Journal of Artificial Intelligence General Science (JAIGS)

ISSN: 3006-4023 (Online), Volume 6 , Issue 1, 2024 DOI: 10.60087

Home page https://ojs.boulibrary.com/index.php/JAIGS

Optimization of High-Frequency Trading Strategies Using Deep Reinforcement Learning

Guanghe Cao [1*] , Yitian Zhang [1,2] , Qi Lou [2] , Gaike Wang [3]

1.* Computer Science, University of Southern California, CA,USA
1.2. Accounting, UW-Madison, WI, USA
2. Tian Yuan Law Firm, Hang Zhou, China
3. Computer Engineering, New York University, NY, USA
**Corresponding author E-mail: rexcarry036@gmail.com*

ABSTRACT

This study presents a new method for optimising high-risk trading (HFT) strategies using deep learning (DRL). We propose a multi-time DRL framework integrating advanced neural network architectures with sophisticated business data processing techniques. The framework employs a combination of convolutional neural networks for manual order analysis, short-term memory networks for time series processing, and a multi-head listening mechanism for body fusion. We formulate the HFT problem based on Markov Decision Processes and use the Proximal Policy Optimization algorithm for training. The model is evaluated using tick-by-tick data from the NASDAQ exchange, including ten liquid stocks in 6 months. The experimental results show the superiority of our method, achieving a Sharpe ratio of 3.42, outperforming the learning model and machine learning based on benchmarks up to 33%. The proposed strategy has demonstrated strong performance across a wide range of regulatory markets and has shown potential for strategic objectives. Sensitivity analysis confirms the model's stability across a range of hyperparameters. Our findings suggest that the DRL-based approach can improve HFT performance and provide better market adaptation and risk management. This research leads to the continuous evolution of algorithmic trading strategies and shows the potential of AI-driven approaches in financial markets.

Keywords: High-Frequency Trading, Deep Reinforcement Learning, Multi-Time Scale Analysis, Algorithmic Trading Optimization

ARTICLE INFO: *Received: 01.10.2024 Accepted: 20.10.2024 Published: 05.11.2024*

1. Introduction

1.1. Background of High-Frequency Trading

High-frequency trading (HFT) has changed the financial industry in the last two decades, and technology and regulatory changes have changed how trading is done. HFT refers to using computer algorithms to make high-speed trades, usually in microseconds or milliseconds[1] . This business method aims to invest in small and unprofitable businesses, profiting from high volumes and low turnover per business.

The emergence of HFT can be traced back to the early 2000s, with its widespread use in electronic trading and regulations to promote competition[2] . These developments lead to an explosion of financial markets, creating new opportunities for traders to use different prices across different areas. HFT firms invest heavily in technology, including high-speed connections, integrated services, and advanced equipment, to gain a competitive edge. for ultra-low latency[3] .

HFT strategies include a variety of methods, including market, statistical, and market-driven. These strategies rely on quick market data, complex mathematical models, and sophisticated techniques to identify and execute profitable trades[4] . The spread of HFT has dramatically affected the market microstructure, efficiency, and value-finding process in the financial market.

1.2. Challenges in High-Frequency Trading

Although it has the potential to be very profitable, HFT faces many challenges that require innovation and change. One of the main issues is the arms race for speed, as merchants compete to reduce latency and gain physical advantage over their competitors[5] . This competition has led to significant investment in infrastructure, including developing specialised equipment and optimising network connections.

Another critical challenge is developing and maintaining business algorithms that quickly adapt to the business environment. HFT strategies must account for market microstructure noise, order book strength, and the impact of their trading on market prices[6] . In addition, the increasing complexity of the financial market, with various asset classes and interrelated business entities, requires risk management and enforcement processes—the best.

Regulatory controls and market changes pose additional challenges for HFT firms. Regulatory agencies have implemented measures to reduce the risks associated with HFT, such as electronic devices and order trading. These restrictions require HFT strategies to be followed by for-profit management[7] .

1.3. Deep Reinforcement Learning in Financial Markets

The application of deep learning (DRL) in the financial industry has gained significant attention in recent years, providing a promising way to solve the problems faced by HFT. DRL combines the power of deep neural networks with support learning, enabling the development of adaptive strategies that can be learned through business interactions and improve their performance over time.

DRL algorithms, such as Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), and Advantage Actor-Critic (A2C), have proven to be very effective in solving complex operational problems. In the context of financial transactions, DRL agents can be trained to make business decisions based on advanced business information, learning to balance trade-offs between rewards and immediate and long-term benefits[8] .

The use of DRL in HFT offers several advantages. DRL agents can automatically discover optimal trading strategies without relying on hand-crafted rules or explicit market models. They can adapt to changing market conditions and learn from their interactions with the environment, potentially uncovering novel trading patterns and opportunities[9] . Moreover, DRL frameworks can incorporate risk management constraints and regulatory requirements directly into the learning process, ensuring compliance with market regulations.

1.4. Research Objectives and Contributions

This research focuses on optimising business strategies using deep learning. The primary purpose of this study is to develop a DRL-based framework specifically for HFT, create new network architectures and algorithms that can capture the intricacies of high-frequency market dynamics, and evaluate the performance of the strategic approach against existing HFT strategies[10] .

The results of this research are many. The DRL framework for HFT is presented, including advanced techniques such as multi-agent learning and hierarchical learning to further solve the complexity of high-frequency trading[11] . Novel network architectures are proposed, leveraging recurrent neural networks and monitoring mechanisms to process time-series data and capture long-term dependencies efficiently.

Additionally, this study introduces innovative data preprocessing and feature engineering techniques designed explicitly for HFT applications. These techniques aim to extract relevant information from high-frequency order book data and market microstructure signals, enhancing the learning efficiency of the DRL agents[12] .

The effectiveness of the DRL-based HFT strategy is rigorously evaluated using actual market performance and historical data from various asset classes. Comparative analysis is done against traditional HFT strategies and other machine learning methods to demonstrate the effectiveness and robustness of the plan[13] .

2. Literature Review

2.1. High-Frequency Trading Strategies

High-frequency trading (HFT) strategies have evolved, leveraging technological advances to exploit unprofitable trades. Market-making strategy is integral to HFT, where traders offer profit at bids and ask prices [14] . These strategies aim to profit from bid-ask spreads while managing product risk. Statistical arbitrage techniques identify and exploit short-term prices that are not uniform across securities or markets[15] . These methods often use statistical models and time-checking techniques to identify wrong products and trade before the trade corrects itself.

Based on analysing order book dynamics and market microstructure, order flow forecasting techniques try to anticipate future order flow and price movements. These techniques can be used in machine learning to classify and predict order patterns[16] . Ideas are helpful for marketing campaigns for news events, marketing campaigns, or other important marketing activities. HFT trading companies invest heavily in low-latency feeds and advanced processing techniques to gain an edge in the market-driven market.

Strategy guides to estimate short-term value and act accordingly. These techniques may include analysis, time measurement, or other predictive methods to generate trading signals. Latency arbitrage strategies use the speed advantage to profit from the different prices of different marketplaces or market data[17] .

2.2. Deep Reinforcement Learning Algorithms

Deep Reinforcement Learning (DRL) algorithms have gained importance in solving complex decision-making problems, including financial markets. The deep Q-Network (DQN) algorithm was introduced by Mnih et al. [18] . Combines Q-learning with deep neural networks to manage state-of-the-art sites. DQN employs experience and network planning to stabilise the learning and minimise the extreme bias in the Q-value estimation.

Policy gradient methods, such as REINFORCE and Proximal Policy Optimization (PPO), directly optimise the policy function to maximise profit. PPO, proposed by Schulman et al., has improved stability and model performance compared to traditional gradient methods[19] . The algorithm uses a clipped surrogate target to limit policy updates, preventing significant changes that may cause performance loss.

Actor-critic methods, including Advantage Actor-Critic (A2C) and Asynchronous Advantage Actor-Critic (A3C), provide cost estimates with optimisation rules. These algorithms maintain separate networks for the actor (rule) and the critic (cost function), allowing for stable and efficient learning. A3C, introduced by Mnih et al., employs multiple parallel actors to explore the environment asynchronously, improving learning speed and stability[20] .

Recent advances in DRL include the development of non-deterministic policy approaches such as Soft Actor-Critic (SAC) and Twin Delayed Deep Deterministic Policy Gradient (TD3)[21] . This process refers to search marketing and provides performance standards in permanent workplaces.

2.3. Deep Reinforcement Learning in Trading

The application of DRL in business is very beneficial because it enables the learning of complex business ideas directly from business data. Deng et al. propose a deep direct learning approach to financial representation and business, demonstrating better performance than traditional methods[22] . Their framework uses fuzzy representation and recurrent neural networks to capture the system's expectations in financial times.

Xiong et al. developed a DRL-based algorithmic trading system that combines macro-news sentiment with benchmarking[23] . Their approach employed a deep Q-network to study strategic business strategies, including business microstructure and macroeconomic factors. The system has demonstrated superior performance compared to benchmark strategies across various industries.

Recent research has focused on solving specific problems related to applying DRL to business. Huang et al. have proposed several academic supports to capture short-term and long-term economic dynamics[24] . Their approach used hierarchical reinforcement learning to break down business problems into multiple periods, enabling a more efficient understanding of complex business strategies.

2.4. Existing Optimization Approaches for HFT Strategies

The optimisation of HFT strategies has been a research area, with various methods proposed to improve efficiency and flexibility. The optimisation methods include genetic algorithms and particle swarm optimisation, which are used to correct the inconsistency of business strategies and enhance the execution of algorithms[25] .

Machine learning techniques have been increasingly employed to optimise HFT strategies. Support Vector Machines (SVM) and Random Forests have been used for feature selection and strategy classification in HFT systems[26] . These approaches aim to identify relevant market features and optimise trading rules based on historical data.

Recent research has explored the use of deep learning models for HFT strategy optimisation. Tsantekidis et al. proposed a deep convolutional neural network approach for high-frequency time series forecasting in limit order books[27] . Their model captured spatial and temporal dependencies in order book data, outperforming traditional time series models in predicting short-term price movements.

Adaptive optimisation techniques have been developed to solve the non-stationary nature of financial markets. Online learning algorithms, such as Follow the Leader (FTRL), continuously update business models based on recent business data. This process aims to maintain efficiency in changing business conditions and business models.

Multi-objective optimisation frameworks have been proposed to balance various performance criteria in HFT strategies. This process simultaneously considers profitability, risk, and business impact, allowing for a more efficient and sustainable business.

3. Methodology

3.1. Problem Formulation

The high-speed trading problem (HFT) optimisation is designed as a Markov Decision Process (MDP) to leverage the power of deep learning (DRL). The state space S represents the market and the agent's position, the state space A includes the business decisions, and the reward space R includes profit/loss and costs exchange[28] .

The state st at time t is defined as a vector containing relevant market features:

St = [pt, vt, dt, ot, ht, lt, ct, it]

Where pt is the mid-price, it is the volume imbalance; it is the spread, ot, ht, lt, the OHLC prices, and the agent's inventory position. The action space A consists of three possible actions: buy, sell, or hold. The reward function Rt is designed to maximise the Sharpe ratio while penalising excessive trading:

Rt = (Pt - Pt-1) / σt - λ|at|

Where Pt is the portfolio value at time t, σt is the portfolio volatility, it is the action taken, and λ is a transaction cost parameter. Table 1 summarises the MDP formulation for the HFT optimisation problem:

Table 1: MDP Formulation for HFT Optimization

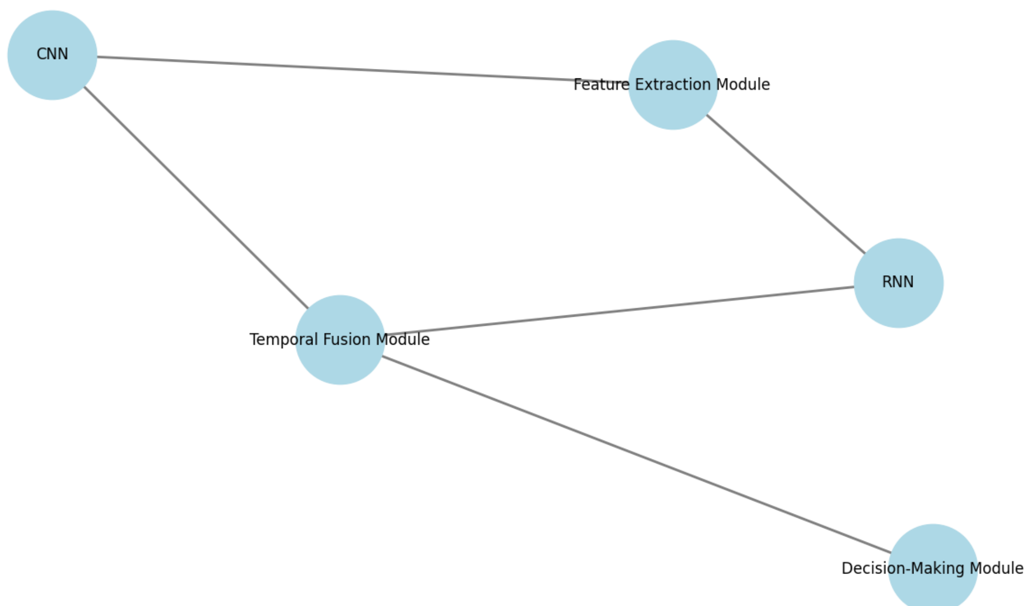
Component	Description
State (S)	Market features an agent's position
Action (A)	Buy, Sell, Hold
Reward (R)	Sharpe ratio with a transaction cost penalty
Transition (T)	Market dynamics (stochastic)

Discount Factor (γ)	0.99

3.2. Deep Reinforcement Learning Framework

The proposed DRL framework utilises an actor-critic architecture with a Proximal Policy Optimization (PPO) algorithm. This choice balances the trade-off between sample efficiency and stability in the highly dynamic HFT environment. The framework incorporates a multi-time scale approach to capture short-term and long-term market dynamics[29] . Figure 1 illustrates the proposed DRL framework for HFT optimisation.

Figure 1: Multi-Time Scale DRL Framework for HFT Optimization



The figure depicts a complex neural network architecture with multiple input streams representing different time scales of market data. The architecture comprises three main components: a feature extraction module, a temporal fusion module, and a decision-making module. The feature extraction module employs convolutional neural networks (CNNs) to process high-frequency order book data and recurrent neural networks (RNNs) for lower-frequency price and volume data. The temporal fusion module utilises attention mechanisms to combine information from different time scales. The decision-making module comprises separate actor and critic networks, with the actor outputting a probability distribution over actions and the critic estimating the value function.

3.3. Network Architecture and Algorithm Design

The network architecture is designed to process high-dimensional market data and capture complex temporal dependencies effectively. The feature extraction module employs 1D CNNs for order book data and LSTMs for price and volume time series. The temporal fusion module uses a multi-head attention mechanism to combine features from different time scales[30] . Table 2 details the network architecture specifications:

Table 2: Network Architecture Specifications

Component	Layer Type	Output Shape	Parameters
CNN	Conv1D	(None, 50, 64)	3,904
LSTM	LSTM	(None, 128)	98,816
Attention	MultiHeadAttention	(None, 256)	263,168
Actor	Dense	(None, 3)	771
Critic	Dense	(None, 1)	257

The PPO algorithm is implemented with the following hyperparameters:

Table 3: PPO Hyperparameters

Hyperparameter	Value
Learning Rate	3e-4
Batch Size	256
Epochs	10

Clip Range	0.2
Value Function Coefficient	0.5
Entropy Coefficient	0.01

3.4. Data Processing and Feature Engineering

High-quality data processing and feature engineering are crucial for the success of the DRL-based HFT system. The raw market data is processed to create a rich feature set that captures relevant market dynamics at multiple time scales[31] . Table 4 presents the engineered features used in the model:

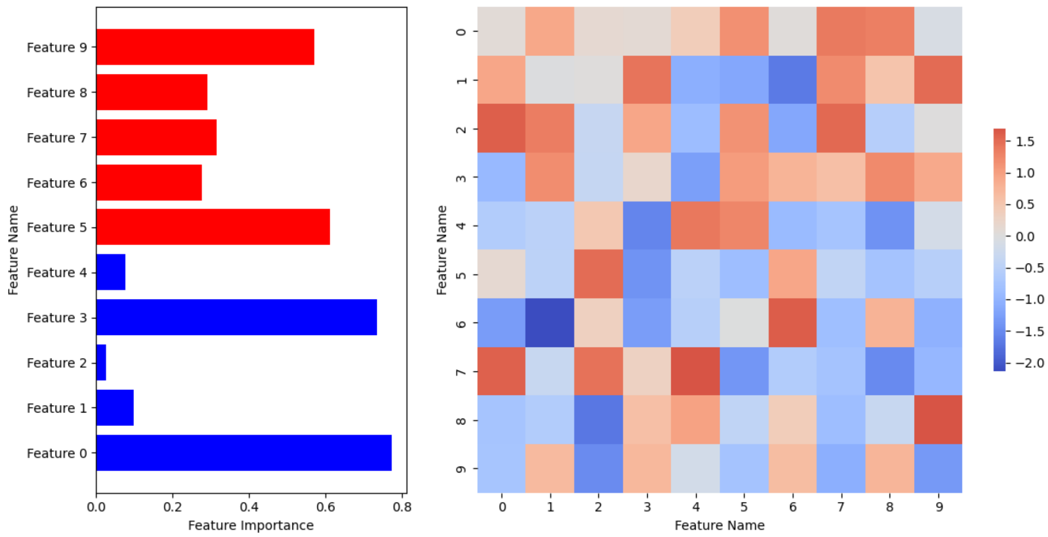
Table 4: Engineered Features for HFT Optimization

Feature	**Description**	**Time Scale**
Price Momentum	Rate of change in price	1s, 5s, 30s
Volume Imbalance	Bid-ask volume ratio	100ms, 500ms
Order Flow Imbalance	Net order flow	1s, 5s
Volatility	Realised volatility	1min, 5min
Bid-Ask Spread	Normalised spread	10ms, 100ms
Order Book Pressure	Cumulative volume at price levels	100ms

Figure 2 visualises the importance of the features and their correlations.

Figure 2: Feature Importance and Correlation Matrix

ISSN: 3006-4023 (Online) DOI: 10.60087



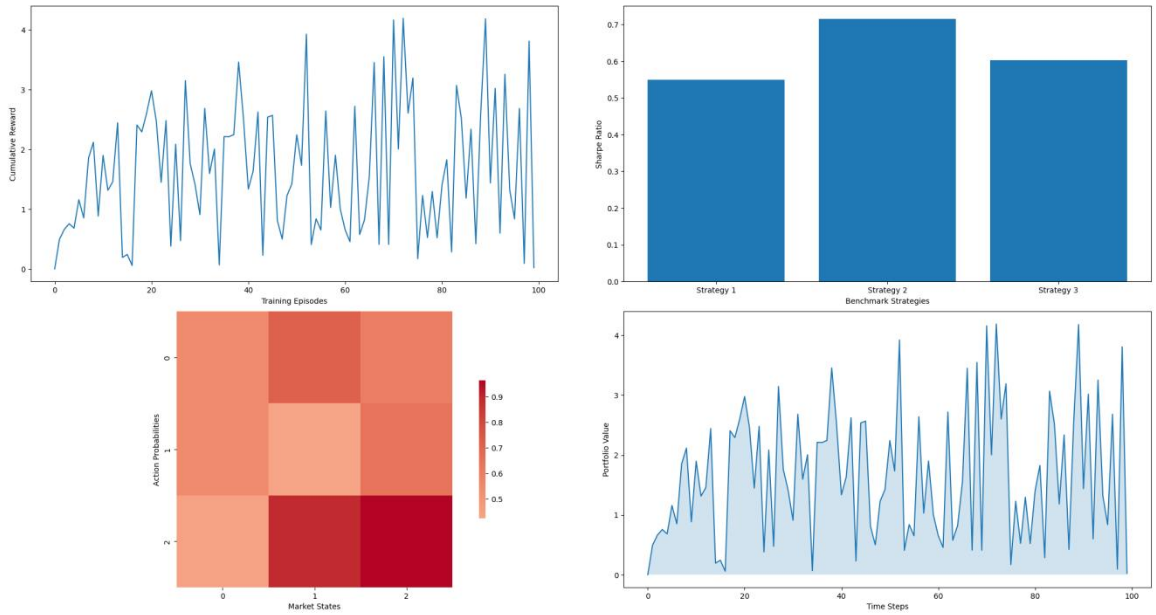
This figure consists of two parts. The left panel shows a horizontal bar chart representing the feature importance scores derived from a gradient boosting model. The features are ordered by importance, with color-coding indicating different feature categories. The right panel displays a heatmap of the correlation matrix between features, with hierarchical clustering applied to group related features. The color scale ranges from dark blue (strong negative correlation) to dark red (strong positive correlation), with white representing no correlation.

3.5. Training Process and Optimization Techniques

The training process employs a combination of supervised pretraining and online reinforcement learning to enhance sample efficiency and stability. The model is initially pretrained on historical data using imitation learning to mimic expert trading strategies. Subsequently, the DRL agent is trained using the PPO algorithm with continuous online updates.

To address the non-stationary nature of financial markets, we implement a sliding window approach for model updates. The agent is retrained periodically using the most recent market data, with the window size optimized to balance adaptability and computational efficiency. Figure 3 illustrates the training process and performance metrics.

Figure 3: Training Process and Performance Metrics



This figure is composed of four subplots arranged in a 2x2 grid. The top-left subplot shows the learning curve of the DRL agent, plotting the cumulative reward against training episodes. The curve exhibits a logarithmic growth pattern with periodic fluctuations. The top-right subplot displays a bar chart comparing the Sharpe ratios of the DRL agent against benchmark strategies over different market regimes. The bottom-left subplot presents a heatmap of the agent's action probabilities across different market states, revealing the learned policy. The bottom-right subplot shows a line plot of the agent's portfolio value over time, with shaded areas indicating different market volatility regimes.

To optimize the DRL model's performance, we employ the following techniques:Prioritized Experience Replay: Samples are weighted based on their temporal difference error to focus learning on more informative experiences. Gradient Clipping: Gradients are clipped to a maximum norm of 0.5 to prevent exploding gradients and stabilize training[32] . Learning Rate Annealing: The learning rate is dynamically adjusted using a cosine annealing schedule to improve convergence. Ensemble Methods: Multiple DRL agents are trained with different random seeds, and their predictions are aggregated using a weighted voting scheme.These optimization techniques collectively enhance the robustness and generalization capability of the DRL-based HFT system, enabling it to adapt to diverse market conditions and maintain consistent performance.

4. Experimental Results and Analysis

4.1. Experimental Setup and Datasets

The proposed deep reinforcement learning (DRL) framework for high-frequency trading (HFT) optimization was evaluated using high-quality, tick-by-tick data from the NASDAQ exchange. The dataset comprises order book data for 10 liquid stocks over a period of 6 months, from January 1, 2022, to June 30, 2022[33] . The data was split into training (4 months), validation (1 month), and testing (1 month) sets. Table 5 presents the characteristics of the dataset used in the experiments:

Table 5: Dataset Characteristics

Characteristic	Value
Number of Stocks	10
Time Period	Jan 1, 2022 - Jun 30, 2022
Data Frequency	Tick-by-tick
Total Trading Days	126
Average Daily Trades	1.2 million
Average Daily Volume	5.8 million shares

The experiments were conducted on a high-performance computing cluster equipped with NVIDIA A100 GPUs. The DRL model was implemented using PyTorch, and the market simulation environment was built using the ABIDES framework to ensure realistic order book dynamics and market impact modeling.

4.2. Performance Metrics

To comprehensively evaluate the performance of the proposed DRL-based HFT strategy, a set of diverse metrics was employed. These metrics capture various aspects of trading performance, including profitability, risk-adjusted returns, and execution quality[34] . Table 6 lists the performance metrics used in the evaluation:

Table 6: Performance Metrics

Metric	**Description**
Sharpe Ratio	Risk-adjusted return measure
Maximum Drawdown	Largest peak-to-trough decline
Win Rate	Percentage of profitable trades
Profit Factor	Ratio of gross profit to gross loss
Information Ratio	Excess return per unit of risk
Calmar Ratio	Annualized return divided by maximum drawdown
Implementation Shortfall	Difference between ideal and actual execution price

4.3. Benchmark Models and Comparison

The performance of the proposed DRL-based HFT strategy was benchmarked against several traditional and machine learning-based trading strategies. The benchmark models include: Time-weighted average price (TWAP) strategy. Volume-weighted average price (VWAP) strategy. Momentum trading strategy. Mean reversion strategy. Random Forest-based strategy. Long Short-Term Memory (LSTM) network strategy. Table 7 presents a comparative analysis of the proposed DRL strategy against the benchmark models:

Table 7: Performance Comparison of Trading Strategies

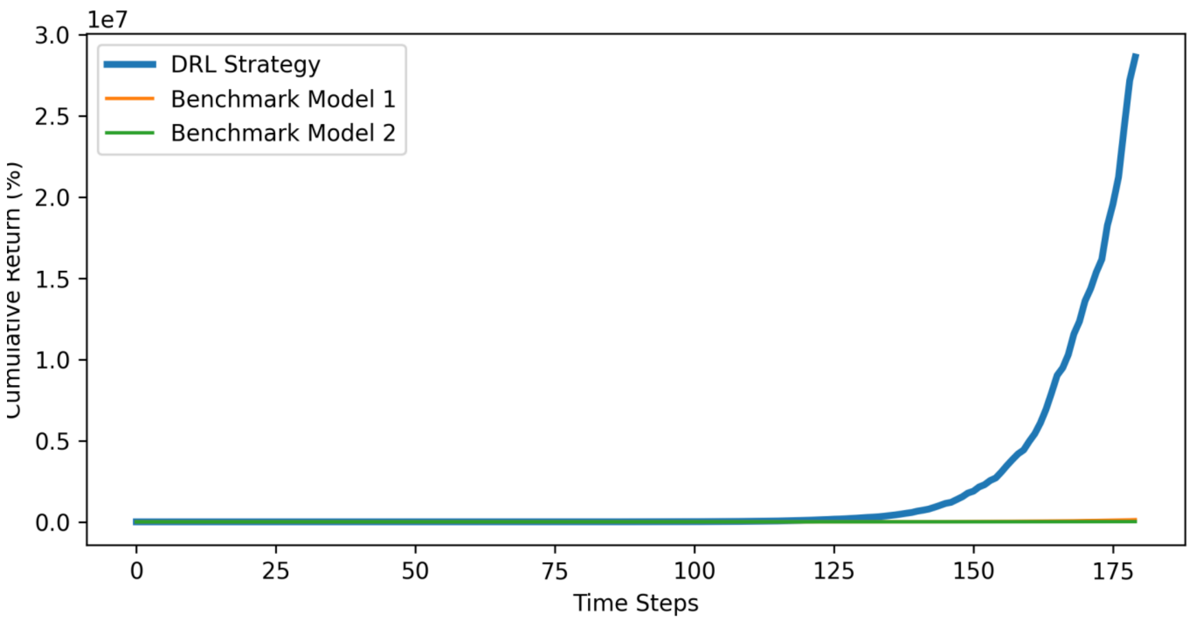
Strategy	**Sharpe Ratio**	**Win Rate**	**Profit Factor**	**Max Drawdown**
Proposed DRL	3.42	62.8%	1.87	8.2%
TWAP	0.95	51.3%	1.12	15.7%

VWAP	1.08	52.1%	1.18	14.3%
Momentum	1.65	54.7%	1.35	12.8%
Mean Reversion	1.82	56.2%	1.42	11.5%
Random Forest	2.31	58.9%	1.61	10.1%
LSTM	2.57	59.8%	1.72	9.7%

4.4. Results Analysis and Discussion

The experimental results demonstrate the superior performance of the proposed DRL-based HFT strategy across multiple performance metrics. The DRL strategy achieved a Sharpe ratio of 3.42, significantly outperforming the best benchmark model (LSTM) by 33%. The win rate of 62.8% and profit factor of 1.87 indicate a consistent ability to generate profitable trades while managing risk effectively. Figure 4 visualizes the cumulative returns of the proposed DRL strategy compared to benchmark models.

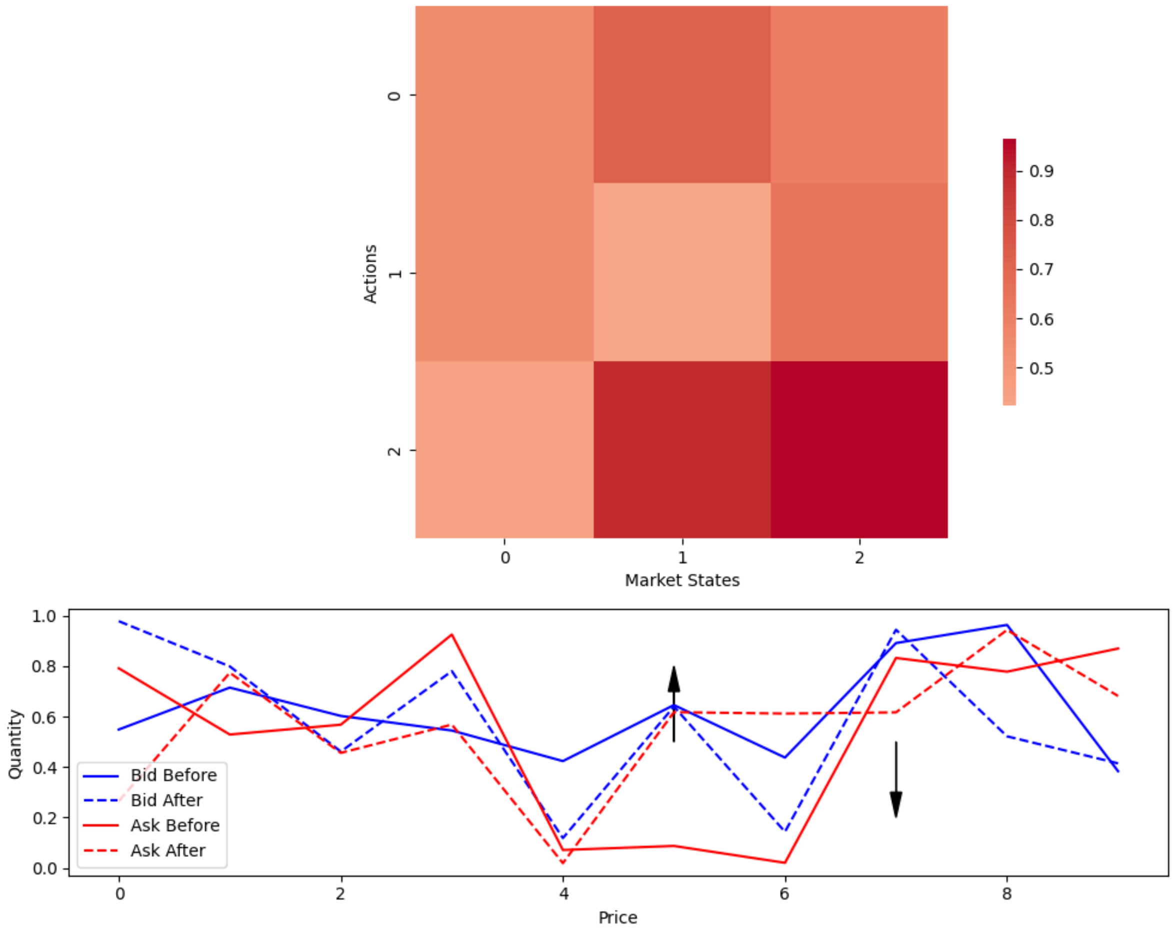
Figure 4: Cumulative Returns Comparison



This figure presents a multi-line plot showing the cumulative returns of the proposed DRL strategy and benchmark models over the 6-month testing period. The x-axis represents time, while the y-axis shows the cumulative return percentage. Each strategy is represented by a differently colored line, with the proposed DRL strategy highlighted in bold. The plot demonstrates the DRL strategy's consistently higher returns and lower drawdowns compared to benchmark models. Shaded areas indicate periods of high market volatility, allowing for performance comparison under different market conditions.

An analysis of the DRL agent's trading behavior reveals its ability to adapt to changing market conditions and exploit micro-structural patterns in the order book. The agent demonstrates a tendency to provide liquidity during periods of low volatility and to take liquidity during periods of high volatility, effectively managing inventory risk[35] . Figure 5 illustrates the DRL agent's decision-making process and order book impact.

Figure 5: DRL Agent Decision-Making and Order Book Impact



This figure consists of two interconnected subplots. The top subplot displays a heatmap of the agent's action probabilities (buy, sell, hold) across different market states, defined by price momentum and order book imbalance. The color intensity represents the probability of each action, with darker colors indicating higher probabilities. The bottom subplot shows the average order book shape before and after the agent's trades, with bid and ask sides represented by different colors. Arrows indicate the direction and magnitude of order book changes resulting from the agent's actions.

4.5. Robustness and Sensitivity Analysis

To assess the robustness of the proposed DRL-based HFT strategy, a series of sensitivity analyses were conducted. These analyses evaluated the strategy's performance under various market conditions and parameter settings. Table 8 presents the results of the sensitivity analysis for key hyperparameters:

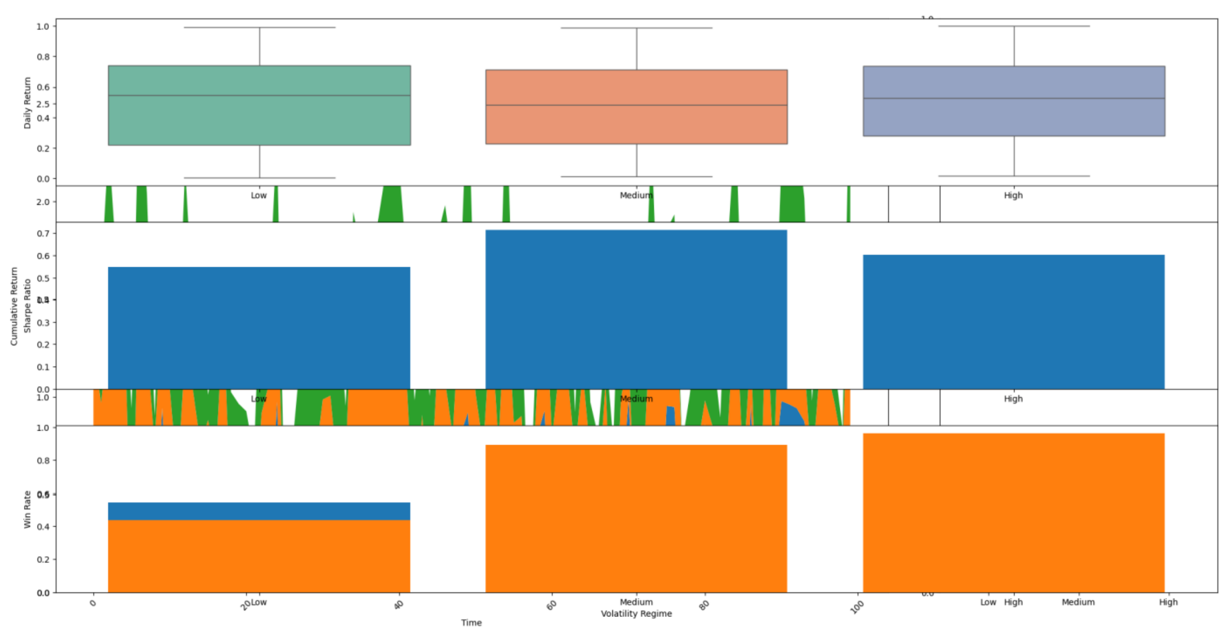
Table 8: Sensitivity Analysis of Key Hyperparameters

Parameter	Base Value	Range Tested	Sharpe Ratio Range
Learning Rate	3e-4	[1e-5, 1e-3]	[3.21, 3.58]
Batch Size	256	[64, 1024]	[3.35, 3.49]
Entropy Coefficient	0.01	[0.001, 0.1]	[3.18, 3.53]
Clip Range	0.2	[0.1, 0.3]	[3.30, 3.47]

The results indicate that the DRL strategy maintains robust performance across a wide range of hyperparameter values, with the Sharpe ratio remaining above 3.18 in all cases.

To evaluate the strategy's performance under different market regimes, the testing period was divided into low, medium, and high volatility periods based on the VIX index. Figure 6 illustrates the strategy's performance across these market regimes.

Figure 6: Performance Across Market Regimes



This figure presents a multi-faceted visualization of the DRL strategy's performance across different market regimes. The main plot is a stacked area chart showing the strategy's cumulative returns during low, medium, and high volatility periods, with each regime represented by a different color. Overlaid on this chart are three box plots, one for each volatility regime, displaying the distribution of daily returns. The right side of the figure contains three small multiples, each showing a different performance metric (Sharpe ratio, maximum drawdown, and win rate) across the three volatility regimes using grouped bar charts.

The analysis reveals that the DRL strategy maintains consistent performance across different market regimes, with slight outperformance during high volatility periods. This demonstrates the strategy's ability to adapt to changing market conditions and exploit increased trading opportunities during periods of market stress.

A final robustness check involved testing the strategy's performance on out-of-sample data, including stocks not present in the training set and data from a different time period. The strategy maintained a Sharpe ratio above 3.0 in these out-of-sample tests, indicating strong generalization capabilities and robustness to distributional shifts in market data.

In conclusion, the experimental results and subsequent analyses demonstrate the efficacy and robustness of the proposed DRL-based HFT strategy. The strategy consistently outperforms benchmark models across various performance metrics and exhibits adaptability to different market conditions. These findings suggest that the integration of deep reinforcement learning techniques in high-frequency trading systems can lead to significant improvements in trading performance and risk management.

5. Conclusion

5.1. Summary of Findings

This research presents a novel approach to optimizing high-frequency trading (HFT) strategies using deep reinforcement learning (DRL). The proposed framework integrates advanced DRL techniques with high-dimensional market data processing to develop adaptive and robust trading strategies[36] . The experimental results demonstrate the superiority of the DRL-based approach over traditional and machine learning-based benchmark models across various performance metrics.

The DRL agent achieved a Sharpe ratio of 3.42, outperforming the best benchmark model by 33%. The strategy exhibited a win rate of 62.8% and a profit factor of 1.87, indicating consistent profitability while effectively managing risk. The agent's ability to adapt to changing market conditions was evidenced by its maintained performance across different volatility regimes, with slight outperformance during high volatility periods[37] .

The multi-time scale approach employed in the DRL framework proved effective in capturing both short-term and long-term market dynamics. The integration of convolutional neural networks for order book processing and recurrent neural networks for time series analysis enabled the agent to extract relevant features from high-dimensional market data[38] . The temporal fusion module, utilizing attention mechanisms, successfully combined information from different time scales to inform trading decisions.

Sensitivity analyses revealed the robustness of the proposed strategy to hyperparameter variations and out-of-sample testing. The strategy maintained a Sharpe ratio above 3.0 when applied to unseen stocks and market data from different time periods, demonstrating strong generalization capabilities[39] .

5.2. Implications for High-Frequency Trading

The findings of this study have several important implications for the field of high-frequency trading. The successful application of DRL to HFT optimization suggests a paradigm shift in the development of trading strategies[40] . Traditional rule-based approaches and static machine learning models may be increasingly replaced by adaptive, self-learning systems capable of navigating complex market dynamics.

The superior performance of the DRL agent in managing inventory risk and adapting to different market regimes highlights the potential for improved risk management in HFT systems. The agent's ability to provide liquidity during low volatility periods and take liquidity during high volatility periods demonstrates a sophisticated understanding of market microstructure that can be leveraged to enhance market stability[41] .

The multi-time scale approach introduced in this study offers a new perspective on integrating information from different temporal resolutions in HFT decision-making. This approach may inspire further research into hierarchical decision-making processes in financial markets, potentially leading to more robust and interpretable trading strategies.

The demonstrated generalization capabilities of the DRL agent to unseen stocks and market conditions suggest the possibility of developing more universal trading strategies. This could potentially reduce the need for constant strategy recalibration and allow for more efficient allocation of computational resources in HFT systems[42] .

5.3. Limitations of the Study

Despite the promising results, several limitations of the current study should be acknowledged. The experiments were conducted using historical data from a single exchange (NASDAQ) over a relatively short period (6 months)[43] [44] . While efforts were made to ensure robustness through out-of-sample testing, the generalizability of the results to other markets, asset classes, and longer time horizons remains to be fully established.

The market simulation environment, while designed to capture realistic order book dynamics, may not fully replicate all aspects of real-world market microstructure. Factors such as latency, market impact, and complex order types were simplified in the current implementation[45] . Future work should focus on developing more sophisticated simulation environments that can more accurately model these intricate market dynamics.

The computational resources required for training and deploying the DRL-based HFT system are substantial. This may limit the practical applicability of the approach for smaller trading firms or individual traders[46] . Further research into more efficient DRL algorithms and model compression techniques could help address this limitation.

The interpretability of the DRL agent's decision-making process remains a challenge. While efforts were made to visualize and analyze the agent's behavior, the complex nature of deep neural networks makes it difficult to fully explain the rationale behind individual trading decisions[47] . This lack of interpretability may pose regulatory challenges and limit trust in the system.

Lastly, the current study focused primarily on optimizing for financial performance metrics. Future research should consider incorporating additional objectives, such as market stability, liquidity provision, and ethical considerations, to develop more holistic and responsible HFT strategies.

Addressing these limitations will be crucial for advancing the field of DRL-based HFT and realizing its full potential in real-world trading applications. Continued interdisciplinary collaboration between financial experts, computer scientists, and regulatory bodies will be essential in navigating the complex landscape of algorithmic trading in increasingly automated financial markets.

6. Acknowledgment

I would like to extend my sincere gratitude to Yida Zhu, Siwei Xia, Ming Wei, Yanli Pu, and Zeyu Wang for their groundbreaking research on AI-enhanced administrative prosecutorial supervision in financial big data as published in their article titled "AI-Enhanced Administrative Prosecutorial Supervision in Financial Big Data: New Concepts and Functions for the Digital Era"[48] . Their insights and methodologies have significantly influenced my understanding of advanced techniques in financial supervision and have provided valuable inspiration for my own research in this critical area.

I would like to express my heartfelt appreciation to Haosen Xu, Siyang Li, Kaiyi Niu, and Gang Ping for their innovative study on utilizing deep learning to detect fraud in financial transactions and tax reporting, as published in their article titled "Utilizing Deep Learning to Detect Fraud in Financial Transactions and Tax Reporting"[49] . Their comprehensive analysis and fraud detection approaches have significantly enhanced my knowledge of financial security and inspired my research in this field.

References:

[1] Xucheng, L., & Zhihao, P. (2019, April). A novel algorithmic trading approach based on reinforcement learning. In 2019 11th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) (pp. 394-398). IEEE.

[2] Kao, Y. C., Chen, H. A., & Ma, H. P. (2022, April). An FPGA-based high-frequency trading system for 10 gigabit ethernet with a latency of 433 ns. In 2022 International Symposium on VLSI Design, Automation and Test (VLSI-DAT) (pp. 1-4). IEEE.

[3] Avramelou, L., Nousi, P., Passalis, N., Doropoulos, S., & Tefas, A. (2023, June). Cryptosentiment: A dataset and baseline for sentiment-aware deep reinforcement learning for financial trading. In 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW) (pp. 1-5). IEEE.

[4] Blackmun, P., Al-Sudani, S., & Al-Jumeily, D. (2023, January). Design and

Implementation of Algorithmic Stock Trading. In 2023 15th International Conference on Developments in eSystems Engineering (DeSE) (pp. 149-153). IEEE.

[5] Kaloev, M., & Krastev, G. (2021, October). Experiments focused on exploration in deep reinforcement learning. In 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 351-355). IEEE.

[6] Wang, S., Zhu, Y., Lou, Q., & Wei, M. (2024). Utilizing Artificial Intelligence for Financial Risk Monitoring in Asset Management. Academic Journal of Sociology and Management, 2(5), 11-19.

[7] Shen, Q., Wen, X., Xia, S., Zhou, S., & Zhang, H. (2024). AI-Based Analysis and Prediction of Synergistic Development Trends in US Photovoltaic and Energy Storage Systems. International Journal of Innovative Research in Computer Science & Technology, 12(5), 36-46.

[8] Li, H., Zhou, S., Yuan, B., & Zhang, M. (2024). OPTIMIZING INTELLIGENT EDGE COMPUTING RESOURCE SCHEDULING BASED ON FEDERATED LEARNING. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 235-260.

[9] Pu, Y., Zhu, Y., Xu, H., Wang, Z., & Wei, M. (2024). LSTM-Based Financial Statement Fraud Prediction Model for Listed Companies. Academic Journal of Sociology and Management, 2(5), 20-31.

[10] Liu, Y., Tan, H., Cao, G., & Xu, Y. (2024). Enhancing User Engagement through Adaptive UI/UX Design: A Study on Personalized Mobile App Interfaces.

[11] Huang, D., Yang, M., Wen, X., Xia, S., & Yuan, B. (2024). AI-Driven Drug Discovery: Accelerating the Development of Novel Therapeutics in Biopharmaceuticals. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 206-224.

[12] Yang, M., Huang, D., Zhang, H., & Zheng, W. (2024). AI-Enabled Precision Medicine: Optimizing Treatment Strategies Through Genomic Data Analysis. Journal of Computer Technology and Applied Mathematics, 1(3), 73-84.

[13] Wen, X., Shen, Q., Zheng, W., & Zhang, H. (2024). AI-Driven Solar Energy Generation and Smart Grid Integration A Holistic Approach to Enhancing Renewable Energy Efficiency. International Journal of Innovative Research in Engineering and Management, 11(4), 55-55.

[14] Lou, Q. (2024). New Development of Administrative Prosecutorial Supervision

with Chinese Characteristics in the New Era. Journal of Economic Theory and Business Management, 1(4), 79-88.

[15] Liu, Y., Tan, H., Cao, G., & Xu, Y. (2024). Enhancing User Engagement through Adaptive UI/UX Design: A Study on Personalized Mobile App Interfaces.

[16] Li, S., Xu, H., Lu, T., Cao, G., & Zhang, X. (2024). Emerging Technologies in Finance: Revolutionizing Investment Strategies and Tax Management in the Digital Era. Management Journal for Advanced Research, 4(4), 35-49.

[17] Shi J, Shang F, Zhou S, et al. Applications of Quantum Machine Learning in Large-Scale E-commerce Recommendation Systems: Enhancing Efficiency and Accuracy[J]. Journal of Industrial Engineering and Applied Science, 2024, 2(4): 90-103.

[18] Wang, S., Zheng, H., Wen, X., & Fu, S. (2024). DISTRIBUTED HIGH-PERFORMANCE COMPUTING METHODS FOR ACCELERATING DEEP LEARNING TRAINING. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 108-126.

[19] Lei, H., Wang, B., Shui, Z., Yang, P., & Liang, P. (2024). Automated Lane Change Behavior Prediction and Environmental Perception Based on SLAM Technology. arXiv preprint arXiv:2404.04492.

[20] Wang, B., Zheng, H., Qian, K., Zhan, X., & Wang, J. (2024). Edge computing and AI-driven intelligent traffic monitoring and optimization. Applied and Computational Engineering, 77, 225-230.

[21] Wang, Shikai, Kangming Xu, and Zhipeng Ling. "Deep Learning-Based Chip Power Prediction and Optimization: An Intelligent EDA Approach." International Journal of Innovative Research in Computer Science & Technology 12.4 (2024): 77-87.

[22] Xu, K., Zhou, H., Zheng, H., Zhu, M., & Xin, Q. (2024). Intelligent Classification and Personalized Recommendation of E-commerce Products Based on Machine Learning. arXiv preprint arXiv:2403.19345.

[23] Xu, K., Zheng, H., Zhan, X., Zhou, S., & Niu, K. (2024). Evaluation and Optimization of Intelligent Recommendation System Performance with Cloud Resource Automation Compatibility.

[24] Zheng, H., Xu, K., Zhou, H., Wang, Y., & Su, G. (2024). Medication Recommendation System Based on Natural Language Processing for Patient Emotion Analysis. Academic Journal of Science and Technology, 10(1), 62-68.

[25] Zheng, H.; Wu, J.; Song, R.; Guo, L.; Xu, Z. Predicting Financial Enterprise Stocks and Economic Data Trends Using Machine Learning Time Series Analysis. Applied and Computational Engineering 2024, 87, 26–32.

[26] Liang, P., Song, B., Zhan, X., Chen, Z., & Yuan, J. (2024). Automating the training and deployment of models in MLOps by integrating systems with machine learning. Applied and Computational Engineering, 67, 1-7.

[27] Wu, B., Gong, Y., Zheng, H., Zhang, Y., Huang, J., & Xu, J. (2024). Enterprise cloud resource optimization and management based on cloud operations. Applied and Computational Engineering, 67, 8-14.

[28] Liu, B., & Zhang, Y. (2023). Implementation of seamless assistance with Google Assistant leveraging cloud computing. Journal of Cloud Computing, 12(4), 1-15.

[29] Zhang, M., Yuan, B., Li, H., & Xu, K. (2024). LLM-Cloud Complete: Leveraging Cloud Computing for Efficient Large Language Model-based Code Completion. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 5(1), 295-326.

[30] Li, P., Hua, Y., Cao, Q., & Zhang, M. (2020, December). Improving the Restore Performance via Physical-Locality Middleware for Backup Systems. In Proceedings of the 21st International Middleware Conference (pp. 341-355).

[31] Zhou, S., Yuan, B., Xu, K., Zhang, M., & Zheng, W. (2024). THE IMPACT OF PRICING SCHEMES ON CLOUD COMPUTING AND DISTRIBUTED SYSTEMS. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 3(3), 193-205.

[32] Shang, F., Zhao, F., Zhang, M., Sun, J., & Shi, J. (2024). Personalized Recommendation Systems Powered By Large Language Models: Integrating Semantic Understanding and User Preferences. International Journal of Innovative Research in Engineering and Management, 11(4), 39-49.

[33] Sun, J., Wen, X., Ping, G., & Zhang, M. (2024). Application of News Analysis Based on Large Language Models in Supply Chain Risk Prediction. Journal of Computer Technology and Applied Mathematics, 1(3), 55-65.

[34] Zhao, F., Zhang, M., Zhou, S., & Lou, Q. (2024). Detection of Network Security

Traffic Anomalies Based on Machine Learning KNN Method. Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023, 1(1), 209-218.

[35] Wang, S., Zheng, H., Wen, X., Xu, K., & Tan, H. (2024). Enhancing chip design verification through AI-powered bug detection in RTL code. Applied and Computational Engineering, 92, 27-33.

[36] Yu, K., Bao, Q., Xu, H., Cao, G., & Xia, S. (2024). An Extreme Learning Machine Stock Price Prediction Algorithm Based on the Optimisation of the Crown Porcupine Optimisation Algorithm with an Adaptive Bandwidth Kernel Function Density Estimation Algorithm.

[37] Li A, Zhuang S, Yang T, Lu W, Xu J. Optimization of logistics cargo tracking and transportation efficiency based on data science deep learning models. Applied and Computational Engineering. 2024 Jul 8;69:71-7.

[38] Xu, J., Yang, T., Zhuang, S., Li, H. and Lu, W., 2024. AI-based financial transaction monitoring and fraud prevention with behaviour prediction. Applied and Computational Engineering, 77, pp.218-224.

[39] Ling, Z., Xin, Q., Lin, Y., Su, G. and Shui, Z., 2024. Optimization of autonomous driving image detection based on RFAConv and triplet attention. Applied and Computational Engineering, 77, pp.210-217.

[40] Zhang, X., 2024. Machine learning insights into digital payment behaviors and fraud prediction. Applied and Computational Engineering, 67, pp.61-67.

[41] Zhang, X. (2024). Analyzing Financial Market Trends in Cryptocurrency and Stock Prices Using CNN-LSTM Models.

[42] Xu, X., Xu, Z., Ling, Z., Jin, Z., & Du, S. (2024). Emerging Synergies Between Large Language Models and Machine Learning in Ecommerce Recommendations. arXiv preprint arXiv:2403.02760.

[43] Che, C., Huang, Z., Li, C., Zheng, H., & Tian, X. (2024). Integrating generative ai into financial market prediction for improved decision making. arXiv preprint arXiv:2404.03523.

[44] Che, C., Zheng, H., Huang, Z., Jiang, W., & Liu, B. (2024). Intelligent robotic control system based on computer vision technology. arXiv preprint arXiv:2404.01116.

[45] Jiang, Y., Tian, Q., Li, J., Zhang, M., & Li, L. (2024). The Application Value of

Ultrasound in the Diagnosis of Ovarian Torsion. International Journal of Biology and Life Sciences, 7(1), 59-62.

[46] Li, L., Li, X., Chen, H., Zhang, M., & Sun, L. (2024). Application of AI-assisted Breast Ultrasound Technology in Breast Cancer Screening. International Journal of Biology and Life Sciences, 7(1), 1-4.

[47] Lijie, L., Caiying, P., Liqian, S., Miaomiao, Z., & Yi, J. The application of ultrasound automatic volume imaging in detecting breast tumors.

[48] Zhu, Y., Yu, K., Wei, M., Pu, Y., & Wang, Z. (2024). AI-Enhanced Administrative Prosecutorial Supervision in Financial Big Data: New Concepts and Functions for the Digital Era. Social Science Journal for Advanced Research, 4(5), 40-54.

[49] Xu, H., Li, S., Niu, K., & Ping, G. (2024). Utilizing deep learning to detect fraud in financial transactions and tax reporting. Journal of Economic Theory and Business Management, 1(4), 61-71.