**DEVELOPMENT OF HYBRID DEEP LEARNING AND REINFORCEMENT LEARNING FOR INTELLIGENT TRADING IN FOREX MARKETS**

**Abstract**

Forex trading, a $6 trillion a day market, remains challenging to work with due to its non-linearity, volatility, and noise. Most traditional approaches like ARIMA and SARIMA fail to capture these dynamics and hence tend to be low in forecasting accuracy. In this study, we propose a hybrid of Convolutional Neural Network - Long Short-Term Memory (CNN-LSTM) for forex price forecasting and Deep Q-Learning (DQL) for trading strategy optimization. The CNN-LSTM model learns spatial-temporal features from past data, measured by MSE, RMSE, and MAE. For decision-making in trading, a DQL agent learns to maximize actions—buy, sell, or hold—through reward-based learning with epsilon-greedy exploration, experience replay, and target networks. Results show CNN-LSTM outperforms traditional models, achieving lower RMSE (0.0025) and MAE (0.0017). The DQL agent achieves a cumulative return of 49.2% and a Sharpe ratio of 2.87, surpassing rule-based methods. An ablation analysis corroborates the necessity of the vital elements such as experience replay and target networks for enabling stable learning. Statistical testing methods such as the Diebold-Mariano test corroborate the prediction prowess of the model. The hybrid model shows great promise for real-time forex trading with high precision and strong risk-adjusted returns. Future studies must integrate macroeconomic variables, sentiment analysis, and multi-asset portfolios to facilitate generalization and trading performance.

**Keywords:** CNN-LSTM, Deep Q-Learning, Forex Trading, Time Series Forecasting, Reinforcement Learning

**1 Introduction**

The Forex market is a decentralized global currency market that trades 24 hours a day, five days a week, with a daily trading volume of more than $6 trillion. It facilitates international trade and investments and is motivated by economic fundamentals, central bank policy, and geopolitical events [1] describes the activities of forex market participants—including brokers, institutional investors, and retail traders—who trade currency pairs such as EUR/USD on decentralized over-the-counter (OTC) networks via platforms like TradingView and MetaTrader, utilizing both technical and fundamental analysis.

Despite its potential for profitability, Forex trading remains risky due to market volatility, leverage, and emotionally driven decisions. Accurate forecasting of price movements is essential for reducing uncertainty and identifying profitable opportunities. However, classical statistical models—such as ARIMA, Linear Regression, and SARIMA—struggle to cope with the high-frequency and non-linear nature of the market, limiting their adaptability in dynamic environments [2].

AI and ML models such as CNNs and LSTMs deliver better predictive capability by discovering complex patterns within financial time series. These models can leverage unstructured data such as news sentiment and geopolitical conflicts to improve the accuracy of predictions. Machine learning in finance has widened, with applications of deep learning architectures and ensemble methods such as XGBoost for price prediction and strategy development.

Reinforcement learning, specifically Deep Q-Learning (DQL), complements trading approaches in terms of constantly tuning decisions according to market interaction. This research suggests a hybrid CNN-LSTM model for accurate Forex forecasting, along with DQL for reward-based improvement of the trading strategy, thus providing real-time flexibility and smarter automated trading systems.

**2 Related Works**

Recent empirical studies underscore the increasing effectiveness of hybrid deep learning (DL) and reinforcement learning (RL) models in Forex trading. [2] proposed a model that integrates deep learning on order books with temporal-difference RL to forecast returns across multiple horizons based on order flow imbalance. Applied to several Forex pairs, including GBP/USD and EUR/USD, the model exhibited promising profitability; however, it required adjustments to account for retail trading costs and market volatility. Similarly, [3] presented a multi-agent reinforcement learning framework utilizing the A3C algorithm, enabling concurrent learning across diverse currency pairs and significantly improving trading performance by accelerating the learning of profitable strategies.

[4] developed a deep long-short memory model that performed credibly well in predicting most Nigerian equity stocks when subjected to reliable historical training dataset. It has an average of 97% predictive accuracy and 3% error rate. In their findings they concluded that LSTM models only need well-cleaned and normalized huge dataset to perform at its best.

[5] suggested a hybrid LSTM-CNN model for Forex direction prediction using high-frequency Bloomberg and OANDA data. LSTM captured temporal dependencies, and CNN learned spatial features from candlestick charts and technical indicators. Random search and Bayesian optimization were employed for tuning, with a 27% RMSE improvement over individual LSTM models. This showed the value of combining sequential and spatial learning for better Forex predictability.

[6] suggested a multi-feature stock price prediction model in the framework of LASSO feature selection and cascade LSTM (Ca-LSTM) architecture design. The framework enhanced the ability of the model to remember beneficial information and capture volatility, outperforming the likes of XGBoost and PSO-SRNN models. Although being computationally complex and parameter-dependent, the model validated multi-feature fusion as one promising financial forecasting direction.

[7] proposed an LSTM-DQN hybrid model for USD/JPY and EUR/USD trading by combining LSTM trend prediction with DQN-based decision optimization. After accounting for slippage and trading fees, this model outperformed single LSTM models by 14.7% versus 8.2% annual return, demonstrating the benefit of RL in trade execution dynamics.

[3]also examined a multi-agent RL approach with A3C on EUR/USD, GBP/USD, and AUD/JPY, where agents were trained asynchronously. This system improved diversification of the strategy, was 22% more profitable than single-agent systems, and converged more quickly. [8] also used GBP/USD and EUR/USD order book data from LSE and Binance to show that their hybrid DL-RL approach—combining DNNs for order imbalance prediction and RL for execution—achieved a 12.5% higher Sharpe ratio and reduced transaction costs by 18%.

Overall, empirical evidence supports the advantages of DL-RL hybrids in Forex: predictive accuracy, adaptive policy learning, and risk-adjusted returns. Challenges remain, though, such as computational heaviness, overfitting risks, and interpretability. Future work must overcome those through the application of Explainable AI (XAI), federated learning for decentralized trading, and quantum deep learning for time series pattern recognition.

**3. Methodology**

This paper proposes a hybrid model for predicting foreign exchange (forex) price movements using a blend of Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Deep Q-Learning models. Historical daily Open, High, Low, Close, and Volume (OHLCV) data of EUR/USD symbol were collected from Deriv.com for the period January 1, 2014, to December 31, 2024. In addition to this, five different technical indicators such as. Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Rate of Change (ROC), Exponential Moving Average (EMA), and Stochastic Oscillator—were computed from the price series for additional input to the model.

Pre-processing of the data included normalization using min-max scaling for allowing equal feature contribution at the time of training. Missing values, occurring due to market closure or interruptions in data feeds, were addressed using forward-fill and linear interpolation techniques. Noise removal and trend clarity on the price data were achieved using EMA smoothing.

A sliding window strategy was used for feature extraction, which converted the time-series data into multivariate input-output pairs that are readily amenable to sequence modeling. A sample reflected a fixed window size spanning several features. Before the windowing operation was applied, the dataset was split sequentially into training, validation, and test sets in order to maintain the temporal ordering and prevent data leakage.

The proposed architecture utilizes CNN layers to obtain spatial features from the structured input data. They are then input to LSTM layers, which capture temporal dependencies. The Deep Q-Learning module incorporated introduces a reward-based optimization process that allows the model to learn decision policies that maximize prediction accuracy and minimize classification errors. This hybrid model takes advantage of the complementary capabilities of CNN, LSTM, and reinforcement learning to adaptively learn the intricate, nonlinear, and non-stationary forex market dynamics, thereby yielding improved predictive performance. Figure 1 describes the summary of methods applied in the study.

Figure 1: Diagram that displays the methods and procedures

### Developmental Tools Utilized for Model Implementation

The model was implemented using Python with TensorFlow/Keras for deep learning, Stable Baselines3 for reinforcement learning, and Pandas/NumPy for data processing. TA-Lib computed technical indicators. Visualization was done using Matplotlib and Seaborn. Training was accelerated using NVIDIA GPU and cloud platforms like Google Colab Pro and Kaggle Kernels.

**Model Evaluation**

Hybrid CNN-LSTM-DQN was evaluated on forecasting performance metrics—RMSE, MAE, and MAPE—and trading performance metrics such as Sharpe Ratio, Maximum Drawdown, and Total Profit [9]. It was compared with benchmark models such as ARIMA [10]**,** LSTM alone, and Moving Average. The Diebold-Mariano test at α = 0.05 was used to test statistical significance [11]. An ablation study evaluated the separate contributions of CNN, LSTM, and DQN components.

## **4 Results and Discussion**

### ****Model Training Dynamics and Early Stopping****

The CNN-LSTM model demonstrated rapid convergence, with both training and validation loss decreasing significantly within the first 50 epochs. Validation loss plateaued around epoch 50, while training loss continued to drop slightly. Early stopping was triggered at epoch 70 (patience = 15), successfully preventing overfitting.



Figure 2: Training and Validation Loss Curves (Early Stopping at Epoch 70)

Table 1: Interpretation of Training & Validation Loss Curves

| **Observation** | **Insight** |
| --- | --- |
| Loss Plateau | Indicates convergence; no gain in prolonged training |
| Minimal Loss Gap (Δ = 0.0015) | Suggests strong generalization performance |
| Dropout & L2 Regularization | Helped stabilize learning and avoid overfitting |

###  ****Forecasting Performance of CNN-LSTM****

####  **Forecast vs. Actual Prices**

The hybrid model tracks real forex prices with high fidelity. It captures both short-term fluctuations and broader trends effectively.



Figure 3: Actual vs. Predicted Close Prices for EUR/USD

#### **Accuracy Metrics Comparison**

CNN-LSTM significantly outperformed ARIMA, standalone LSTM, and Moving Average (MA) models across all error metrics.

**Table 2**: Forecasting Accuracy Comparison

| **Model** | **RMSE** | **MAE** | **MAPE (%)** |
| --- | --- | --- | --- |
| ARIMA | 0.0048 | 0.0035 | 0.82 |
| Standalone LSTM | 0.0032 | 0.0021 | 0.54 |
| Moving Average | 0.0051 | 0.0040 | 0.91 |
| **CNN-LSTM (Hybrid)** | **0.0025** | **0.0017** | **0.39** |

**Observation**: The hybrid model reduced RMSE by **51.2% vs ARIMA** and **21.8% vs standalone LSTM**.

#### **Statistical Significance (Diebold-Mariano Test)**

**Table 3**: Statistical Comparison of Forecasting Accuracy

| **Benchmark** | **DM Statistic** | **p-value** | **RMSE Reduction (%)** |
| --- | --- | --- | --- |
| ARIMA | 4.72 | <0.001 | 51.2 |
| LSTM | 3.21 | 0.002 | 21.8 |
| Moving Average | 5.89 | <0.001 | 49.6 |

**Conclusion**: CNN-LSTM superiority is statistically validated (p < 0.05) across all benchmarks.

### ****Forecasting Error Distribution****

The error histogram reveals slight negative skewness (γ₁ = –0.45) and high kurtosis (γ₂ = 4.7), indicating a fat-tailed error distribution with rare but significant outliers—likely during geopolitical shocks.



Figure 4: Histogram of Forecasting Errors

**Observation**: Although mean error is low (RMSE = 0.0025), tail risks persist, suggesting the use of Value-at-Risk (VaR) for risk control.

### ****Trading Strategy Optimization with Deep Q-Learning****

####  **DQN Performance vs. Benchmarks**

**Table 4**: Trading Metrics Comparison

| **Metric** | **Hybrid DQL** | **MA Crossover** | **Buy-and-Hold** |
| --- | --- | --- | --- |
| Sharpe Ratio | **2.87** | 1.45 | 1.12 |
| Max Drawdown (%) | **8.9** | 12.7 | 18.3 |
| Cumulative Profit (%) | **49.2** | 31.8 | 24.5 |
| Win Rate (%) | 62.1 | 54.3 | — |

**Conclusion**: DQL optimizes return-risk balance significantly better than traditional methods.

#### **Q-Learning Dynamics**



Figure 5: Exploration-Exploitation Curve over Training Episodes



Figure 6: Q-Value Convergence

* Exploration was emphasized early (ϵ = 1.0), but as ϵ decayed to 0.1, exploitation dominated.
* Q-values stabilized after episode 300, indicating reliable policy convergence.

### ****Ablation Study Results****

#### **Impact on Forecasting**

**Table 5**: Forecasting Impact of Model Component Removal

| **Variation** | **RMSE** | **MAPE (%)** | **R² Score** |
| --- | --- | --- | --- |
| Full CNN-LSTM | 0.0025 | 0.39 | 0.91 |
| Without CNN | 0.0032 | 0.54 | 0.85 |
| Without LSTM | 0.0041 | 0.67 | 0.76 |
| Without CNN & LSTM | 0.0056 | 0.82 | 0.61 |

**Observation**: LSTM contributes more to temporal coherence; CNN enhances spatial pattern extraction. Their synergy is critical.

####  **Impact on Trading Strategy**

**Table 6**: DQN Performance under Component Removal

| **Component Removed** | **Sharpe Ratio** | **Max Drawdown (%)** | **Cumulative Profit (%)** |
| --- | --- | --- | --- |
| Full DQL Model | 2.87 | 8.9 | 49.2 |
| No Experience Replay | 1.92 | 14.3 | ↓ |
| No Target Network | 2.15 | 11.6 | ↓ |
| No Epsilon Decay | 1.45 | 17.8 | ↓ |

**Observation**: Each DQL component is essential for balancing risk and return. Removing them significantly degrades performance.

**5 Conclusion**

This study investigated the use of CNN-LSTM for forex price forecasting and Deep Q-Learning (DQL) for trading strategy optimization. Traditional models like ARIMA and Moving Averages struggle with the forex market’s non-linearity and volatility. CNN-LSTM effectively captured spatial-temporal features, outperforming baseline models in RMSE, MAE, and MAPE. DQL-based strategies showed higher profitability and lower drawdowns than rule-based approaches.

Ablation studies validated the importance of Experience Replay, Target Networks, and Epsilon Decay in stabilizing DQL. The findings demonstrate that hybrid deep learning models significantly enhance forecasting accuracy, while reinforcement learning improves decision-making by dynamically adapting to market trends.

Future research should explore hyperparameter tuning using Bayesian Optimization, extend models to multi-asset trading, and investigate hybrid reinforcement learning with multi-agent or transformer-based architectures.

# References

|  |  |
| --- | --- |
| [1]  | M. A. Khan, "Forex Trading: A Beginner's Guide.," *Springer,* 2018.  |
| [2]  | Jaddu, M., & Bilokon, P., "Deep Learning and Reinforcement Learning for Optimal Forex Execution: Evidence from LSE and Binance Order Books.," *Quantitative Finance,* vol. 23, no. 9, p. 1478–1492, 2023.  |
| [3]  | Sarani, H., & Rashidi-Khazaee, "A. Multi-Agent Reinforcement Learning for Diversified Forex Strategies Using A3C," *Expert Systems with Applications,* vol. 236, p. 120419, 2024.  |
| [4]  | Ojo, Adebola K., and Ifechukwude Jude Okafor, "Forecasting Nigerian Equity Stock Returns Using Long Short-Term Memory Technique," *Journal of Advances in Mathematics and Computer Science,* vol. 39, no. 7, pp. 45-54, 2024.  |
| [5]  | Faru, S. H., Waititu, A., & Nderu, L., " A Hybrid Neural Network Model Based on Transfer Learning for Forecasting Forex Market," *Journal of Data Analysis and Information Processing,* vol. 11, no. 2, pp. 103-120, 2023.  |
| [6]  | X. Chen, Y. Li, H. Wang, and Z. Zhang, "A multi-feature stock price prediction model based on multi-feature calculation, LASSO feature selection, and Ca-LSTM network," *Connection Science,* vol. 36, no. 1, p. 2286188, 2024.  |
| [7]  | Madhulatha, T., & Ghori, Q. , "LSTM-DQN Hybrid Approach for Intelligent Forex Trading: A Case Study on USD/JPY and EUR/USD," *Journal of Financial Data Science and Machine Learning,* vol. 6, no. 1, p. 45–59, 2024.  |
| [8]  | Jaddu, K. S., & Bilokon, P. A., "Combining Deep Learning on Order Books with Reinforcement Learning for Profitable Trading," *arXiv preprint arXiv:2311.02088,* 2023.  |
| [9]  | Zhang, X., Li, Y., & Xu, Y. , "A hybrid approach for financial time series forecasting using CNN and LSTM," *Journal of Financial Data Science,* vol. 3, no. 2, p. 57–74, 2021.  |
| [10]  | Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. , "Time Series Analysis: Forecasting and Control," *Wiley,* 2015.  |
| [11]  | Diebold, F. X., & Mariano, R. S. , "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics,* vol. 13, no. 3, p. 253–263, 1995.  |