

KẾT HỢP CÁC THUẬT TOÁN HỌC TĂNG CƯỜNG SÂU CHO GIAO DỊCH CHỨNG KHOÁN TỰ ĐỘNG

Phan Đại Dương - 19520482

Hồ Mỹ Hạnh - 19521470

Huỳnh Văn Hùng - 19521564

Tóm tắt

- Lớp: CS519.M11
- Link Github của nhóm:
<https://github.com/phandaiduonghcb/CS519.M11>
- Link YouTube video: <https://youtu.be/eGbBuziJAZw>



Phan Đại Dương



Huỳnh Văn Hùng

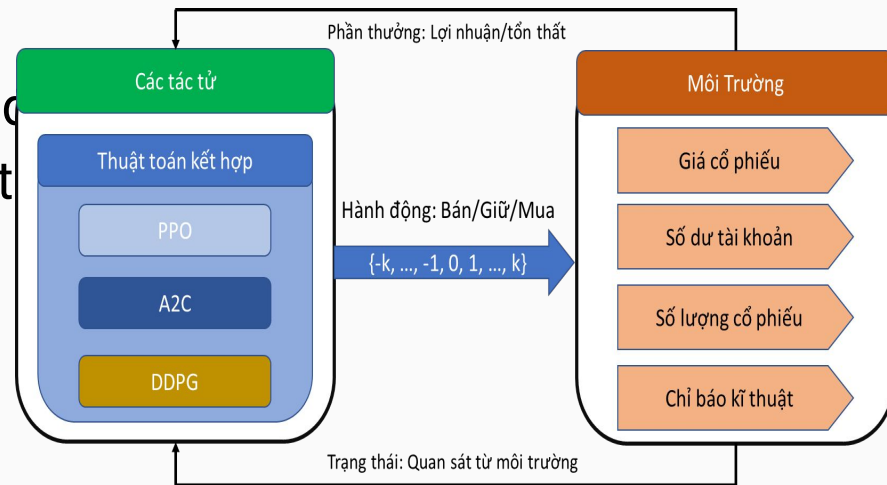


Hồ Mỹ Hạnh

Giới thiệu

Lĩnh vực chứng khoán đang ngày càng thu hút nhiều nhà đầu tư tham gia và nhiều chiến lược giao dịch cũng được đề xuất tuy nhiên chưa mang lại hiệu suất tốt.

➔ Chúng tôi đề xuất một phương pháp kết hợp các thuật toán học tăng cường sâu nhằm tìm ra được chiến lược giao dịch tối ưu.



Giới thiệu

Input

Giá đóng cửa điều chỉnh	Số cổ phiếu đang giữ	MACD	RSI	CCI	ADX	Số dư tài khoản
168.28999	332	-7.32456	37.90567	-54.34375	15.40578	53628
...						
67.845657	173	-0.80151	45.73895	-95.12863	20.04667	

Output

Hành động
-105
...
57

Mục tiêu

- Mô hình hoá bài toán và xây dựng môi trường thị trường chứng khoán để huấn luyện, đánh giá các thuật toán.
- Cài đặt các thuật toán A2C [1], DDPG [2], PPO [3] và huấn luyện trên môi trường thị trường chứng khoán đã được xây dựng
- Xây dựng được thuật toán học kết hợp dựa trên ba thuật toán học tăng cường là A2C, DDPG, PPO giúp tối đa hoá được lợi nhuận khi đầu tư.

Nội dung và Phương pháp

- Mô hình hoá bài toán: Sử dụng mô hình Markov Decision Process kết hợp với một số ràng buộc có thể ảnh hưởng tới việc giao dịch trong thực tế.
- Xây dựng môi trường thị trường chứng khoán:
 - Không gian trạng thái ngoài các thông tin cơ bản sẽ có thêm các chỉ báo kỹ thuật.
 - Không gian hành động phụ thuộc và số lượng cổ phiếu có thể mua và bán và số lượng cổ phần.
 - Cài đặt kỹ thuật “load-on-demand”

Nội dung và Phương pháp

- Các thuật toán A2C, DDPG, PPO sẽ được cài đặt và huấn luyện dựa trên [1], [2], [3].
- Thuật toán kết hợp gồm ba bước:
 - Chọn số n để huấn luyện lại các tác tử mỗi n tháng.
 - Đánh giá ba tác tử của A2C, PPO, DDPG với dữ liệu ba tháng gần nhất và chọn tác tử cho tỉ lệ Sharpe cao nhất.
 - Sử dụng tác tử đã chọn để dự đoán và giao dịch trong ba tháng tiếp theo.

➡ Ý tưởng là sử dụng tác tử khi xu hướng của thị trường thuộc thể mạnh của nó.

Nội dung và Phương pháp



Kết quả dự kiến

- Các thuật toán A2C, DDPG, PPO và thuật toán học kết hợp cho hiệu suất tốt hơn nhiều khi so sánh với chỉ số DJIA và chiến lược phân bổ danh mục đầu tư với phương sai nhỏ nhất.
- Tỷ lệ Sharpe của thuật toán học kết hợp sẽ cao nhất.
- Nhờ vào việc sử dụng mức ngại rủi ro, A2C, DDPG, PPO và thuật toán học kết hợp sẽ cho hiệu suất tốt hơn nhiều so với DJIA và chiến lược phân bổ danh mục đầu tư với phương sai nhỏ nhất khi thị trường bị sụp đổ (3 tháng đầu năm 2020).

Tài liệu tham khảo

- [1]. Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Timothy P. Lillicrap, Tim Harley, David Silver, Koray Kavukcuoglu: “Asynchronous Methods for Deep Reinforcement Learning”. ICML 2016: 1-8
- [2]. Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, Daan Wierstra: “Continuous control with deep reinforcement learning”. ICLR 2016: 1-8
- [3]. John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov: “Proximal Policy Optimization Algorithms”. CoRR 2017: 1-8