Comparing Predictive, Reinforcement, and Generative Deep Models for Statistical Arbitrage

Atharva Dhuri
Department of Computer Engineering
NMIMS, University
Mumbai, India
atharva.dhuri45@nmims.in

Rouhin Gandhi
Department of Computer Engineering
NMIMS, University
Mumbai, India
rouhingandhi123@gmail.com

Dr. Sofia Francis
Department of Computer Engineering
NMIMS, University
Mumbai, India
sofia.francis@nmims.edu

Vidit Agrawal
Department of Computer Engineering
NMIMS, University
Mumbai, India
vidit.agrawal67@nmims.in

Yash Narayan
Department of Computer Engineering
NMIMS, University
Mumbai, India
yash.narayan45@nmims.in

***Abstract*— Statistical arbitrage seeks to exploit short-term mispricing between correlated assets in pursuit of market-neutral profits. Standard econometric models fall short of estimating nonlinear dependencies and regime switches, characteristic features of complex financial data. The current paper discusses the different deep learning paradigms (predictive, reinforcement, and generative) in statistical arbitrage. The predictive framework utilizes LSTM, CNN-LSTM, and Attention-LSTM to forecast spreads; the reinforcement component uses a Deep Q-Network for optimizing sequential trades; and the generative module utilizes GAN and WGAN-GP to synthesize realistic log-return data for stress testing. Experiments conducted on six NSE equities have shown complementary strengths across these paradigms and underlined the trade-off between predictive accuracy and trading profit, with an emphasis on the joint value for modern quantitative finance.**

Keywords - Statistical Arbitrage, Deep Learning, LSTM, CNN-LSTM, DQN, WGAN-GP, Reinforcement Learning, Generative Models, Financial Time Series.

I. INTRODUCTION

Financial markets are complex adaptive systems that exhibit dynamic, nonlinear, and stochastic behavior, which is driven by a constant interplay of macroeconomic and microstructural factors.. As traditional econometric models are applied only in stationary and linear systems, it has motivated the integration of utilising deep learning techniques in qualitative finance, for strategies like statistical arbitrage, which helps identify mispricing between correlated assets to yield market neutral returns.

Linear models like Engle Granger cointegration test and Ornstein Uhlenbeck processes have been used in Statistical Arbitrage to identify and utilise mean reverting spreads. However, as financial data is complex and continuously changing, these conventional approaches are unable to meet expected standards in profitability and robustness. Thus, deep learning models, with their ability to find latent dependencies based patterns in time series data and its sequential approach, end up serving as a promising alternative.

The latest advancements demonstrate how various deep learning models offer unique strengths for quantitative trading. Long Short Term Memory (LSTM) based predictive models help study temporal dependencies in asset spreads to forecast arbitrage signals. We can optimize the sequential trading decisions by using reinforcement learning agents like DQNs; we can optimize the cumulative reward when applying such agents. Generative models, such as GAN, aid in synthesizing realistic data on market conditions for stress tests and analysis of potential patterns.

In parallel, contemporary market conditions have made modern environments increasingly difficult for traditional statistical arbitrage: high-frequency noise, structural breaks, and rapidly shifting volatility regimes. The presence of algorithmic trading, tighter spreads, and fragmentation of liquidity makes the price series deviate from ideal stochastic assumptions; therefore, static models cannot be relied upon in a practical setting. Moreover, financial time series often contain heavy tails, volatility clustering, jumps, and non-stationarity, which intrinsically violate the basic assumptions of classical linear models. These challenges illustrate the need for adaptive methods that are able to learn directly from data without resorting to strong parametric constraints. Deep learning models naturally meet these demands since they successfully model nonlinear interactions, extract hierarchical temporal features, and adapt to changing market conditions. The ability of deep learning models to incorporate multivariate signals, technical indicators, and contextual information further improves the robustness of the generated arbitrage signals. As a result, deep learning applied to statistical arbitrage shall offer a deeper understanding of market micro-dynamics, improve generalization across regime shifts, and lead to the development of data-driven trading systems capable of survival in complex current financial ecosystems.

The key contributions of this study are summarized as follows:

1. Unified Deep Learning Framework: A framework that helps in evaluating predictive, reinforcement, and generative models for statistical arbitrage.
2. Empirical Multi-Model Analysis: We implement the framework and evaluate the models across six stocks listed on NSE, with a focus on forecasting accuracy, trading profitability, and synthetic data realism.
3. Accuracy–Profitability Trade off Demonstration: Quantitative evidence highlighting the fact that, even though models afford high predictive accuracy, it does not necessarily correspond to superior financial returns. It shows the trade-off between forecast precision and trading performance.
4. Stability of Generative Modeling: The Wasserstein GAN produced the most stable and realistic-based

synthetic spreads that improve backtesting reliability and robustness under dynamic market conditions.

II. Literature Review

In the past few years, significant growth has been recorded in deep learning-based methods for financial forecasting, automated trading, and generation of data. The works published in recent years have considered predictive, reinforcement, and generative paradigms that are complementary in modeling market complexity, volatility, and nonlinear dependencies. The following section synthesizes the most relevant research between 2023 and 2025, underlining comparative insights and methodological developments for each paradigm.

A. *Predictive Models (LSTM, CNN-LSTM, and Attention-Based Architectures)*

Deep learning models using sequences remain the backbone of modern financial forecasting. Barua et al. [1] compared several recurrent and convolutional networks, including RNN, LSTM, CNN, GRU, and Attention-LSTM on Indian NIFTY 50 stocks, and found that CNN and GRU captured both short-term volatility and long-range dependencies better than their conventional recurrent counterparts. Kabir et al. [2] further improved predictive robustness with the hybrid LSTM–Transformer–MLP model that combined temporal memory with self-attention and yielded superior results across equities, cryptocurrencies, and indices.

Scaling this to multivariate learning, Dong and Liang [3] proposed a hybrid CNN-LSTM-GNN, which simultaneously extracts spatial-temporal features and cross-asset relationships using a graph neural network and that outperforms the baselines on the Chinese A-share market. Chen et al. [4] developed the Enhanced Multi-Aspect Attention Transformer (EMAT), which simultaneously captures trend, volatility, and temporal-decay dynamics through specialized attention heads, consistently surpassing conventional LSTM and Transformer models.

Additional works reinforce these trends. Gajamannage et al. [5] designed a sequentially trained dual-LSTM that retrains on its own forecasts for long-horizon stability, achieving higher real-time accuracy than ARIMA and Kalman filters. Zhu et al. [6] optimized LSTM hyper-parameters via Cuckoo Search and Zebra algorithms to forecast arbitrage spreads with a MAPE near 1.1 %, demonstrating that bio-inspired optimization can substantially reduce prediction error. Sebastian and Tantia [7] applied an attention-augmented multivariate LSTM to Indian stock indices, confirming statistically significant accuracy gains through the Diebold–Mariano tests. Finally, Zeng et al. [8] proposed a CNN–Transformer hybrid that merges local pattern extraction and global contextual modeling, outperforming deep and statistical benchmarks on S&P 500 intraday forecasting.

One such cross-cutting pattern in predictive modeling research is that financial time series have the best use of a predictive modeling architecture that encodes a combination of both model-driven pattern extraction and longer-range modeling of dependencies. Pure LSTMs are sensitive to slow changes in volatility, but implementations of CNN layers can be used to capture localized motion and noise reduction as well as capture longer-term temporal structure. Mechanisms of attention also contribute to the adaptability of the models since they enable them to give attention to the most informative periods in history and particularly during regime changes. Recent studies also emphasize the fact that multivariate forecasting, which includes cross-asset correlations, technical indicators and macroeconomic variables, has been able to overcome the shortcomings of univariate models. However, even with the increased degree of sophistication, predictive models continue to face a number of enduring issues, such as overfitting, vulnerability to market noise, and reliability breakdown in times of market reversals. These restrictions give a good argument on why the complementary paradigms- namely, reinforcement learning to optimize decisions, and generative models to stress test- should be used in creating a genuinely resilient statistical arbitrage framework.

Collectively, these papers confirm that the combination of recurrent, convolutional and attention mechanisms hybridization provides more expressiveness in the models to improve generalization and resilience to market regime changes.

B. *Reinforcement Learning for Trading Agents (DQN, PPO, and MARL Variants)*

Reinforcement learning (RL) has developed into an information-driven data paradigm of optimization of policies and not explicit price prediction. Another model is the Self-Rewarding Double DQN that was introduced by Huang et al. [9] to adaptively change its reward according to expert measures like the Sharpe ratio, and it attained more than 1100 percent cumulative return on the Nasdaq data. The implementation of an improved DDQN model on the NIFTY 50 index with a Sharpe ratio of 0.74 and 73 percent win rate on 15 minutes bars suggested the relevance of deep RL in Indian markets (Mishra et al. [10]).

Globally, Kong and So [11] compared seven RL algorithms (A2C, PPO, DDPG, SAC, TD3, TRPO, ACKTR) and found that bigger ensembles did not always achieve greater profitability, which highlights the essence of rewards construction. Zou et al. [12] made the use of cascaded LSTM feature extractors along with PPO agents to obtain higher returns and Sharpe ratios on the Dow 30 and SSE 50 indices. The quantitative trading process that Xu et al. [13] developed is a partially observable Markov decision process, and the authors introduced QTNet, which takes into account imitation learning to strike a balance between exploration and exploitation under high-frequency environments. In agreement with this, Wei et al. [14] suggested a multi-agent reinforcement learning environment based on MAPPO and Value-Decomposition Networks that enabled collaborative agents to yield better outcomes than single-agent DQN methods in terms of annualized 9.8% returns and transaction cost.

In reinforcement learning literature, practical challenges also indicate that these models are constrained by a number of practical issues and limitations to applying them directly to live trading. The commonest criticism is that deep RL algorithms are unstable in the presence of non-stationary market conditions-their policies degrade at an alarming rate on either side of times of elevated volatility or structural breaks. The other issue is the low and noisy reward signal accredited to financial markets which can greatly slow down convergence and push agents to suboptimal behaviour under no reward engineering. Numerous studies also highlight that RL agents are sensitive to transaction costs, slippage and execution delay-which are all important in the real world but are typically neglected during training. More recent work has thus moved in the direction of using risk-aware objectives, entropy regularization and hybrid structures combining supervised pre-training and RL fine-tuning. Such trends highlight the importance of strong, sample efficient agents capable of adapting to the changing regimes without a significant change in their trading policy.

Combined, these studies confirm that reinforcement learning can translate predictive signals into profitable actions, provided reward shaping and risk management are carefully engineered.

C. Generative Models for Synthetic Financial Data (GAN, WGAN-GP, and Diffusion Models)

Generative modeling gives access to efficient tools for data augmentation, stress-testing, and simulation. Ramzan et al.[15] conducted an analysis of GAN-based synthesis of stock market time series and showed high statistical similarity between real and synthetic data distributions. Zakharov et al.[16] achieved higher realism by regime clustering, with a custom QuantGAN trained on different market regimes to simulate regime-consistent series that resulted in higher quality forecasts of stock prices on different market regimes. Then, Ramzan et al.[17] demonstrated WGAN-based stress testing on S&P 500 and FTSE 100 market indices with the aim to validate Value-at-Risk calculation on synthetic data versus real-world data.

More contemporary innovations go beyond GANs. Takahashi and Mizuno in [18] used diffusion probabilistic models (DDPMs) with wavelet-based spectrogram data to replicate stylized facts, namely volatility clustering and heavy tailed distributions, producing very realistic multivariate financial data. Sattarov et al. in [19] developed FinDiff, a diffusion model to synthesize mixed-type financial data tables for stress tests, outperforming GAN baselines on both fidelity and privacy preservation goals. Rizzato et al. in [20] designed Jinkou, a conditional GAN model to synthesize scenario-specific data subject to macroeconomic constraints such as inflation and oil prices, useful for realistic multi-asset stress scenarios. Finally, Sedai et al. in [21] used GANs to generate more macroeconomic crisis data to increase the accuracy of bank capital stress tests.

The overall tendency of recent generative modeling research is the increase to less biased models that are more able to capture the heavy tailed, volatile character of financial time series and regime dependence. Mode collapse can be a problem with traditional GANs, which can be not suitable in the context of reproduction of extreme events or uncommon market states. The constraint has led to the popularity of WGAN-GP and diffusion-based methods that offer more amenable optimization scales and better distributional fidelity. In addition, conditional modeling is becoming a major focus of generative studies, with macroeconomic variables, volatility measures, or regime names as the conditioning variables to generate scenario-specific synthetic data. With this kind of conditionality, more specific stress testing can be done and sound model validation across varied environments can be achieved. Another emerging direction involves using synthetic data to pre-train or augment predictive and reinforcement learning models, helping them generalize better in situations with limited historical examples. Overall, these advances reinforce the role of generative modeling as a critical component in building resilient, data-rich quantitative trading frameworks.Within these works, generative models show the capability to synthesize data with a high level of fidelity for risk assessment tasks and to complement predictive/reinforcement models by exploring the space beyond what can be trained on directly.

D. Summary and Research Gap

The literature reviewed shows that there is a trend towards integrating hybrid and cross-paradigm models. There is an increasing trend in predictive models to integrate convolutional, recurrent, and attention mechanisms to improve the learning of features with respect to the temporal-spatial domains [1]–[8]. The literature on reinforcement learning models [9]–[14] focuses on reward shaping and cooperative game play. Generative research [15]–[21] extends beyond classical GANs to WGAN-GP and diffusion models that deliver statistically faithful synthetic data. Nevertheless, very few studies directly compare these paradigms within a single experimental pipeline, leaving unresolved questions regarding their trade-offs in accuracy, profitability, and robustness. The present work addresses this gap by providing a unified empirical comparison of predictive, reinforcement, and generative deep-learning architectures applied to statistical arbitrage.

III. METHODOLOGY

This paper implements a framework to evaluate the various domains under deep learning used in a statistical arbitrage pipeline. Figure 1 displays the overall workflow, from data acquisition and preprocessing, to training, evaluation, and comparative analysis across multiple performance dimensions.

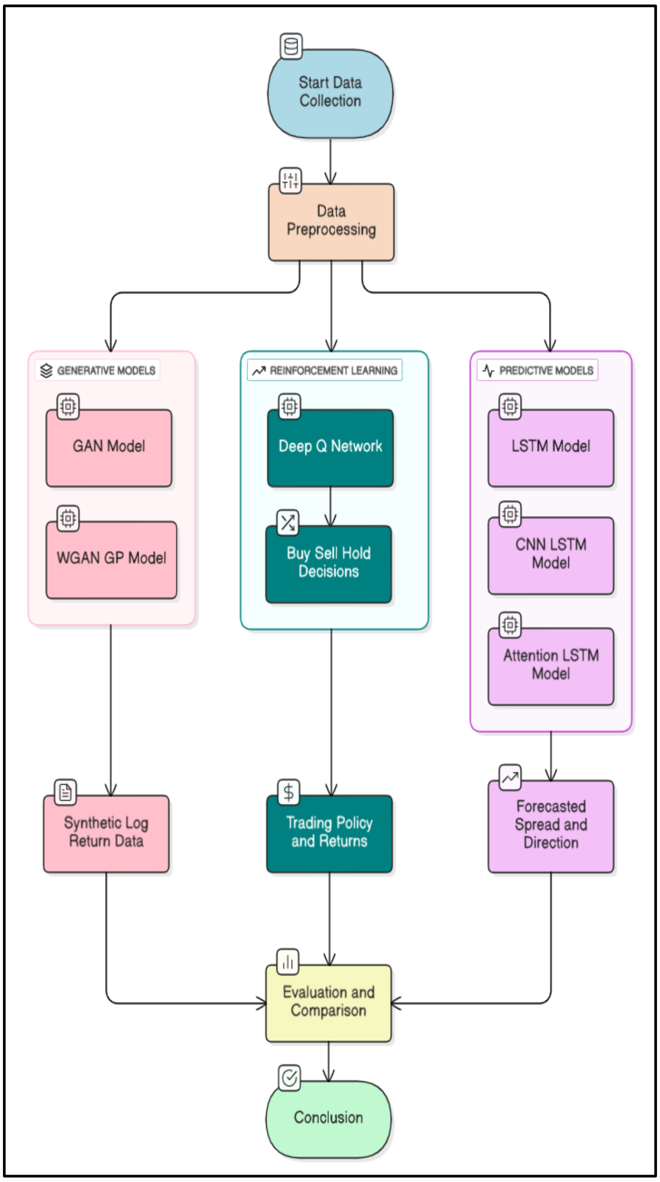

Fig. 1. Proposed deep learning framework for statistical arbitrage

A. Data Collection and Pre-Processing

Using public APIs , the data collected for this framework were daily closing prices of six NSE-listed equities RELIANCE.NS, TCS.NS, HDFCBANK.NS, INFY.NS, ICICIBANK.NS, and HINDUNILVR.NS. The values were converted to Log-returns to make it stable by standardizing percentage changes in price. Additional statistical and technical indicators were derived to further enhance the feature representation, including the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands. In order to ensure numerical stability during gradient based optimization we normalize each feature via x-score scaling. The processed data were split into 80 % training, 10 % validation, and 10 % testing windows. Time series windows of length 60 were generated to capture temporal dependencies where each window produced one forecasted spread or trading signal label. This dataset serves as a structured input for all the models within this framework.

B. Predictive Model Architecture

Three supervised sequence learning models were used to forecast predicted spread direction as well as the scale.

1. Long Short-Term Memory (LSTM): A baseline model consisting of stacked LSTM layers with dropout regularization, is used in modeling sequential dependencies in long term and mitigating vanishing gradients related issues. A sigmoid activation function is applied in the output layer to classify the binary output of arbitrage opportunities.

2. Convolutional LSTM (CNN-LSTM): A hybrid architecture which in order to extract spatial patterns in feature spaces in short term applies 1-D convolutions, followed by LSTM layers for temporal sequence modeling. The convolutional stage improves short-term pattern capture and noise resilience, while recurrent layers preserve long-term structure. This joint spatial-temporal modeling enhances stability under regime shifts.

3. Attention-Enhanced LSTM: It acts as an extension of the LSTM that applies an attention mechanism where adaptive weights are assigned to the prior hidden states. This assists the model in focusing only on the crucial time steps under dynamic market conditions.

The Adam optimizer (learning rate = 0.001) and binary-cross-entropy loss for 100 epochs with early stopping based on validation loss is applied in each predictive model. The performance was evaluated via Accuracy, Precision, Recall and F-1 score to assess directional forecasting reliability under imbalanced market data.

C. Reinforcement Learning Model

The Deep Q-Network (DQN) agent was modeled to make decisions in an environment where there was the latest spread, indicators, and market momentum. There were only three actions where one would Buy, Sell, or Hold, and the reward function was the same as the Profit and Loss, with the aim of ensuring risk-adjusted returns.

The NN architecture was trained with Q-learning with experience relay and target network updates where the agent employed the ε-greedy policy to govern balanced exploration with a decay from 1.0 to 0.05 over 10^4 episodes.For the training the architectures consisted of three fully connected layers (256–128–64 neurons) with ReLU activation functions. For the agent to find quantified profitability, performance and capital preservation, evaluations were done with regards to Cumulative Return, Sharpe Ratio, and Maximum Drawdown in order to study economic relevance beyond just its accuracy of prediction.

For incorporating a more realistic trading environment, the RL state space was augmented from raw spreads to include technical indicators of rolling volatility and short-term trend signals. Transaction costs were added to penalize heavy trading, while a mild drawdown penalty favored risk-sensitive behavior. Important hyperparameters explored and tuned through grid search for improving stability included the learning rate, replay buffer size, and target update frequency.

By enriching the market context and accounting for practical constraints, the adjustments enabled the agent to learn more robust policies that were economically meaningful.

D. Generative Modeling with GAN and WGAN-GP

This modelling technique aims to synthesize realistic financial time-series data for stress-testing and data-augmentation purposes.

1. Standard GAN: In which there exists a generator and discriminator trained via adversarial loss. The generator map noise vectors $z \in \mathbb{R}^{100}$ to synthetic return sequences of length 60, while Discriminator determines which is real and which is generated. However, the vanilla GAN suffered from mode collapse and unstable convergence.
2. Wasserstein GAN with Gradient Penalty (WGAN-GP): Introduces the Earth-Mover (Wasserstein-1) distance as the optimization objective and a gradient-penalty term to enforce 1-Lipschitz continuity. The generator contained approximately 715 k parameters with residual blocks, while the critic used 318 k parameters. Training demonstrated smooth loss convergence and higher fidelity synthetic returns that closely matched real-data statistics (mean, variance, kurtosis).

Additional pre-processing steps were performed before training the generative models in order to enhance further the robustness of modelling. Log returns were standardized to remove scale differences, and rolling volatility estimates were included as conditioning variables during experimentation, which allowed the generator to learn volatility-dependent patterns. The training process also used a number of additional techniques, including gradient clipping and mini-batch standard deviation, in order to reduce instability and further encourage the diversity of generated samples. In WGAN-GP, multiple critic updates per generator iteration were made to better approximate the Wasserstein distance. These refinements collectively improved training stability, with the result that sharper regime shifts and more realistic tail behaviors could be captured by the models in synthetic financial time series.

For evaluation of the generated samples Wasserstein distance and mean-squared error (MSE) between real and synthetic data were done. This helped verify that WGAN-GP is more stable and generates more realistic data compared to the standard GAN.

IV. RESULTS AND DISCUSSION

This section presents the comparative analysis of the three paradigms predictive, reinforcement, and generative evaluated using identical data partitions and market conditions.

A. Predictive Model Performance

The predictive models were trained on rolling windows of length of 60 days consisting of stock returns with corresponding technical indicators. Table I lists out the comparative metrics.

TABLE I. PERFORMANCE OF PREDICTIVE MODELS

Aspects	Models		
	LSTM	**CNN-LSTM**	**Attention-LSTM**
Accuracy (%)	52.6	**70.2**	**81.3**
Precision	0.51	0.66	0.00
Recall	0.54	0.36	0.00
F1-Score	0.52	0.47	0.00
Observations	Baseline; stable but limited generalization.	Best overall balance; captures spatial and temporal features.	High accuracy but failed to detect minority class (class imbalance).

Out of these models CNN-LSTM achieved the preferred balance between accuracy and recall, demonstrating the benefits of enhanced short and long term dependency modeling achieved by convolutional and recurrent layers. The Attention-LSTM model showed the highest numerical accuracy, its performance was biased towards the dominant class, resampling or focal loss adjustments in imbalanced datasets will be needed for this LSTM. The basic LSTM provided moderate but stable forecasts, which makes it end up serving as a baseline.

B. Reinforcement Learning: Trading Performance

The Deep Q-Network (DQN) agent by directly optimizing decision policies demonstrated the highest profitability in place of relying solely on static results. Backtesting over the same equity basket yielded a cumulative return of +42.98%, Sharpe ratio of 0.73 and maximum drawdown of 11.2%.

Compared to the predictive models, DQN exhibited lower sensitivity to noisy market inputs and dynamically adjusted its buy/sell behavior based on market momentum. This reinforces the observation that reinforcement learning frameworks can outperform static predictors in environments with high regime volatility by learning context-aware trading policies.

C. Generative Modeling Results

The GAN and WGAN-GP models were evaluated on their ability to synthesize realistic data. Table II measures the model's performance between real and synthetic return distributions.

TABLE II. GENERATIVE MODEL EVALUATION

Aspects	Models	
	GAN	WGAN-GP
Generator Params	52.6	**81.3**
Critic Params	0.51	0.00
Wasserstein Distance ↓	0.54	0.00
MSE ↓	0.52	0.00
Observations	Training unstable; partial mode collapse..	Stable convergence; realistic, smooth synthetic data.

The WGAN-GP presented smaller values for Wasserstein distance and MSE, which indicated higher accuracy and robustness than the vanilla GAN model. The sequences obtained from WGAN-GP almost approximated the real market data values. The results can aid in validating current work concerning the significance of gradient penalties and Wasserstein distances in ensuring stable adversarial learning in financial data settings.

D. Cross-Paradigm Comparison and Discussion

To contextualize performance, Table III compares the three types of models on their respective evaluation based dimensions.

TABLE III. OVERALL MODEL COMPARISON

Aspects	Paradigm		
	Predictive	***Reinforcement***	***Generative***
Representative Model	CNN-LSTM	DQN	WGAN-GP
Key Metric	Accuracy: 70.2%	Return: +42.98%	W. Distance: 0.082
Score / Observation	Balanced prediction and temporal learning	Adaptive decision-making; best profitability	High-fidelity synthetic data; robust convergence

Thus, we see that there is no single paradigm that outperforms in every aspect. The predictive models work well in terms of directional performance but fail in handling market imbalance and overfitting tasks. The reinforcement models optimize profits in varying market conditions but involve significant exploration and reward engineering work. The generative models play an indirect role in improving data exploration and stress testing.

The above result of the accuracy-profitability trade-off makes it obvious that although the models with the highest predictive accuracy may be the best, they may not translate to high returns on the trades. Quite the contrary, the maximization of rewards in reinforced learning agents obviously causes improved performance with reduced precision. In addition, the introduction of generative models assists in the robustness and synthetic data set test of the future models.

Combined, the three paradigms demonstrate the ability of a single deep learning pipeline to reinforce statistics arbitrage strategies beyond what is obtained separately by each model. With predictive models, it is possible to generate the same signal repeatedly; the agents of reinforcement learning can translate the signal into adaptive trade execution and the generative models can offer an assessment of the strength of strategies in various or extreme market environments. The complementary interaction is also becoming well represented in quantitative finance research in a number of hybrid frameworks that tie forecasting, decision-making and stress testing into a unified knot. This multi-layered design is not only useful to alleviate the flaws of the individual models but also to take the development of trading systems to a higher level of tolerance and scalability.

V. CONCLUSION

For this purpose, a new integrated deep learning model has been designed and tested with the predictive model, reinforcement model, and generative model paradigms in statistical arbitrage related to six equities that were listed on the National Stock Exchange. The model combines LSTM models for forecasting, Deep Q-Network for dynamic trading, and WGAN with GP for generative tasks.

The results showed that there was no paradigm with overarching supremacy in terms of objectives. The CNN-LSTM was best in terms of well-rounded predictive capability via efficient spatial-temporal feature extraction, while the DQN agent has better profitability in terms of direct optimization of cumulative rewards with dynamic market adjustments. The WGAN-GP was able to synthesize realistic return distributions, ensuring model robustness in terms of data scarcity and regime variations.

In general, the literature emphasizes the significance of the accuracy-profitability trade-off – while models designed to enhance accuracy may not necessarily serve to maximize profitability, whereas the agents engaged in reinforcement learning in pursuit of reward functions can provide higher levels of economic performance. When combined together, the models for predictions improve signal production, dynamic agents facilitate execution, while generative models not only facilitate stress testing and data augmentation but can even address data challenges with the capability to simulate stress scenarios related to market data in extreme market settings.

VI. FUTURE SCOPE

The future work can be extended on the basis of enhanced architecture and integration techniques. There are various models based on the transformer architecture, such as Temporal Fusion Transformers (TFT) and Longformer, which can detect the seasonal patterns in the market not captured by recurrence models. There is multi-agent reinforcement learning, which can help in arbitrage across multiple assets if they are correlated.

Another area is the hybrid predictive/generative pipelines, in which synthetic data from WGAN-GP can be used to pre-train predictive and reinforcement learning agents to help them generalize better and perform well even in volatile market settings with low market liquidity. The final aspect would be to implement the proposed system on real market data flows to estimate the latency and scalability of the proposed system and the level of sensitivity to transaction costs. This gives rise to autonomous and explainable arbitrage systems that can suit real market settings.

The other area is the hybrid predictive/generative pipelines where the synthesized data of WGAN-GP may be employed to pre-train predictive and reinforcement learning agents to assist them in generalizing better and perform well even in the volatile market conditions with low market liquidity. The last would be the implementation of the proposed system on actual market data flows to determine the anticipated performance of the proposed system in terms of latency and scalability and degree of sensitivity to transaction costs. This leads to arbitrary and explainable arbitrage regimes that may be applicable in actual markets. Future research directions also involve incorporation of regime-detection mechanisms which would automatically define structural breaks, clusters of volatility, or macroeconomic changes before models are run. To this end, it can be considered that the incorporation of regime classifiers into an arbitrage pipeline can potentially allow a predictive, RL, and generative component to dynamically change or re-calibrate their strategy according to the current market conditions. The other equally promising direction is the creation of explainable AI in deep learning models which measure feature importance, trade reasoning or risk attribution. This is quite basically an obligatory condition to the adoption of such models in actual institutional trading. Online learning and continual learning architectures would be implemented in such a manner that the models could be updated at a runtime when data is fed in through a stream to avoid degradation during silent market periods. Lastly, the further widening of the single structure with portfolio-based optimization, cross-asset hedging and multi-market arbitrage would quite literally broaden its range of use and bring the system a step closer to a completely autonomous, quantitative trading machine of scalable size.

REFERENCES

[1] M. Barua et al., "Comparative Analysis of Deep Learning Models for Stock Price Prediction in the Indian Market," FinTech, vol. 3, no. 4, pp. 551–568, 2024.

[2] M. R. Kabir et al., "LSTM–Transformer-Based Robust Hybrid Deep Learning Model for Financial Time Series Forecasting," Sci, vol. 7, no. 1, Art. 7, 2025.

[3] J. Dong and S. Liang, "Hybrid CNN-LSTM-GNN Neural Network for A-Share Stock Prediction," Entropy, vol. 27, no. 8, Art. 881, 2025.

[4] Y. Chen et al., "EMAT: Enhanced Multi-Aspect Attention Transformer for Financial Time Series Forecasting," Entropy, vol. 27, no. 10, Art. 1029, 2025.

[5] K. Gajamannage et al., "Real-Time Forecasting of Time Series in Financial Markets Using Sequentially Trained Dual-LSTM," Expert Syst. Appl., vol. 223, Art. 119879, 2023.

[6] M. Zhu et al., "Improving LSTM Networks for Arbitrage Spread Forecasting: Integrating Cuckoo and Zebra Algorithms," PeerJ Comput. Sci., vol. 10, 2024.

[7] A. Sebastian and V. Tantia, "Multi-Variate LSTM with Attention Mechanism for the Indian Stock Market," Int. J. Inf. Manag. Data Insights, vol. 5, no. 2, Art. 100350, 2025.

[8] Z. Zeng et al., "Financial Time Series Forecasting Using CNN and Transformer," in Proc. AAAI Workshop on AI for Financial Services, 2023.

[9] Y. Huang et al., "A Self-Rewarding Mechanism in Deep Reinforcement Learning for Trading Strategy Optimization," Mathematics, vol. 12, no. 24, Art. 4020, 2024.

[10] R. G. Mishra et al., "A Deep Reinforcement Learning Framework for Strategic Indian NIFTY 50 Index Trading," AI, vol. 6, no. 8, Art. 183, 2025.

[11] M. Kong and J. So, "Empirical Analysis of Automated Stock Trading Using Deep Reinforcement Learning," Appl. Sci., vol. 13, no. 1, p. 633, 2023.

[12] J. Zou et al., "A Novel Deep Reinforcement Learning-Based Automated Stock Trading System Using Cascaded LSTM Networks," Expert Syst. Appl., vol. 242, Art. 122801, 2024.

[13] M. Xu et al., "Deep Reinforcement Learning for Quantitative Trading," arXiv:2312.15730, Dec. 2023.

[14] M. Wei et al., "Multi-Agent Reinforcement Learning for High-Frequency Trading Strategy Optimization," preprint, Nov. 2024.

[15] F. Ramzan et al., "Generative Adversarial Networks for Synthetic Data Generation in Finance: Evaluating Statistical Similarities and Quality Assessment," AI, vol. 5, no. 2, pp. 667–685, 2024.

[16] K. Zakharov et al., "Synthetic Financial Time Series Generation with Regime Clustering," J. Adv. Inf. Technol., vol. 14, no. 6, pp. 1372–1381, 2023.

[17] D. E. Allen et al., "GANs and Synthetic Financial Data: Calculating VaR," Appl. Econ., vol. 57, no. 37, pp. 5680–5695, 2025.

[18] T. Takahashi and T. Mizuno, "Generation of Synthetic Financial Time Series by Diffusion Models," Quant. Finance, early access, 2025.

[19] T. Sattarov et al., "FinDiff: Diffusion Models for Financial Tabular Data Generation," in Proc. ACM ICAIF 2023, 2023.

[20] M. Rizzato et al., "Generative Adversarial Networks Applied to Synthetic Financial Scenarios Generation," Physica A, vol. 623, Art. 128899, 2023.

[21] A. Sedai et al., "Augmenting Bank and Economic Data with Generative Adversarial Networks," Procedia Comput. Sci., vol. 263, pp. 175–182, 2025.