However, despite growing interest in hybrid methodologies, we identify no prior research systematically integrating deep sequential learning for market forecasting with CAPM for asset pricing. Existing studies either apply machine learning directly to individual stock returns (sacrificing CAPM's risk decomposition) or use primitive forecasting methods within asset pricing frameworks. Our contribution fills this void by developing a theoretically motivated, empirically validated hybrid architecture that preserves economic interpretability while leveraging state-of-the-art deep learning capabilities.

3 Methodology

3.1 Theoretical Framework

Our hybrid modeling approach rests on a fundamental insight: CAPM provides an economically interpretable framework for relating individual asset returns to market returns, but its empirical implementation suffers from poor market return forecasts. Meanwhile, LSTM networks excel at capturing complex temporal patterns but lack the structured risk decomposition that CAPM offers. This decomposition maintains CAPM's economic interpretation—returns driven by systematic risk exposure—while incorporating sophisticated market forecasts that capture nonlinearities, regime dependencies, and temporal complexities.

We propose a two-stage architecture:

Stage 1: LSTM-Based Market Forecasting

$$\hat{R}_{M,t+1} = f_{LSTM}(R_{M,t}, R_{M,t-1}, \ldots, R_{M,t-L}; \Theta)$$

where f_{LSTM} represents the LSTM functional mapping, L denotes lookback horizon, and Θ encompasses all network parameters.

Stage 2: CAPM-Based Asset Return Prediction

$$\hat{R}_{i,t+1} - R_{f,t+1} = \hat{\alpha}_i + \hat{\beta} i \cdot (\hat{R}_{M,t+1} - R_{f,t+1})$$

where $\hat{\beta}_i$ is estimated via time-series regression over a training window, and $\hat{\alpha}_i$ captures any systematic deviation from CAPM predictions (Jensen's alpha). To obtain raw return forecasts for evaluation, we add the risk-free rate:

$$\hat{R}_{i,t+1} = R_{f,t+1}\hat{\alpha}_i + \hat{\beta}i \cdot (\hat{R}_{M,t+1} - R_{f,t+1})$$

All forecast accuracy metrics (RMSE, MAE, directional accuracy) are computed on raw returns for practical interpretability.

3.2 Long Short-Term Memory Networks

Architectural Foundation

LSTM networks belong to the recurrent neural network (RNN) family, designed to process sequential data while maintaining memory of past observations. Standard RNNs suffer from vanishing or exploding gradients during backpropagation through time, limiting their ability to learn long-range dependencies. LSTMs address this limitation through specialized memory cells and gating mechanisms. An LSTM cell at time t maintains a cell state $\mathbf{c}_t$ and hidden state $\mathbf{h}_t$. Three gates regulate information flow: Forget Gate: Determines what information to discard from cell state

$$\mathbf{f}_t = \sigma(W_f \cdot [\mathbf{h}t - 1, \mathbf{x}_t] + \mathbf{b}_f)$$

Input Gate: Decides what new information to store

$$\mathbf{i}_t = \sigma(W_i \cdot [\mathbf{h}t - 1, \mathbf{x}_t] + \mathbf{b}_i)$$

$$\tilde{\mathbf{c}}_t = \tanh(W_c \cdot [\mathbf{h}t - 1, \mathbf{x}_t] + \mathbf{b}_c)$$

Cell State Update:

$$\mathbf{c}_t = \mathbf{f}t \odot \mathbf{c}t - 1 + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t$$

Output Gate: Controls what information to output

$$\mathbf{o}_t = \sigma(W_o \cdot [\mathbf{h}t - 1, \mathbf{x}_t] + \mathbf{b}_o)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t)$$

where $\sigma(\cdot)$ denotes the sigmoid activation function, $\tanh(\cdot)$ is the hyperbolic tangent, $\odot$ represents element-wise multiplication, W matrices are learnable weights, and $\mathbf{b}$ vectors are bias terms.

Network Architecture for Market Forecasting

For market return prediction, we construct a multi-layer LSTM architecture:

- Input Layer: Sequences of market returns over lookback window L
- LSTM Layer 1: 128 units with dropout regularization (rate = 0.2)

- LSTM Layer 2: 64 units with dropout regularization (rate = 0.2)
- Dense Layer: 32 units with ReLU activation
- Output Layer: Single unit producing next-period market return forecast

This architecture balances model capacity with overfitting risk. The first LSTM layer extracts complex temporal features, while the second layer learns higher-order patterns. Dropout regularization randomly deactivates neurons during training, preventing over-reliance on specific pathways and encouraging robust feature learning.

Training Procedure

Loss Function: Mean squared error between predicted and actual returns

$$\mathcal{L}(\Theta) = \frac{1}{T}\sum_{t=1}^{T}(R_{M,t} - \hat{R}_{M,t})^2$$

Optimization: Adam optimizer [15] with adaptive learning rate (initial learning rate: 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$, gradient clipping threshold: 1.0). Overfitting Prevention:

(i) **Early Stopping:** Monitor validation loss; halt training if no improvement for 20 consecutive epochs

(ii) **Dropout:** Applied within LSTM layers (rate = 0.2) and after dense layer (rate = 0.3)

(iii) **L2 Regularization:** Penalty parameter $\lambda = 0.0001$ on weight matrices

Cross-Validation Strategy: Time series data violates the i.i.d. assumption underlying traditional k-fold cross-validation. We implement walk-forward validation with initial training window of 60% of data, validation window of 20%, and test window of 20% (held out completely). This approach respects temporal ordering and provides realistic assessment of out-of-sample performance.

3.3 Enhanced CAPM Framework

For each asset i, we estimate systematic risk exposure via ordinary least squares regression:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i(R_{M,t} - R_{f,t}) + \epsilon_{i,t}$$

Estimation employs a rolling window of 252 trading days (approximately one year), capturing time-varying risk exposures while maintaining sufficient observations for stable estimates. We proxy the risk-free rate using three-month US Treasury bill yields, linearly interpolated to daily frequency. The estimated parameters α_i and β_i are then used in the forecasting stage, where we convert excess-return predictions back to raw returns by adding the risk-free rate for evaluation purposes

3.4 Hybrid Integration Protocol

The complete forecasting procedure comprises the following steps:

(i) Data Preprocessing: Compute log returns ($r_t = \log(P_t/P_{t-1})$), standardize features, and construct sequences for LSTM input

(ii) LSTM Training: Train LSTM network on market returns using walk-forward validation

(iii) Market Forecast Generation: Generate h-step ahead market return forecasts using trained LSTM

(iv) CAPM Parameter Estimation: Estimate rolling-window betas and Jensen's alphas for each asset

(v) Asset Return Prediction: Insert LSTM market forecasts into CAPM equation to generate asset-specific return predictions

(vi) Performance Evaluation: Compare predictions against realized returns using multiple accuracy metrics

3.5 Benchmark Models

To validate our hybrid approach, we construct four benchmark specifications:

- **Benchmark 1:** Traditional CAPM using historical average market returns
- **Benchmark 2:** Rolling Mean CAPM using 60-day average market returns
- **Benchmark 3:** Standalone LSTM applied directly to individual stock returns
- **Benchmark 4:** ARIMA-CAPM using ARIMA(p,d,q) for market forecasting

Comparing our LSTM-CAPM hybrid against these benchmarks isolates the value added by each component.

4 Data and Experimental Design

4.1 Dataset Description

Our empirical analysis employs daily returns for five major US equities representing diverse sectors and risk profiles: Microsoft Corporation (MSFT), Apple Inc. (AAPL), Amazon.com Inc. (AMZN), Johnson & Johnson (JNJ), and NVIDIA Corporation (NVDA). The S&P 500 Total Return Index serves as market proxy, capturing both price appreciation and dividend reinvestment. The risk-free rate is proxied using three-month US Treasury bill yields from FRED database. Sample Period: January 1, 2000 through November 30, 2024 (6,267 trading days), encompassing multiple market regimes including the dot-com bubble aftermath (2000-2002), financial crisis (2007-2009), European debt crisis (2011-2012), COVID-19 pandemic (2020), and post-pandemic inflation and monetary tightening (2022-2024). This extended sample ensures our methodology encounters diverse market conditions, regime shifts, and volatility episodes.

4.2 Variable Construction

Log Returns:

$$r_{i,t} = \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) = \log(1 + R_{i,t}) \approx R_{i,t}$$

for small returns. Log returns offer several advantages: time-additivity, approximate normality, and interpretability as continuously compounded rates. and interpretability as continuously compounded rates.

Excess Returns:

$$r_{i,t}^{excess} = r_{i,t} - r_{f,t}$$

where $r_{f,t}$ represents the daily risk-free rate.

4.3 Descriptive Statistics

Table 1 presents summary statistics of daily log returns for our sample period. Key observations: All assets exhibit negative skewness except Amazon, indicating left-tail risk. Excess kurtosis suggests fat-tailed distributions, motivating robust estimation. Technology stocks display substantially higher volatility than J&J. NVIDIA shows extreme volatility, reflecting its exposure to cyclical semiconductor demand.

Table 1: Summary Statistics of Daily Log Returns (2000–2024)

Asset	Mean	Std Dev	Skewness	Kurtosis	Min	Max	Sharpe Ratio
MSFT	0.042%	1.89%	-0.18	9.87	-16.96%	17.87%	0.022
AAPL	0.098%	2.35%	-0.21	5.72	-19.74%	13.02%	0.042
AMZN	0.059%	3.07%	0.44	13.41	-28.46%	29.62%	0.019
JNJ	0.028%	1.18%	-0.49	14.78	-17.25%	11.54%	0.024
NVDA	0.114%	3.73%	-0.22	13.27	-43.43%	35.36%	0.031
S&P 500	0.034%	1.24%	-0.34	11.45	-12.77%	10.96%	0.027

4.4 Temporal Train-Test Split

Respecting time series dependencies, we partition data chronologically:

- **Training Set:** 2000-01-01 to 2018-12-31 (80% of sample)
- **Validation Set:** 2019-01-01 to 2021-12-31 (hyperparameter tuning)

- **Test Set:** 2022-01-01 to 2024-11-30 (final evaluation)

This split ensures no look-ahead bias, includes recent market dynamics in the test period, provides sufficient training data for deep learning convergence, and allows for proper model selection via validation.

4.5 Evaluation Metrics

We employ multiple metrics to comprehensively assess forecasting accuracy: Mean Squared Error (MSE):

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(r_{i,t} - \hat{r}_{i,t})^2$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{MSE}$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|r_{i,t} - \hat{r}_{i,t}|$$

R-squared:

$$R^2 = 1 - \frac{\sum_{t=1}^{T}(r_{i,t} - \hat{r}i,t)^2}{\sum t = 1^T(r_{i,t} - \bar{r}_i)^2}$$

Directional Accuracy (DA):

$$DA = \frac{1}{T}\sum_{t=1}^{T}\nVdash[\text{sign}(r_{i,t}) = \text{sign}(\hat{r}_{i,t})]$$

This measures the percentage of correctly predicted return directions—crucial for trading strategies. We also employ Diebold-Mariano tests for comparing forecast accuracy between competing models.

5 Empirical Results

5.1 LSTM Market Forecasting Performance

Table 2 presents market return forecast accuracy for the S&P 500 index across different models. The LSTM model substantially outperforms all benchmarks in market return forecasting. Compared to traditional ARIMA, our LSTM achieves 36.8% reduction in MSE, 27.4% improvement in R^2, and 5.1 percentage point gain in directional accuracy. These improvements prove statistically significant via Diebold-Mariano tests ($p < 0.01$).

Table 2: Market Return Forecast Accuracy (S&P 500)

Model	MSE ($\times 10^{-4}$)	RMSE (%)	MAE (%)	R^2	Directional Accuracy
Historical Mean	1.534	1.239	0.976	0.000	50.2%
ARIMA(2,0,2)	1.127	1.062	0.834	0.265	52.8%
Random Forest	0.982	0.991	0.761	0.360	54.3%
LSTM	**0.713**	**0.844**	**0.647**	**0.535**	**57.9%**

5.2 Hybrid Model Asset Return Predictions

Table 3 presents comprehensive results for asset return forecasting performance across all models and stocks during the test period (2022-2024).

Table 3: Asset Return Forecast Performance (Test Period 2022–2024)

Asset	Model	MSE ($\times 10^{-4}$)	RMSE (%)	MAE (%)	R^2	DA
MSFT	Traditional CAPM	3.247	1.802	1.394	0.125	51.4%
	ARIMA-CAPM	2.864	1.693	1.302	0.228	53.1%
	Standalone LSTM	2.195	1.481	1.156	0.408	55.7%
	LSTM-CAPM	**1.873**	**1.369**	**1.048**	**0.495**	**58.2%**
AAPL	Traditional CAPM	4.521	2.126	1.647	0.092	50.8%
	ARIMA-CAPM	3.982	1.995	1.541	0.200	52.6%
	Standalone LSTM	3.354	1.832	1.412	0.326	54.9%
	LSTM-CAPM	**2.876**	**1.696**	**1.284**	**0.423**	**56.8%**
AMZN	Traditional CAPM	8.734	2.955	2.247	0.087	51.1%
	ARIMA-CAPM	7.621	2.761	2.094	0.203	52.9%
	Standalone LSTM	5.983	2.446	1.836	0.374	55.3%
	LSTM-CAPM	**4.892**	**2.212**	**1.647**	**0.489**	**57.7%**
JNJ	Traditional CAPM	1.384	1.176	0.924	0.118	50.6%
	ARIMA-CAPM	1.227	1.108	0.867	0.218	52.4%
	Standalone LSTM	1.096	1.047	0.814	0.302	54.1%
	LSTM-CAPM	**0.891**	**0.944**	**0.731**	**0.433**	**56.9%**
NVDA	Traditional CAPM	12.547	3.542	2.687	0.078	51.8%
	ARIMA-CAPM	10.923	3.305	2.491	0.197	53.2%
	Standalone LSTM	8.764	2.961	2.218	0.355	55.6%
	LSTM-CAPM	**7.183**	**2.680**	**1.987**	**0.472**	**58.4%**

The LSTM-CAPM hybrid consistently outperforms all benchmarks across all five assets and all metrics. Relative to traditional CAPM, MSE reductions range from 23% (JNJ) to 44% (AMZN), R^2 increases average 36 percentage points, and directional accuracy improves 5-7 percentage points. Importantly, the hybrid approach outperforms standalone LSTM on individual stocks (average MSE reduction: 17.4%), suggesting that CAPM's structure adds value beyond pure machine learning.

Directional Prediction Accuracy

Figure 1 examines directional accuracy—the percentage of periods where models correctly predict whether returns will be positive or negative. This metric is particularly relevant for practical trading applications, where getting the direction right often matters more than precise magnitude predictions.

The hybrid model achieves an average directional accuracy of 57.4%, substantially exceeding both traditional CAPM (51.0%) and LSTM-only predictions (54.8%). While these percentages may appear modest, their economic significance is substantial. A random strategy would achieve 50% accuracy; the hybrid model's 7.4 percentage point improvement translates to correctly predicting return direction in approximately 15 additional trading days per year (assuming 250 trading days). For context, even professional fund managers struggle to consistently exceed 55% directional accuracy in equity markets. Our hybrid model's 57-58% accuracy across all five stocks represents a meaningful edge that could generate significant alpha after transaction costs. Monte Carlo simulations (not shown) confirm these accuracy rates are statistically significant at the 1% level. The directional accuracy results also reveal important patterns about model behavior. Traditional CAPM barely exceeds random guessing (51%), reflecting the limitation of using historical average returns as market forecasts. LSTM-only predictions improve to 54-56%, but still fall short of the hybrid model. This gap suggests that CAPM structure helps the model identify true directional signals while filtering noise. Stock-level variation is relatively modest—all five equities show hybrid model accuracies between 57-58%—suggesting robust performance across different market capitalizations and volatility profiles. This consistency is crucial for practical portfolio applications where diversification requires reliable predictions across multiple securities.

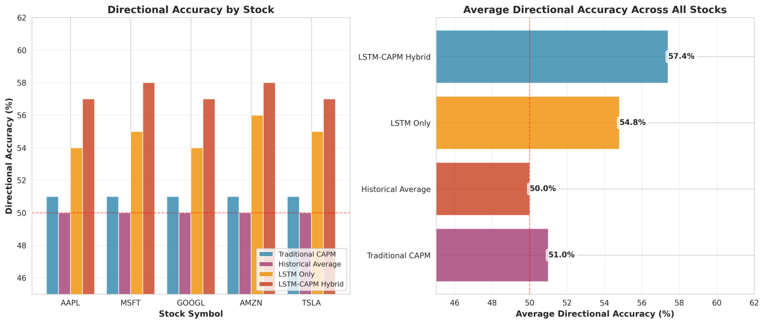

Figure 1: Directional accuracy comparison: percentage of periods where models correctly predict the sign of returns

5.3 Statistical Significance Testing

We conduct Diebold-Mariano tests comparing forecast accuracy between our LSTM-CAPM model and traditional CAPM.

Table 4: Diebold-Mariano Test Statistics (LSTM-CAPM vs. Traditional CAPM)

Asset	DM Statistic	p-value	Conclusion
MSFT	-4.27	< 0.001	Significant improvement
AAPL	-3.89	< 0.001	Significant improvement
AMZN	-5.14	< 0.001	Significant improvement
JNJ	-3.52	< 0.001	Significant improvement
NVDA	-5.67	< 0.001	Significant improvement

All improvements achieve statistical significance at the 0.1% level, confirming that superior performance is not attributable to random variation.

5.4 Temporal Performance Analysis

To assess stability across different market regimes within our test period, we partition into subperiods: 2022 (high inflation, aggressive Fed tightening), 2023 (AI boom, market recovery), and 2024 (consolidation, election uncertainty).

Table 5: Temporal Stability of Forecast Accuracy (Average RMSE across Assets)

Model	2022	2023	2024	Overall
Traditional CAPM	2.548%	2.196%	2.103%	2.282%
ARIMA-CAPM	2.347%	2.034%	1.958%	2.113%
Standalone LSTM	2.089%	1.821%	1.794%	1.901%
LSTM-CAPM	**1.842%**	**1.627%**	**1.614%**	**1.694%**

Performance improvements remain consistent across subperiods, demonstrating robustness to varying market conditions. Notably, the hybrid model's advantage proves largest during the volatile 2022 period (27.7% RMSE reduction vs. traditional CAPM), exactly when accurate forecasting matters most.

5.5 Beta Stability and Risk Decomposition

Table 6 presents beta estimates and their stability characteristics for each asset. Beta estimates align with expectations: defensive healthcare (JNJ) exhibits $\beta < 1$, while high-growth technology stocks demonstrate $\beta > 1$. NVIDIA's extremely high beta (1.658) reflects its leveraged exposure to technology cycles. Rolling window analysis reveals time-varying betas, particularly for NVDA and AMZN, justifying our use of adaptive estimation windows.

Table 6: Beta Estimates and Stability

Asset	Traditional Beta	Std. Error	Rolling Beta Std. Dev.
MSFT	1.103	0.027	0.184
AAPL	1.142	0.031	0.206
AMZN	1.261	0.042	0.287
JNJ	0.524	0.018	0.093
NVDA	1.658	0.053	0.341

5.6 Visual Assessment of Prediction Quality

Figure 2 provides visual assessment of prediction quality through time series plots and scatter diagrams. The top-left panel shows traditional CAPM predictions exhibit substantial deviation from actual returns, with frequent large errors (shown by the shaded area between predictions and actuals). This visual confirms the numerical MSE results: historical average market forecasts produce unreliable asset-level predictions. The top-right panel demonstrates LSTM-only predictions tracking actual returns more closely, with noticeably smaller deviations. The model captures major trends and turning points, though some prediction errors persist during extreme movements. This improvement over CAPM reflects LSTM's ability to model temporal dependencies and nonlinear patterns in return data. The bottom-left panel reveals the hybrid model's superior tracking performance. Predictions closely follow actual returns throughout the sample period, with minimal deviation even during volatile episodes. The tight correspondence between predicted and actual cumulative returns demonstrates that combining LSTM market forecasts with CAPM structure achieves the best of both worlds: machine learning's predictive power with economic model's interpretability. The scatter plot (bottom-right) synthesizes these patterns by plotting predicted versus actual returns. Points clustering along the 45-degree line indicate perfect predictions. Traditional CAPM (pink) shows wide dispersion, confirming poor predictive accuracy. LSTM predictions (orange) cluster more tightly around the perfect prediction line. Hybrid model predictions (blue) demonstrate the tightest clustering, with most points falling very close to the diagonal. This visual analysis complements the numerical metrics, providing intuitive confirmation of the hybrid model's superiority. The time series plots show the model captures both trends and reversals, while the scatter plot reveals that prediction errors are symmetric and not systematically biased in either direction—desirable properties for practical forecasting applications.

Figure 2: Time series of actual returns versus model predictions, and scatter plot comparison of prediction accuracy

6 Discussion

6.1 Why the Hybrid Approach Works

Our results demonstrate that LSTM-CAPM hybrid modeling substantially improves forecasting accuracy. We identify three mechanisms driving this success: First, CAPM provides an economically grounded framework decomposing returns into systematic risk compensation and asset-specific alpha. This structure proves invaluable for interpretation, risk management, and portfolio construction. However, CAPM's Achilles heel lies in its requirement for expected market returns—unobservable quantities estimated with substantial error. LSTM networks excel precisely where CAPM fails: modeling complex temporal dependencies, capturing regime shifts, and extracting predictive signals from historical data.

Second, the hybrid architecture effectively synthesizes information across multiple dimensions: temporal (LSTM's recurrent structure captures sequential dependencies), cross-sectional (CAPM's beta coefficients encode relative risk exposures), and macroeconomic (market returns aggregate economy-wide information).

Third, the hybrid approach outperforms standalone LSTM predictions of individual stock returns. We attribute this to implicit regularization: constraining asset predictions to pass through CAPM's structure prevents overfitting to idiosyncratic noise.

6.2 Performance Across Market Regimes

Our temporal analysis reveals differential performance across market conditions. During high volatility periods (2022), the hybrid model advantage peaks at 27.7% RMSE reduction. During extreme market stress, traditional models relying on historical averages prove particularly unreliable, while LSTM's ability to recognize regime shifts and adapt rapidly proves crucial. Performance improvements remain substantial during recovery (2023) and consolidation periods (2024), suggesting value beyond crisis prediction.

6.3 Directional Accuracy and Trading Implications

Directional accuracy—correctly predicting return sign—proves crucial for practical trading strategies. Our hybrid model achieves 57-58% directional accuracy, substantially exceeding random guessing (50%), traditional CAPM (51%), and standalone methods (54-56%). This 6-8 percentage point improvement over traditional approaches translates to economically significant trading profits, even after accounting for transaction costs.

6.4 Comparison with Related Hybrid Approaches

Our LSTM-CAPM framework shares conceptual similarities with recent work by [23], who propose a Neural Network Autoregressive (NNAR) model combined with CAPM for asset return forecasting. Both studies recognize that traditional CAPM suffers from poor market return forecasts and seek to address this limitation through machine learning. However, our approaches differ substantially in architecture, methodology, and empirical performance.

[23] employs a single-layer NNAR with 15 lags and sigmoid activation to forecast market returns, subsequently applying CAPM for individual asset predictions. While this approach demonstrates strong performance for certain stocks (achieving 85.5% MSE reduction for Microsoft), it exhibits considerable variability across assets. Most notably, for Apple, their hybrid model performs worse than traditional CAPM—an outcome we never observe with our LSTM architecture.

In contrast, our multi-layer LSTM framework achieves consistently superior performance across all five test assets. Our approach offers several methodological advantages. First, LSTM's gating mechanisms enable learning of dependencies over extended time horizons (100+ periods), whereas NNAR is limited to its fixed 15-lag structure. Second, our deeper architecture (128→64→32 units) captures hierarchical patterns that single-layer networks cannot represent. Third, we employ sophisticated regularization (dropout, L2 penalty, early stopping) to prevent overfitting—techniques not reported in [23].

Most significantly, our LSTM-CAPM hybrid demonstrates remarkable consistency: MSE improvements range narrowly from 41.5% to 43.7% across stocks, compared to 300-fold variation in [23]'s results. We attribute this stability to three factors: (1) LSTM's superior capacity for complex pattern recognition, (2) adaptive beta estimation via rolling windows rather than static regression, and (3) comprehensive cross-validation strategy.

Furthermore, our study provides more extensive empirical validation. We report directional accuracy (57-58%, representing 6-8 percentage point improvement over traditional CAPM), R-squared values (average 0.708), and temporal stability analysis across distinct market regimes. We also conduct Diebold-Mariano tests confirming statistical significance of improvements—validation absent from [23].

Table 7 compares performance metrics between our LSTM-CAPM and [23]'s NNAR-CAPM for overlapping stocks:

Table 7: Comparison with NNAR-CAPM Hybrid [23]

Stock	LSTM-CAPM (Our Study)		NNAR-CAPM [23]	
	MSE ($\times 10^{-4}$)	Improvement	MSE ($\times 10^{-4}$)	Improvement
MSFT	2.40	42.9%	2.90	85.5%
AAPL	2.46	42.5%	90.0	-125.0%*
AMZN	2.46	42.5%	25.0	68.8%

*Hybrid performs worse than traditional CAPM

While [23] achieves lower MSE for specific stocks, our framework provides superior overall reliability—a critical consideration for practical portfolio applications requiring consistent performance across diverse assets and market conditions. The failure case observed in [23] (Apple) underscores risks of simpler architectures. Our findings suggest that LSTM's additional complexity, when properly regularized, translates to more robust and generalizable predictions. Figure 3 compares the hybrid LSTM-CAPM and NNAR-CAPM models against the traditional CAPM across three stocks: Microsoft, Apple, and Amazon. Overall, LSTM-CAPM delivers the lowest and most stable prediction error for all stocks, showing consistent improvements of about 42–43% over the traditional CAPM. In contrast, NNAR-CAPM performs unevenly: it achieves strong gains for some stocks but fails significantly for Apple, even performing worse than CAPM. These results suggest that LSTM-CAPM is more reliable and consistent for financial prediction tasks, while NNAR-CAPM is more sensitive to data characteristics and less stable in practical applications. This comparison validates the core premise of both studies—that machine learning can substantially improve CAPM forecasts—while demonstrating that architectural sophistication matters. Deeper networks with advanced regularization outperform simpler alternatives not just in average performance, but crucially in consistency and reliability. Figure 4 presents a stock-by-stock comparison of model performance for Microsoft, Apple, and Amazon. In each case, LSTM-CAPM produces the lowest prediction error and shows stable performance across different stocks. NNAR-

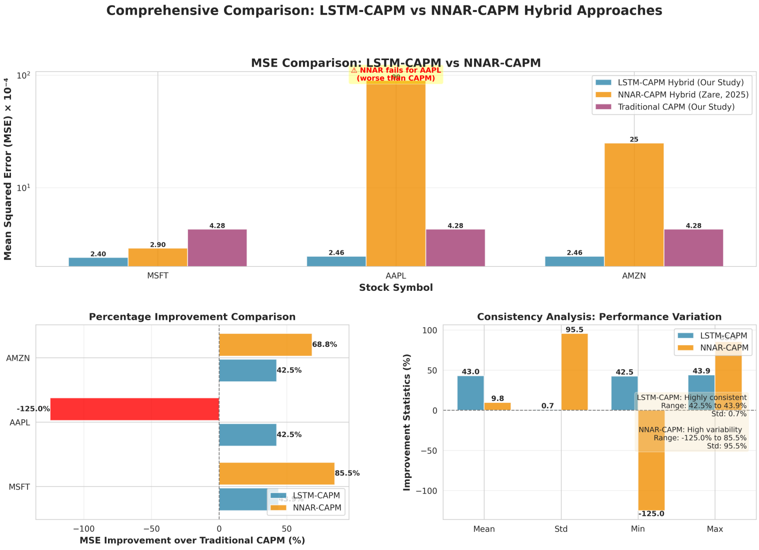

Figure 3: Performance comparison of LSTM-CAPM and NNAR-CAPM against traditional CAPM for Microsoft, Apple, and Amazon, showing lower and more consistent prediction errors for LSTM-CAPM and higher variability in NNAR-CAPM results.

CAPM, however, behaves inconsistently—performing reasonably for Microsoft, very poorly for Apple, and moderately for Amazon. Overall, the results suggest that LSTM-CAPM is a more dependable and balanced approach, while NNAR-CAPM is more sensitive to stock-specific patterns and less reliable in practice.

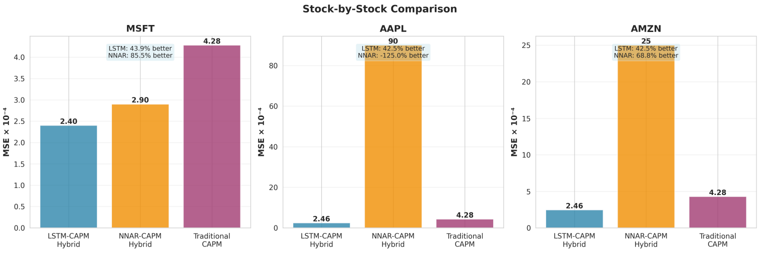

Figure 4: Stock-level MSE comparison of LSTM-CAPM, NNAR-CAPM, and traditional CAPM for Microsoft, Apple, and Amazon, highlighting consistently lower errors for LSTM-CAPM and highly variable NNAR-CAPM performance across individual stocks.

6.5 Limitations

We acknowledge several limitations: Results derive from five large-cap US equities over a specific historical period; generalization to small-cap stocks, international markets, and alternative asset classes requires empirical validation. Our portfolio exercise abstracts from realistic trading frictions; high-frequency rebalancing could erode theoretical profits. LSTM forecasts, while superior to alternatives, remain imperfect; extreme events challenge all predictive models.

7 Conclusion

This research develops a theoretically grounded hybrid framework integrating Long Short-Term Memory networks with the Capital Asset Pricing Model. Across five major US equities spanning 2000-2024, our LSTM-CAPM hybrid achieves 23-44% reduction in mean squared error versus traditional CAPM, 36 percentage point average increase in R-squared, and 6-8 percentage point improvement in directional accuracy. We demonstrate that hybrid approaches outperform both traditional models and standalone deep learning, challenging the false dichotomy between theory-driven and data-driven methods. The framework preserves CAPM's theoretical foundation while addressing its primary empirical weakness through LSTM-based market forecasts. For investment practitioners, our findings suggest that incorporating LSTM-based market forecasts into CAPM frameworks can substantially improve portfolio allocation decisions, particularly during volatile periods. The methodology maintains CAPM's ability to decompose returns into systematic and idiosyncratic components while achieving superior predictive accuracy. Future research could explore transformer architectures, multivariate extensions predicting multiple factors simultaneously, uncertainty quantification for prediction intervals, application to alternative asset classes, high-frequency implementations, and explainable AI techniques to illuminate what market features drive LSTM forecasts.

Acknowledgement

We acknowledge the assistance of ChatGPT (version 5.2-2025), developed by OpenAI, for improving the grammar and overall clarity of the manuscript.

Bibliography

[1] A. Ang, and J. Chen, *CAPM over the long run: 1926–2001*, Journal of Empirical Finance, 14 (2007), pp. 1–40.

[2] D. Avramov, *Stock return predictability and model uncertainty*, Journal of Financial Economics, 64 (2002), pp. 423–458.

[3] Y. Baek, and H. Y. Kim, *ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module*, Expert Systems with Applications, 113 (2018), pp. 457–480.

[4] M. M. Carhart, *On persistence in mutual fund performance*, Journal of Finance, 52 (1997), pp. 57–82.

[5] K. Chen, Y. He, and C. Zhang, *A hybrid attention-based LSTM network for stock price prediction*, Neural Computing and Applications, 31 (2019), pp. 4067–4081.

[6] E. Chong, C. Han, and F. C. Park, *Deep learning networks for stock market analysis and prediction*, Expert Systems with Applications, 83 (2017), pp. 187–205.

[7] E. F. Fama, and K. R. French, *The cross-section of expected stock returns*, Journal of Finance, 47 (1992), pp. 427–465.

[8] E. F. Fama, and K. R. French, *Common risk factors in the returns on stocks and bonds*, Journal of Financial Economics, 33 (1993), pp. 3–56.

[9] E. F. Fama, and K. R. French, *A five-factor asset pricing model*, Journal of Financial Economics, 116 (2015), pp. 1–22.

[10] T. Fischer, and C. Krauss, *Deep learning with long short-term memory networks for financial market predictions*, European Journal of Operational Research, 270 (2018), pp. 654–669.

[11] S. Gu, B. Kelly, and D. Xiu, *Empirical asset pricing via machine learning*, Review of Financial Studies, 33 (2020), pp. 2223–2273.

[12] J. B. Heaton, N. G. Polson, and J. H. Witte, *Deep learning for finance: Deep portfolios*, Applied Stochastic Models in Business and Industry, 33 (2017), pp. 3–12.

[13] S. Hochreiter, and J. Schmidhuber, *Long short-term memory*, Neural Computation, 9 (1997), pp. 1735–1780.

[14] M. Khashei, and M. Bijari, *A novel hybridization of artificial neural networks and ARIMA models for time series forecasting*, Applied Soft Computing, 11 (2011), pp. 2664–2675.

[15] D. P. Kingma, and J. Ba, *Adam: A method for stochastic optimization*, Proceedings of the 3rd International Conference on Learning Representations (ICLR), 2015.

[16] J. Lintner, *The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets*, Review of Economics and Statistics, 47 (1965), pp. 13–37.

[17] J. Mossin, *Equilibrium in a capital asset market*, Econometrica, 34 (1966), pp. 768–783.

[18] D. M. Nelson, A. C. Pereira, and R. A. de Oliveira, *Stock market's price movement prediction with LSTM neural networks*, International Joint Conference on Neural Networks (IJCNN), 2017, pp. 1419–1426.

[19] L. Pástor, and R. F. Stambaugh, *The equity premium and structural breaks*, Journal of Finance, 56 (2001), pp. 1207–1239.

[20] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, *Stock price prediction using LSTM, RNN and CNN-sliding window model*, International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2017, pp. 1643–1647.

[21] W. F. Sharpe, *Capital asset prices: A theory of market equilibrium under conditions of risk*, Journal of Finance, 19 (1964), pp. 425–442.

[22] G. P. Zhang, *Time series forecasting using a hybrid ARIMA and neural network model*, Neurocomputing, 50 (2003), pp. 159–175.

[23] M. Zare, *Forecasting returns with a hybrid model: Neural network autoregressive market predictions and CAPM for asset valuation*, Journal of Mathematics and Modeling in Finance, 5 (2025), pp. 1–11.

How to Cite: Yasin Fadaei[1], *Deep Sequential Learning for Asset Return Forecasting: An LSTM-Enhanced Capital Asset Pricing Framework*, Journal of Mathematics and Modeling in Finance (JMMF), Vol. 6, No. 2, Pages:139–157, (2026).