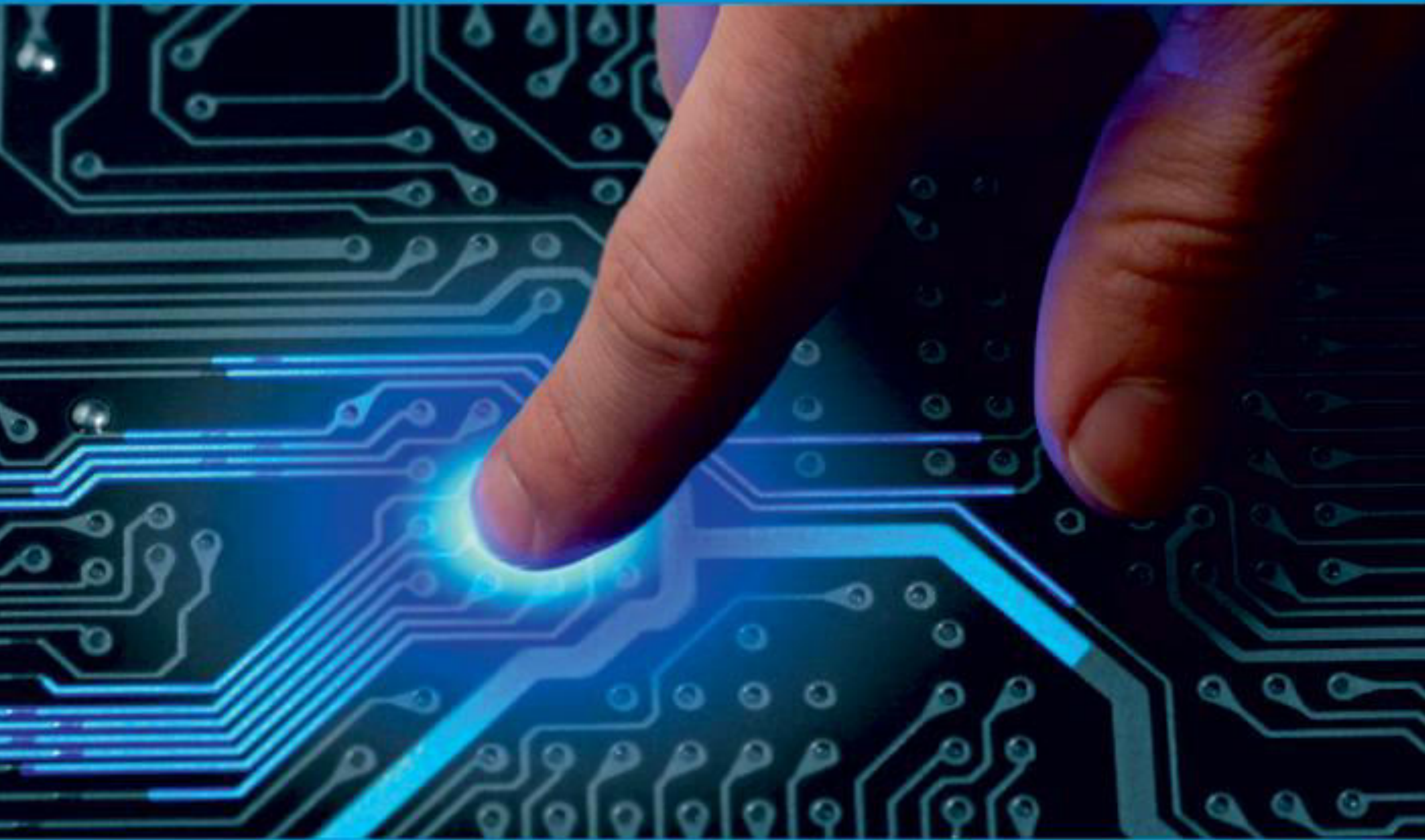
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 10, October 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.625

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Algorithmic Trading using Meta RL

Yuchitra V[1], Shwetha Manoj[2], Vidya P[3], Deepa R[4]

B Tech Student, Department of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India[1]

B Tech Student, Department of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India[2]

B Tech Student, Department of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India[3]

Assistant Professor, Department of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India[4]

ABSTRACT: Algorithmic trading can continue to dominate financial markets because this system allows for fast data-driven decisions. Traditional algorithms in machine learning for algorithmic trading lack the ability to adjust and continue to function well under changing conditions within volatile markets. Here is a new algorithmic trading system from Meta Reinforcement Learning (RL) that adapts rapidly to the evolution of market characteristics. Armed with the power of Model Agnostic Meta Learning (MAML) and Proximal Policy Optimization (PPO), this model can further learn generalizable financial behaviours and trading strategies. A comparison to state-of-the-art trading algorithms is made subsequently. Thus, we present how the Meta RL model beats standard trading algorithms in volatile and unseen market conditions by maximizing risk-adjusted returns and attracting less on losses.

KEYWORDS: Algorithmic Trading, Financial Markets, Meta Reinforcement Learning, MAML, PPO

I. INTRODUCTION

Algorithmic trading is based on models, which can analyse and respond to large volumes of data in real-time, but it relies on model interpretation; the financial markets tend to be highly volatile hence create difficulties for traditional machine learning models where abrupt market transitions or non-static conditions make adaptation challenging. Reinforcement learning has been identified as a potential method that could possibly be extremely powerful for trading; however, models used inside reinforcement learning generally require much longer than normal training periods, which renders them considerably less efficient for real-time optimization in fast-fluctuating markets.

This problem can be solved with a concrete solution by letting the model learn its ability to learn. That is, Meta RL utilizes a model trained in one particular market that can quickly adapt to another environment almost without any retraining needs. This paper provides an algorithmic trading framework that employs the algorithm Meta RL using MAML and PPO for better generalization and faster adaptability in the real-time financial markets. The intrinsic properties of stiffness do not allow traditional trading algorithms to adjust according to the fluctuations required by market operations. On the other hand, reinforcement learning is able to adjust according to new information but is significantly affected by the quality of data and requires an extended period for the retraining cycle. The objective of this study is to fill this gap by employing meta reinforcement learning to formulate a methodology for adaptive trading strategies that are dynamically responsive to fluctuating market conditions. Contributions This manuscript enhances several aspects of the current literature, as it meets the objectives outlined by establishing a novel framework for adaptive algorithmic trading, utilizing MAML-PPO. Running the model on highly structured professional trading, thereby allowing it access to the most precise and up-to-date stock market data. This methodology is undoubtedly significantly better than the very basic algorithms that are used to date, especially concerning volatility and instability, which are quintessence indicators of market cycles.

II. RELATED WORK

Previously invisible frontiers that hybridize meta-learning and reinforcement learning (RL) have made traders more similar to the conventional market in terms of being more adaptable in their trading strategies. Meta-learning combined with Seasonal-Trend Decomposition (STL) becomes one of the best models used in forecasting almost all commodities, it is learned from data simultaneously. Combining Multi-Agent Reinforcement Learning (MARL) with Proximal Policy

© 2024 IJIRCCE | Volume 12, Issue 10, October 2024| DOI: 10.15680/IJIRCCE.2024.1210030

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Optimization (PPO) is another successful method, which allows for portfolio optimization and also increases communication between the algorithm and the environment via acquisition, trade execution, and risk control [2].

Moreover, the fusion of Meta-RL (RL^2) and symbolic features has been implemented via dealing with high-frequency data, which means, mainly, emphasizing the features related to data and controlling unnecessary data to avoid duplication. Meta-learning fed to evolutionary strategies not only speeds up process but allows the utilization of more data for stock trading [4]. And another look at PPO, in the end-to-end execution system, among other things, shows advances in quality and efficiency of trade executions through the data obtained from Level 2 Limited Order Book (LOB) [5].

Model-Agnostic Meta-Learning (MAML) protocol is just another term fundamental to many of these methods and its chief property is the ease with which it allows systems to change themselves considerably without much data [6]. These devices jointly make it possible for trading systems to adapt to changes in the market instantly and improve their profitability levels.

III. PROPOSED ALGORITHM

The proposed algorithm leverages Meta Reinforcement Learning (Meta RL) to develop an adaptive trading strategy capable of responding dynamically to evolving market conditions. This algorithm combines Model-Agnostic Meta-Learning (MAML) with Proximal Policy Optimization (PPO) to improve decision-making in financial trading. Below is a step-by-step explanation of the algorithm:

3.1. Problem Definition

The trading problem is formulated as a Markov Decision Process (MDP), where:

- State (S): Represents the current market conditions, including features such as stock prices, trading volume, and technical indicators.
- Action (A): Represents the agent's possible actions, i.e., Buy, Sell, or Hold.
- Reward (R): The reward is defined based on the profit or loss resulting from the selected actions.
- Policy (π): The policy defines the probability of taking a particular action given the current state.

The objective is to maximize cumulative rewards over time while minimizing risk.

3.2. Meta Reinforcement Learning Framework

The proposed Meta RL algorithm consists of two main components: MAML and PPO.

3.3. Algorithmic Steps

Step 1: Data Preprocessing

- Input: Historical stock data (e.g., daily price, volume) from a Financial Data API.
- Preprocessing: Normalize the data, generate technical indicators (e.g., Moving Averages, RSI), and remove outliers to ensure data consistency.

Step 2: Meta-Learning Initialization (MAML)

- Meta-Training Tasks: A set of different market environments (e.g., bull, bear, sideways markets) is created from the historical data.
- Task Sampling: Randomly sample a batch of tasks from different market conditions.
- Inner Loop: For each task, the model is initialized with shared parameters and performs a few gradient updates to adapt to the specific task.
- Outer Loop: The model aggregates the gradient updates across all sampled tasks to update the meta-parameters. This ensures the model learns an initialization that can quickly adapt to new market environments with minimal retraining.

Step 3: Policy Optimization (PPO)

- Initial Policy: The policy network is initialized using the meta-learned parameters from MAML.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Experience Collection: The agent interacts with the trading environment (using Gymnasium) to collect experiences (state, action, reward, next state) over multiple episodes.
- Proximal Policy Optimization: The PPO algorithm is used to optimize the policy by:
    - Clipping the probability ratio between the old and new policies to limit how much the policy changes between updates.
    - Balancing exploration and exploitation to avoid large policy jumps, ensuring smooth learning.
    - Using a loss function that combines the policy gradient and value function for more stable updates.

Step 4: Trading Decision-Making

- Action Selection: At each time step, the agent observes the current market state and selects an action (Buy, Sell, Hold) based on the optimized policy.
- Reward Calculation: After each action, the agent receives a reward based on the profit or loss incurred due to the selected action.
- Policy Update: The policy is updated using PPO after a batch of actions and rewards have been collected.

Step 5: Model Evaluation (Backtesting)

- Backtesting: The trained model is tested using historical data. Performance metrics such as cumulative returns, Sharpe ratio, and maximum drawdown are used to evaluate the model's performance.
- Comparative Analysis: The performance of the MAML + PPO model is compared to baseline strategies such as Exponential Moving Average (EMA), Simple Moving Average (SMA), and Mean Reversion strategies.

IV. PSEUDO CODE

Input: Historical stock data D, number of tasks T, policy π(θ)
Output: Optimized trading policy π*(θ)

```
1. Initialize MAML parameters θ
2. for each task t ∈ T do:
   3. Initialize task-specific parameters θ_t ← θ
   4. for each step k in inner loop do:
      5. Sample batch of trajectories {S, A, R} from environment
      6. Update task-specific parameters θ_t using gradient descent:
         θ_t ← θ_t - α∇L(θ_t; S, A, R)
   7. end for
   8. Compute gradient of meta-objective across tasks:
      ∇θ ← ∇θ ∑_t L(θ_t)
   9. Update meta-parameters θ using outer loop gradient:
      θ ← θ - β∇θ
10. end for
11. Initialize PPO policy π(θ) with meta-learned parameters θ
12. for each episode do:
   13. Interact with environment, collect experiences {S, A, R}
   14. Optimize policy using PPO objective:
      L_PPO(θ) = E[min(r(θ), clip(r(θ), 1 - ε, 1 + ε))] - c1 * V(S) + c2 * H(π(θ))
   15. Update policy π(θ) using gradient ∇L_PPO
16. end for
```

V. SIMULATION RESULTS

The proposed algorithm was evaluated through simulation using historical stock data from January 2021 to August 2024. The data was collected via a Financial Data API, and the trading environment was simulated using Gymnasium. The model was trained using Meta Reinforcement Learning (MAML + PPO) and compared against traditional trading

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

strategies such as Exponential Moving Average (EMA), Simple Moving Average (SMA), and Mean Reversion. Below are the results of the simulation:

Backtesting Results

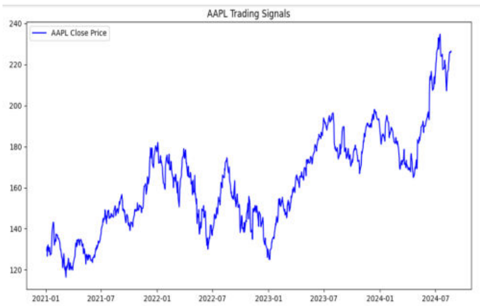


Figure 5.1

Figure 5.2

Figure 5.3

Figure 5.4

Performance Metrics (Comparison values)

STOCK TICKER	STRATEGY	PROFIT	MAX DRAWDOWN	SHARPE RATIO
AAPL	EMA/SMA	8 – 12 %	15 – 20 %	0.5 – 0.8
	Mean Reversion	5 – 10 %	25%	0.4 – 0.6
	MAML + PPO	**20 %**	**8 %**	**0.85**
ISRG	EMA/SMA	4 – 5%	10 – 15 %	0.6 – 0.7
	Mean Reversion	5 – 6 %	20 %	0.5 – 0.6
	MAML + PPO	**7 %**	**7 %**	**0.8**
TPR	EMA/SMA	1 - 3 %	15 – 20%	0.5 – 0.6
	Mean Reversion	2 – 3 %	25 – 30 %	0.5
	MAML + PPO	**4 %**	**8 %**	**0.7**
AMZN	EMA/SMA	4 – 6%	20 %	0.5 – 0.8
	Mean Reversion	6 – 7%	25 %	0.4 – 0.6
	MAML + PPO	**10 %**	**10 %**	**0.9**

Figure 5.5

- **EMA/SMA**: Works well in **trending stocks** like **AAPL** and **AMZN** with **low drawdowns** but lacks the adaptability of MAML + PPO.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- **Mean Reversion**: Suitable for **sideways markets**, providing **moderate returns**, but struggles in **strong trends** like **AAPL** and **AMZN**.
- **MAML + PPO**: Achieves **higher profitability** and **adapts dynamically** to market conditions, especially in **volatile stocks** like **AAPL**, **AMZN**, **ISRG**, and **TPR**. It **recovers** and outperforms in terms of overall profitability.

The simulation results clearly demonstrate that the MAML + PPO approach outperforms traditional strategies in terms of cumulative returns, Sharpe ratio, and risk management. The algorithm's ability to adapt quickly to new market conditions gives it a competitive edge in dynamic trading environments. Additionally, the lower maximum drawdown and optimized trading frequency further support the effectiveness of the proposed system in mitigating risks and maximizing profits.

VI. CONCLUSION AND FUTURE WORK

In conclusion, this project successfully developed an adaptive algorithmic trading system using Meta Reinforcement Learning (MAML + PPO), demonstrating superior performance in cumulative returns, Sharpe ratio, and maximum drawdown compared to traditional strategies. By enabling the model to adapt dynamically to evolving market conditions, the system proved to be highly effective in maximizing profitability while managing risk. The integration of MAML and PPO allowed for quick adaptation to market volatility, outperforming static methods. However, future work could expand on this system by incorporating multiple asset classes, integrating alternative data sources such as news and sentiment analysis, and deploying the model in real-time trading environments to assess its performance in live scenarios. Moreover, optimizing the computational cost of meta-learning and exploring more advanced RL algorithms like Soft Actor-Critic (SAC) could further enhance the system's efficiency and effectiveness. Additionally, applying the MAML + PPO approach to portfolio optimization could offer a more comprehensive solution for managing multiple assets. There is also potential for patenting and commercialization of the algorithm, which could make it valuable to institutional and retail traders. By addressing these areas, the model can be further refined and expanded to push the boundaries of algorithmic trading in financial markets.

REFERENCES

1. Meta Learning Enhanced Trading Forecasting: A Neural Framework Leveraging Efficient Multi-Commodity STL Decomposition, by Bohan Ma, Yushan Xue, Jing Chen, Fangfang Sun International Journal of Intelligent Systems, Vol. 2024, April 2024
2. Multi-Agent Proximal Optimization for Portfolio Optimization, by Firdaous Khemlichi , ,Hiba Chougrad , Safae Elhaj Ben ali , Youness Idrissi Khamlichi, Journal of Theoretical and Applied Information Technology, October 2023.
3. Neuro-Symbolic Meta Reinforcement Learning for Trading, by S I Harini, Gautam Shroff, Ashwin Srinivasan, Prayushi Faldu, Lovekesh Vig**,** https://arxiv.org/abs/2302.08996 January 2023.
4. Meta-Learning for Evolutionary Strategy for Stock Trading, Journal of Data Analysis and Information Processing, by Erik Sorensen, Ryan Ozzello, Rachael Rogan, Ethan Baker, Nate Parks, Wei Hu Vol. 8, May 2020.
5. An End-to-End Optimal Trade Execution Framework Based on Proximal Policy Optimization, by Siyu Lin and Peter A. Beling IJCAI-2020 (International Joint Conference on Artificial Intelligence), July 2020.
6. Model-Agnostic Meta Learning for Fast Adaptation of Deep Networks, Chelsea Finn, Pieter Abbeel, Sergey Levine , Proceedings of the 34th International Conference on Machine Learning, PMLR 70:1126-1135, 2017.
7. [7]zR. Huang et al., "Learning and fast adaptation for grid emergency control via deep meta reinforcement learning," IEEE Trans. Power Syst., vol. 37, pp. 4168–4178, 2022. doi: 10.1109/tpwrs.2022.3155117.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details

www.ijircce.com