QUANTITATIVE TRADING IN SET MARKET WITH FinRL

Mahatsawat SARAPHANT[1*], Somporn PUNPOCHA[1] and Bumroong PUANGKIRD[2]

1 Faculty of Science and Technology, the University of the Thai Chamber of Commerce, Thailand; Mahatsawat45@gmail.com (M. S.) (Corresponding Author); somporn_punpocha@yahoo.com (S. P.)

2 Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Thailand; Bumroong.pu@kmitl.ac.th

ARTICLE HISTORY

Received: 9 June 2025 **Revised:** 23 June 2025 **Published:** 7 July 2025

ABSTRACT

The objective of this study is to develop and evaluate the performance of a quantitative trading strategy in the Stock Exchange of Thailand (SET) by applying Deep Reinforcement Learning (DRL) techniques through the FinRL library to create automated trading agents for Thai equities. The study focuses on a sample of 10 stocks listed in the SET index, using historical data from 2010 to 2024. Analytical variables include the Turbulence Index and technical indicators, which enable the agents to adapt their strategies in response to market volatility. The results demonstrate the potential of DRL in assessing investment performance over the period from 2022 to 2024, using performance metrics such as Cumulative Return, Sharpe Ratio, Annualized Volatility, and Maximum Drawdown. The experiment compares three DRL models—A2C, PPO, and DDPG—against the Mean-Variance Optimization (MVO) method and the SET index. Findings reveal that all three DRL models significantly outperform both the MVO portfolio and the SET index. Among them, the PPO model delivers the highest cumulative return at 23.2714%, followed by A2C at 22.1183% and PPO at 21.0826%. In contrast, the MVO approach yields a maximum return of 9.1662%, while the SET index records a negative return of -16.0943%.

Keywords: Deep Reinforcement Learning (DRL), Quantitative Trading, FinRL

CITATION INFORMATION: Saraphant, M., Punpocha, P., & Puangkird, P. (2025). Quantitative Trading in SET Market with FinRL. *Procedia of Multidisciplinary Research, 3*(7), 6

การพัฒนากลยุทธ์การลงทุนเชิงปริมาณในตลาดหลักทรัพย์แห่งประเทศไทย โดยใช้ไลบรารี FinRL

มหัสวัต สาระพันธ์[1*], สมพร ปั่นโภชา[1] และ บำรุง พวงเกิด[2]

1 คณะวิทยาศาสตร์และเทคโนโลยี, มหาวิทยาลัยหอการค้าไทย; Mahatsawat45@gmail.com (มหัสวัต) (ผู้ประพันธ์บรรณกิจ); somporn_punpocha@yahoo.com (สมพร)

2 คณะวิศวกรรมศาสตร์, สถาบันเทคโนโลยีเจ้าคุณทหารลาดกระบัง; Bumroong.pu@kmitl.ac.th

บทคัดย่อ

การศึกษานี้มีเป้าหมายเพื่อพัฒนาและประเมินประสิทธิภาพของกลยุทธ์การลงทุนเชิงปริมาณในตลาดหลักทรัพย์แห่งประเทศไทย โดยประยุกต์ใช้เทคนิค Deep Reinforcement Learning (DRL) ผ่านไลบรารี FinRL เพื่อสร้างตัวแทนการซื้อขายอัตโนมัติสำหรับการลงทุนในหุ้นไทย กลุ่มตัวอย่างประกอบด้วยหุ้น 10 ตัวในดัชนี SET โดยใช้ข้อมูลย้อนหลังตั้งแต่ปี 2010 ถึง 2024 และใช้ตัวแปรช่วยวิเคราะห์ ได้แก่ Turbulence index และ Technical Indicators ส่งผลให้ Agent ปรับกลยุทธ์ได้เหมาะสมกับความผันผวนของตลาด ผลลัพธ์แสดงให้เห็นถึงศักยภาพของการใช้ DRL ประเมินผลตอบแทนในช่วงปี 2022 ถึง 2024 โดยใช้ตัวชี้วัด ได้แก่ ผลตอบแทนสะสม (Cumulative Return), อัตราส่วนชาร์ป (Sharpe Ratio), ความผันผวน (Annualized Volatility) และการขาดทุนสูงสุด (Maximum Drawdown) การทดลองเปรียบเทียบระหว่างโมเดล DRL ได้แก่ A2C, PPO และ DDPG กับวิธี Mean-Variance Optimization (MVO) และดัชนี SET พบว่า โมเดล DRL ทั้งสามสามารถสร้างผลตอบแทนได้ดีกว่าพอร์ต MVO และ SET อย่างมีนัยสำคัญ โดย PPO ให้ผลตอบแทนสะสมสูงที่สุดอยู่ที่ 23.2714% รองลงมา คือ โมเดล A2C อยู่ที่ 22.1183% และโมเดล DDPG อยู่ที่ 21.0826% ในขณะที่วิธี MVO ให้ผลตอบแทนสูงที่สุดอยู่ที่ 9.1662% และดัชนี SET อยู่ที่ -16.0943%

คำสำคัญ: การเรียนรู้เสริมแรงเชิงลึก, การลงทุนเชิงปริมาณ, ไลบรารี FinRL

ข้อมูลการอ้างอิง: มหัสวัต สาระพันธ์, สมพร ปั่นโภชา และ บำรุง พวงเกิด. (2568). การพัฒนากลยุทธ์การลงทุนเชิงปริมาณในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้ไลบรารี FinRL. *Procedia of Multidisciplinary Research, 3*(7), 6

บทนำ

ในปัจจุบันตลาดการเงินเป็นหนึ่งในช่องทางการลงทุนที่สำคัญและเป็นที่นิยมในหมู่นักลงทุนทั่วโลก ไม่ว่าจะเป็นหุ้น (Stock) ตราสารหนี้ (Bond) ออปชั่น (Option) Forex และคริปโตเคอเรนซี่ (Cryptocurrency) ซึ่งให้ผลตอบแทนที่ดีกว่าการฝากเงินในธนาคาร ทำให้นักลงทุนต้องการเข้ามาลงทุนในตลาดการเงินมากขึ้น แต่การลงทุนนั้นมาพร้อมกับความเสี่ยงหรือความผันผวนเป็นสิ่งที่นักลงทุนต้องเจอพร้อมกับผลตอบแทน แต่อย่างไรก็ตามความผันผวนของตลาดหุ้นและความซับซ้อนของปัจจัยที่ส่งผลต่อราคาหุ้น ทำให้การตัดสินใจลงทุนกลายเป็นเรื่องที่ท้าทายสำหรับนักลงทุน การลงทุนมีหลากหลายแบบไม่ว่าจะเป็นการวิเคราะห์พื้นฐาน (Fundamental Analysis) และการวิเคราะห์ทางเทคนิค (Technical Analysis) ซึ่งการลงทุน 2 รูปแบบนี้ซึ่งนักลงทุนมักจะใช้ความรู้ ประสบการณ์ และอารมณ์ในการตัดสินใจ ทำให้เกิดปัญหาของการเทรดแบบดั้งเดิมมีหลายประการ ได้แก่ อคติทางพฤติกรรม (Behavioral Biases) นักลงทุนส่วนใหญ่มีแนวโน้มที่จะตัดสินใจผิดพลาดจากอารมณ์หรือความเชื่อที่ไม่เป็นเหตุเป็นผล เช่น ความมั่นใจในความสามารถของตัวเองมากเกินไป (Overconfidence Bias) กลัวการขาดทุนมากกว่าการได้กำไร (Loss Aversion) และการเทรดตามคนหมู่มากโดยไม่วิเคราะห์ข้อมูลเอง (Herding Effect)

ในปัจจุบันการใช้เทคโนโลยีทางด้านปัญญาประดิษฐ์ (Artificial Intelligence) และการเรียนรู้ของเครื่อง (Machine Learning) ได้รับความสนใจมากขึ้นและเป็นที่นิยมอย่างกว้างขวางทั่วโลก เนื่องจากมีศักยภาพในการวิเคราะห์และตัดสินใจข้อมูลจำนวนมากได้รวดเร็วและมีประสิทธิภาพเทคนิคที่มีการนำมาใช้ในตลาดการเงินคือ การเรียนรู้แบบเสริมแรง (Reinforcement Learning -RL) เป็นหนึ่งในวิธีของการเรียนรู้ของเครื่องจักร (Machine Learning) ซึ่งช่วยให้นักลงทุนสามารถพัฒนากลยุทธ์การลงทุนโดยอาศัยการเรียนรู้จากผลตอบแทนในอดีต มาวิเคราะห์ข้อมูลเพื่อใช้ในการสร้างโมเดลการเทรดแบบอัตโนมัติ แต่ Reinforcement Learning (RL) ใช้ Q-table ในการจดจำการเรียนรู้ ทำให้เก็บข้อมูลได้ไม่เยอะมาก ดังนั้นจึงได้ปรับเปลี่ยนโดยใช้การเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning: DRL) ซึ่ง DRL นั้นใช้ Neural Network เพื่อประมวลผลข้อมูลขนาดใหญ่ เช่น ราคา, ปริมาณ, Indicator หลายตัว หรือแม้แต่ข้อมูลเชิงเวลา (Time-series) ทำให้ได้กว่าการใช้ RL ในการสร้างกลยุทธ์การลงทุน ซึ่งการใช้ DRL นั้นสามารถสร้างกลยุทธ์การลงทุนที่เรียนรู้ Policy ที่ดีที่สุดจากประสบการณ์ และลดอคติของนักลงทุน แม้ DRL จะได้รับความนิยมในตลาดโลก แต่ยังมีงานวิจัยในบริบทของตลาดหุ้นไทยค่อนข้างจำกัด แต่อย่างไรก็ตามความแม่นยำของกลยุทธ์ดังกล่าวยังขึ้นอยู่กับการคาดการณ์แนวโน้มตลาดในอนาคต ซึ่งเป็นปัจจัยสำคัญที่ส่งผลต่อความสำเร็จของการลงทุน

ซึ่ง FinRL เป็นไลบรารีที่ออกแบบมาเพื่อให้ทุกคนสามารถเข้าถึงได้และเรียนรู้การใช้งานได้อย่างถูกวิธี โดย FinRL พัฒนาเพื่อใช้งานด้านการเงินโดยตรง และช่วยให้การใช้ Deep Reinforcement Learning (DRL) สำหรับการซื้อขายหุ้นอัตโนมัติเป็นไปได้อย่างมีประสิทธิภาพ ด้วยเครื่องมือที่ถูกพัฒนาให้ใช้ควบคู่กับการวิเคราะห์เชิงปริมาณ (Quantitative Finance) ซึ่งโมเดลใน FinRL ทั้ง 3 ได้แก่ A2C PPO และ DDPG สามารถเรียนรู้จากข้อมูลตลาดในอดีต ปรับกลยุทธ์การลงทุน และดำเนินการซื้อขายหุ้นโดยอัตโนมัติ บนสภาพแวดล้อมจำลอง (Environment) ในการทดลองโมเดลที่ใช้ในการเทรด ดังนั้นโมเดล DRL ช่วยสร้างระบบซื้อขายที่ไม่ต้องพึ่งการแทรกแซงของมนุษย์ ลดอคติในการลงทุน (behavioral bias) และสามารถเทรดได้แม้ในภาวะตลาดที่ผันผวน

การทบทวนวรรณกรรม

การลงทุนเชิงปริมาณ (Quantitative Trading) ได้รับความสนใจอย่างมากในช่วงหลายปีที่ผ่านมาและกลายเป็นแนวทางสำคัญในการวางกลยุทธ์ซื้อขายหลักทรัพย์ โดยเน้นการใช้แบบจำลองทางคณิตศาสตร์และอัลกอริทึม เพื่อช่วยในการตัดสินใจลงทุนอย่างเป็นระบบ (Chan, 2013) โดยเฉพาะเมื่อมีการนำปัญญาประดิษฐ์และการเรียนรู้ของเครื่องเข้ามาประยุกต์ใช้อย่างกว้างขวาง หนึ่งในเทคนิคที่โดดเด่นคือ Deep Reinforcement Learning (DRL) ซึ่งสามารถเรียนรู้กลยุทธ์การลงทุนแบบอัตโนมัติผ่านการโต้ตอบกับสภาพแวดล้อมในตลาดการเงิน เพื่อให้ตัวแทนการลงทุน (agent) สามารถเรียนรู้และตัดสินใจจากข้อมูลย้อนหลังและสถานะของตลาดได้อย่างมีประสิทธิภาพ Thomas (2015) ได้เสนอ

รากฐานของปัญหา Markov Decision Process (MDP) ซึ่งเป็นโครงสร้างพื้นฐานของ DRL ที่สามารถประยุกต์ใช้ในตลาดการเงินได้ดีงานวิจัยของ Ang et al. (2020) ได้พัฒนาไลบรารี FinRL ที่รองรับอัลกอริทึม DRL หลายประเภท เช่น DDPG, A2C และ PPO และได้นำไปใช้กับดัชนี Dow30 เพื่อศึกษาผลตอบแทนเชิงเปรียบเทียบ โดยพบว่ากลยุทธ์ DRL มีศักยภาพเหนือกว่าวิธีการดั้งเดิมในหลายกรณี และ Liu et al. (2021) ได้พัฒนา FinRL ซึ่งเป็นเฟรมเวิร์คแบบโอเพ่นซอร์สที่ออกแบบมาเพื่อรองรับการพัฒนาและทดลองกลยุทธ์การซื้อขายด้วย DRL โดยแบ่งโครงสร้างออกเป็น 3 ชั้น ได้แก่ Environment Layer, Agent Layer และ Application Layer จุดเด่นของ FinRL คือความสามารถในการโหลดข้อมูลตลาดจริงผ่าน API, สนับสนุนอัลกอริทึม DRL ชั้นนำ เช่น DDPG, PPO, A2C, และ TD3 พร้อมทั้งรองรับการ backtest โดยอัตโนมัติ อีกทั้งงานของ Liu et al. ยังเสนอแนวทาง training-testing-trading pipeline เพื่อลดช่องว่างระหว่างการฝึกโมเดลกับการใช้งานในตลาดจริง นอกจากนี้ยังรองรับการปรับ reward function ให้คำนึงถึงความเสี่ยง เช่น Sharpe Ratio และ turbulence index ซึ่งเหมาะสำหรับการลงทุนในตลาดที่มีความผันผวนสูง

ในส่วนของประเทศไทย แม้จะมีงานวิจัยจำนวนน้อยที่ใช้ DRL โดยตรง แต่ก็มีแนวโน้มเพิ่มขึ้น งานของ ณัฐพงษ์ เมืองไพศาล และ สมพร ปั่นโภชา (2564) ได้ศึกษาและประยุกต์ใช้อัลกอริทึม DDPG เพื่อหาสัดส่วนที่เหมาะสมในการลงทุนในหุ้นกลุ่ม SET50 โดยใช้ข้อมูลย้อนหลัง 10 ปี ผลการทดลองแสดงให้เห็นว่าแบบจำลอง DDPG ให้ผลตอบแทนและ Sharpe ratio ที่ดีกว่า benchmark แบบ equal-weight ในหลายกรณี และยังสามารถปรับสัดส่วนพอร์ตได้ตามลักษณะของข้อมูลที่เปลี่ยนไปในเวลา

สมมติฐานการวิจัย

1) การใช้โมเดล Deep Reinforcement Learning (DRL) ทำให้สร้างผลตอบแทนของพอร์ตได้ดีกว่าการลงทุนตามดัชนี SET

2) การใช้ Deep Reinforcement Learning (DRL) ในการทำโมเดลทำให้ผลตอบแทนของพอร์ตได้ดีกว่าการลงทุนตามวิธีแบบดั้งเดิม Mean Variance Optimization

3) การใช้ turbulence index และ momentum indicators จะช่วยเพิ่มความแม่นยำของกลยุทธ์ DRL ในสภาวะตลาดที่ผันผวน

วิธีดำเนินการวิจัย

งานวิจัยนี้ใช้แนวการเทรดทางเชิงปริมาณ (Quantitative Trading) โดยอาศัยเทคนิคการเรียนรู้ของเครื่องแบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning -DRL) สำหรับการซื้อขายหุ้นแบบอัตโนมัติ (Automated Trading) ในตลาดหลักทรัพย์แห่งประเทศไทย (SET) ให้มีประสิทธิภาพ โดยมีการทดลองแบ่งขั้นตอนออกเป็น 2 ส่วนหลัก ได้แก่ 1) การพัฒนาเพื่อฝึกฝนตัวแทน (Agent) ของอัลกอริทึม DRL 3 ตัว ได้แก่ A2C, PPO และ DDPG เปรียบเทียบกับกลยุทธ์แบบดั้งเดิม MVO กับดัชนี SET และ 2) การทดสอบประสิทธิภาพย้อนหลัง (Back Test)

การพัฒนาเพื่อฝึกฝนตัวแทน (Agent) ของอัลกอริทึม DRL 3 ตัวได้แก่ A2C, PPO และ DDPG เปรียบเทียบกับกลยุทธ์แบบดั้งเดิม MVO กับดัชนี SET มีกระบวนการดังนี้

1) เลือกชุดข้อมูลราคาหุ้น 10 ตัวในดัชนี SET ที่ต้องการทดสอบและดัชนีชี้วัดทางเทคนิค (Technical Indicators) โดยข้อมูลที่ได้จาก Yahoo Finance เป็นข้อมูลรายวันโดยใช้ชื่อย่อหลักทรัพย์ในตลาดหลักทรัพย์ไทย (SET)

2) ตรวจสอบและกำจัด missing values และจัดเรียงข้อมูลให้อยู่ในรูปtime-series โดยใช้ multi-index (date และ ticker)

3) ทำการสร้าง Feature Engineering สร้างดัชนีชี้วัดทางเทคนิค (Technical Indicators) ได้แก่ Moving Average Convergence Divergence (MACD), Relative Strength Index (RSI) และ Exponential Moving Average (EMA) 49, 89 และ 200

4) การสร้าง Turbulence Index ซึ่งใช้วัดความผันผวนของตลาดเพื่อควบคุมความเสี่ยงในช่วงตลาดมีความไม่แน่นอนหรือผันผวนสูงกว่าปกติ เช่น โควิด-19

$$Turbulence_t = \left(r_t - \mu \right)^T \sum^{-1} \left(r_t - \mu \right)$$

โดย r_t = เวกเตอร์ของ return ณ เวลา t (ของหลายๆ หุ้น)

μ = เวกเตอร์ค่าเฉลี่ยของ return จากข้อมูลย้อนหลัง

$\sum$ = covariance matrix ของ return ในอดีต

$\sum^{-1}$ = inverse ของ covariance matrix

ค่าที่ได้คือ scalar ที่บอกความ “ห่าง” ของ return ณ เวลา t จากความปกติในอดีต โดยมักใช้เป็นเกณฑ์ในการตัดสินใจปรับกลยุทธ์การเทรดหรือป้องกันความเสี่ยงใน FinRL ซึ่งค่า Turbulence Index คำนวณจากระดับความผันผวนของตลาด โดยอ้างอิงจากการกระจายตัวของราคาหุ้นในช่วงเวลาต่างๆ โดย

- ค่า Turbulence สูง หมายถึงตลาดมีความปั่นป่วนสูง อาจเกิดวิกฤต
- ค่า Turbulence ต่ำ หมายถึงตลาดมีความมั่นคงมากขึ้น

5) แบ่งข้อมูลออกเป็น 2 ช่วง ได้แก่ Train set และ Trade set 1. Training set สำหรับการฝึก Agent (ปี 2010-2022) และ Testing set สำหรับทดลอง Agent (ปี 2022-2024)

6) ใช้ StockTradingEnv ของ FinRL ในการจำลองตลาดหุ้น โดยจำลองพอร์ตที่มีเงินเริ่มต้น 1,000,000 บาท และสามารถซื้อ-ขายหุ้นได้ตามกลยุทธ์ของ Agent จากนั้น Environment จะส่งข้อมูล observation ที่ประกอบด้วยราคาหุ้น ปริมาณเงินสดในพอร์ต สถานะของพอร์ต และ indicators ต่างๆแล้วสุดท้าย Reward function (การให้ รางวัล หรือผลตอบแทน แก่ Agent เพื่อช่วยให้ Agent เรียนรู้การกระทำที่มีประสิทธิภาพมากขึ้นในอนาคต) ซึ่งReward function ในงานวิจัยคำนวณจาก การเปลี่ยนแปลงมูลค่าพอร์ตการลงทุน (Investment Portfolio Value) ในแต่ละรอบเวลา

$$r(s, a, s') = v' - v$$

$r(s, a, s')$ ค่าตอบแทนคือความต่างระหว่างมูลค่าพอร์ตใหม่ (v') กับมูลค่าก่อนหน้า (v) หลังจาก action (a) ถูกกระทำเป็น reward function ที่ตรงไปตรงมา ใช้วัด กำไรที่แท้จริง

7) กำหนด Time Step สำหรับการฝึกฝนของโมเดลไว้ที่ 300,000 รอบ ให้แต่ละอัลกอริทึมได้เรียนรู้และปรับค่า hyperparameter เพื่อพัฒนาการเทรดให้มีประสิทธิภาพมากที่สุด

ตารางที่ 1 แสดงค่า Hyperparameters ของแต่ละโมเดลที่ใช้ในการทดลอง

Hyperparameters	**A2C**	**DDPG**	**PPO**
n_steps	10	-	2048
Learning rare	0.0001	0.0001	0.0005
Batch size	-	128	128
Buffer size	-	100,000	-
Ent_coef	-	-	0.01
Clip range	-	-	0.2

การทดสอบประสิทธิภาพย้อนหลัง (Back Test)

1) นำโมเดลที่ปรับค่าแล้วนี้ไป Back test ให้ Agent ดำเนินการซื้อขายจริงตามข้อมูลในอดีตในชุดข้อมูล Trade set ที่กำหนดไว้

2) แล้วสุดท้ายนำผลที่เทรดไปบันทึกข้อมูลพอร์ตและคำนวณตัวชี้วัดทางการเงิน ได้แก่ 2.1) Cumulative Return, 2.2) Sharpe Ratio, 2.3) Annual Volatility และ 2.4) Maximum Drawdown เพื่อเปรียบเทียบกับวิธีการลงทุนแบบดั้งเดิมอย่าง Mean Variance Optimization (MVO) และดัชนี SET

- Cumulative Return คือ ผลตอบแทนสะสมทั้งหมดของพอร์ตการลงทุนตลอดช่วงเวลาหนึ่ง ซึ่งแสดงให้เห็นว่าการลงทุนเติบโตขึ้นเท่าใดจากจุดเริ่มต้นจนถึงสิ้นสุดช่วงเวลามีสมการดังนี้

$$R = \frac{v - v_0}{v_0}$$

R คือ Cumulative return

v คือมูลค่าพอร์ตการลงทุนสุดท้าย (Final portfolio value)

v_0 คือมูลค่าตั้งต้นของพอร์ต (Initial portfolio value)

- Sharpe Ratio คือ อัตราส่วนที่ใช้วัด "ผลตอบแทนส่วนเพิ่มต่อหน่วยความเสี่ยง" ของการลงทุน ช่วยให้นักลงทุนสามารถเปรียบเทียบประสิทธิภาพของพอร์ตที่มีระดับความเสี่ยงต่างกันได้อย่างมีมาตรฐาน มีสมการดังนี้

$$S_T = \frac{R_p - R_f}{\sigma_p}$$

S_T คือค่า Sharpe ratio

R_p คือผลตอบแทนของ Portfolio

R_f คืออัตราผลตอบแทนไร้ความเสี่ยง (risk-free rate)

σ_p คือส่วนเบี่ยงเบนมาตรฐานของผลตอบแทน (ความผันผวน)

- Annualized Volatility คือ ค่าที่ใช้วัด "ระดับความผันผวนเฉลี่ยของผลตอบแทน" ต่อปี ซึ่งใช้เป็นตัวแทนของ “ความเสี่ยง” ของสินทรัพย์หรือพอร์ตการลงทุน มีสมการดังนี้

$$\sigma_a = \sqrt{\frac{\sum_{i=1}^{n}(r_i - \bar{r})^2}{n-1}}$$

σ_a คือ Annualized volatility (ค่าความผันผวนเฉลี่ยรายปี)

r_i คือ ผลตอบแทนรายปีในปีที่ i

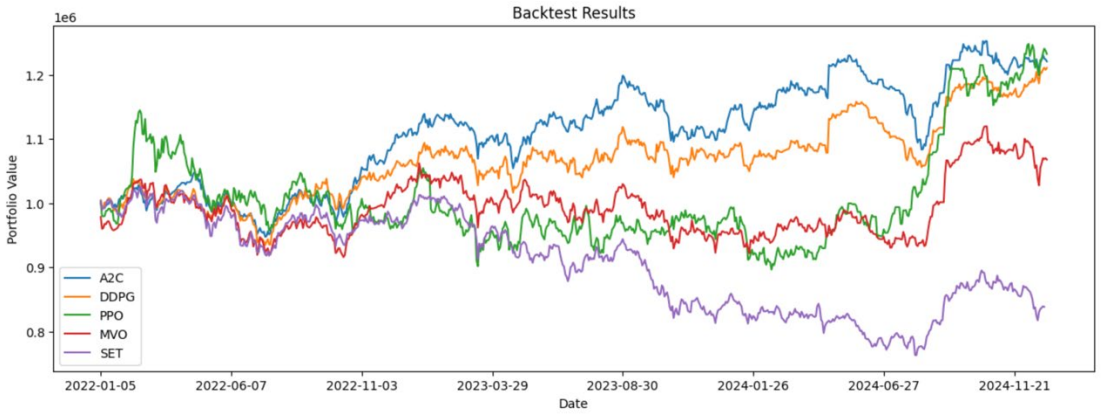
$\bar{r}$ คือ ค่าเฉลี่ยของผลตอบแทนรายปี

n คือ จำนวนปี

- Maximum Drawdown (MDD) คำนวณโดยการเอาผลขาดทุนจากค่าสูงสุด (Peak Value) ของพอร์ตการลงทุนในช่วงเวลานั้นๆนับไปจนถึงจุดต่ำสุด (Valley Value) ของพอร์ตการลงทุน จากนั้นนำไปหารกับค่า Peak Value อีกครั้งก็จะได้ค่า Maximum Drawdown เป็นเปอร์เซ็นต์ มีสมการดังนี้

$$MDD(\%) = \frac{(Valley\ Value - Peak\ Value)}{Valley\ Value}$$

ผลการวิจัย



ภาพที่ 1 กราฟแสดง Portfolio Value A2C DDPG PPO MVO และดัชนี SET โดยเริ่มต้นจากเงินต้น 1,000,000 บาท ระหว่างปี 2022-2024

ตารางที่ 2 ผลการเปรียบเทียบของโมเดลการลงทุนแต่ละแบบ

	A2C	DDPG	PPO	Min-Var	SET
Initial Value (บาท)	1M	1M	1M	1M	1M
Final Value (บาท)	1,221,183	1,210,826	1,232,714	1,067,940	839,057
Cumulative return	22.1183%	21.0826%	23.2714%	6.7947%	-16.0943%
Annual volatility	10.133%	10.2026%	16.0415%	11.7587%	11.3632%
Sharpe ratio	0.735079	0.701717	0.533039	0.317931	-0.481391
Max drawdown	-11.9306%	-9.9216%	-21.7007%	-13.2167%	-25.6356%

สรุปและอภิปรายผลการวิจัย

ในการศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อพัฒนาและประเมินประสิทธิภาพของกลยุทธ์การลงทุนเชิงปริมาณในตลาดหลักทรัพย์แห่งประเทศไทย โดยประยุกต์ใช้เทคนิค Deep Reinforcement Learning (DRL) ภายใต้กรอบการทำงานของ FinRL ซึ่งเป็นเครื่องมือโอเพนซอร์สที่ใช้สำหรับฝึกและประเมินนโยบายการลงทุนโดยอาศัยการเรียนรู้แบบเสริมแรงลึก (DRL agent) งานวิจัยได้ดำเนินการทดสอบด้วยข้อมูลราคาหุ้นย้อนหลังของกลุ่ม SET ร่วมกับข้อมูลทางเทคนิคและตัวแปรภาวะตลาดอย่างความปั่นป่วนของตลาด (Turbulence Index) โดยทำการฝึก agent ด้วยข้อมูลในช่วงปี 2010-2022 และนำไปทดสอบย้อนกลับ (Backtest) กับข้อมูลช่วงปี 2022-2024 โดยในแต่ละโมเดลได้ปรับ Hyperparameters และเซ็ตจำนวนรอบการฝึกไว้ที่ 300,000 รอบ

ผลการทดลองพบว่าโมเดล DRL ทั้งสามแบบสามารถสร้างผลตอบแทนสะสมที่สูงกว่า SET Index และกลยุทธ์ MVO อย่างมีนัยสำคัญ โดยเฉพาะโมเดล PPO ที่ให้ผลตอบแทนสะสมสูงสุดที่ 23.2714% รองลงมา คือ โมเดล A2C ให้ผลตอบแทนสะสมอยู่ที่ 21.0826% และโมเดล DDPG ให้ผลตอบแทนสะสมอยู่ที่ 21.0826% ในขณะที่เมื่อวิเคราะห์จากค่า Sharpe Ratio ค่าSharpe Ratio ของโมเดล A2C มีค่าสูงที่สุดที่ 0.735079 แสดงถึงประสิทธิภาพในการบริหารความเสี่ยงต่อผลตอบแทนของโมเดล A2C ได้ดีกว่าโมเดลอื่นๆ รองลงมา คือ โมเดล DDPG มีค่า 0.701717 และโมเดล PPO มีค่า 0.533039 เนื่องมาจากโมเดล PPO มี Maximum Drawdown สูงที่สุด และมีความผันผวนสูงที่สุดอยู่ที่ 16.0415% ในขณะที่ส่วนของโมเดล DDPG มีลักษณะการฟื้นตัวที่รวดเร็วจาก drawdown และมีผลตอบแทนใกล้เคียงกับ A2C

การที่กลยุทธ์ DRL สามารถมีประสิทธิภาพเหนือกว่าทั้งดัชนี SET และ MVO ได้อย่างมีนัยสำคัญ แสดงให้เห็นถึงศักยภาพของการใช้ AI และ machine learning ในการเรียนรู้เชิงลึกเพื่อการลงทุน โดยเฉพาะ DDPG ซึ่งเป็น agent แบบ policy-gradient ที่สามารถจัดการกับปัญหาที่มี continuous action space ได้ดี และสามารถจูนโมเดลให้ตอบสนองกับความเปลี่ยนแปลงของตลาดอย่างมีประสิทธิภาพ นอกจากนี้ PPO ซึ่งเป็น agent ที่ได้รับความนิยมสูงสุดในการใช้งานทั่วไป ก็แสดงศักยภาพในการให้ผลตอบแทนที่เสถียร โดยมีการควบคุมการอัปเดทตนโยบายไม่ให้เปลี่ยนแปลงมากเกินไปในแต่ละรอบของการฝึก ส่วน A2C แม้จะมีโครงสร้างที่ง่ายที่สุดในบรรดา agent ทั้งหมด แต่ก็สามารถให้ผลตอบแทนที่ดีและมีความผันผวนต่ำกว่า MVO อย่างชัดเจน

ข้อค้นพบที่สำคัญจากการศึกษานี้ คือ ประโยชน์ของการประมวลผลข้อมูลล่วงหน้าหรือ Feature Engineering โดยเฉพาะการใช้ตัวแปรอย่าง turbulence index ซึ่งสามารถบ่งชี้ความไม่แน่นอนของตลาดในแต่ละช่วงเวลาได้อย่างชัดเจน เมื่อนำตัวแปรเหล่านี้เข้าโมเดล DRL จะช่วยให้ agent สามารถปรับกลยุทธ์การลงทุนให้สอดคล้องกับระดับความเสี่ยงของตลาดในขณะนั้นได้อย่างทันท่วงที อีกทั้งการใช้ momentum ยังมีส่วนช่วยให้ agent สามารถจับเทรนด์ของตลาดได้ดีขึ้นและลดโอกาสการขาดทุนในช่วง drawdown

อย่างไรก็ตาม จากการทดลองยังพบว่าตัวแทน (agent) บางตัวมีความไวต่อพารามิเตอร์ เช่น n steps, learning rate, discount factor, reward scaling, batch size, buffer size และจำนวนรอบในการฝึก (timesteps) หากไม่จูนค่าเหล่านี้อย่างเหมาะสม อาจทำให้ประสิทธิภาพของโมเดลลดลง อีกทั้งการ backtest ยังอาจมีข้อจำกัดในแง่ของข้อมูลที่ใช้ ซึ่งแม้ว่าจะสะท้อนถึงสภาวะตลาดในอดีตได้ดี แต่ก็ไม่สามารถการันตีผลลัพธ์ในอนาคตได้

ข้อเสนอแนะที่ได้รับจากการวิจัย

จากผลการศึกษาการพัฒนากลยุทธ์การลงทุนเชิงปริมาณในตลาดหลักทรัพย์แห่งประเทศไทยโดยใช้ไลบรารี FinRL ผู้วิจัยมีข้อเสนอแนะ ดังนี้

1) ในการลงทุนต้องติดตามปัจจัยภายนอกอื่นๆด้วย เช่น การเมืองในและต่างประเทศ สงคราม ภัยพิบัติ ล้วนก่อให้เกิดล้วนก่อให้เกิดความผันผวนที่สูงมากด้วย

2) ตลาดหุ้นทั่วโลกแตกต่างกัน เพราะฉะนั้นในการใช้โมเดลต้องปรับพารามิเตอร์เพิ่มเติมให้เหมาะสมกับแต่ละตลาด

3) ควรทดลองใช้โมเดล DRL อื่นๆ ที่ซับซ้อนมากขึ้น เช่น SAC (Soft Actor-Critic), TD3, หรือ Transformer-based RL เพื่อดูว่าประสิทธิภาพสามารถพัฒนาได้มากขึ้นหรือไม่ในตลาดไทย

4) ควรต้องลองปรับค่า timesteps เพิ่มมากขึ้นไป เพื่อให้ agent ได้เรียนรู้ใน environment นานมากขึ้น

5) ควรมีการนำ DRL ไปทดลองในสภาวะจริงแบบ Paper Trading หรือ Deploy แบบจำลองในพอร์ตจำลองก่อนใช้งานจริง เพื่อศึกษาพฤติกรรมโมเดลภายใต้สภาพตลาดสด และทำการ fine-tune ค่าพารามิเตอร์เพิ่มเติมให้เหมาะสมกับลักษณะของหุ้นหรือสินทรัพย์เป้าหมาย

เอกสารอ้างอิง

นันทรัตน์ ตรีพรชัยศักดิ์, นงนภัส แก้วพลอย และ กนกพรรณ แก้วเนตร. (2560). *การเปรียบเทียบมูลค่าความเสี่ยงของกลุ่มหลักทรัพย์ภายใต้การลงทุนที่มีความเสี่ยงแตกต่างกัน*. การประชุมนำเสนอผลงานวิจัยระดับบัณฑิตศึกษา ครั้งที่ 12 (หน้า 1496-1507), มหาวิทยาลัยรังสิต, 17 สิงหาคม 2560,.

โสรยา แพสวุรรณ์. (2562). *การพยากรณ์อัตราแลกเปลี่ยนเงินตราต่างประเทศโดยใช้การเรียนรู้เชิงลึก*. การศึกษาค้นคว้าอิสระปริญญาวิทยาศาสตรมหาบัณฑิต, มหาวิทยาลัยหอการค้าไทย.

ณัฐพงษ์ เมืองไพศาล และ สมพร ปั่นโภชา. (2564). *การจัดสรรเงินลงทุนในพอร์ตการลงทุนด้วยวิธี Deep Deterministic Policy Gradients*. ใน การประชุมนำเสนอผลงานวิจัยบัณฑิตศึกษาระดับชาติ ครั้งที่ 16 (หน้า 254-265). บัณฑิตวิทยาลัย มหาวิทยาลัยรังสิต.

Li, Y. (2017). Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., ... & Wierstra, D. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*.

Liu, X.-Y., Yang, H., Chen, Q., Zhang, R., Yang, L., Xiao, B., & Wang, C. D. (2022). FinRL: *A deep reinforcement learning library for automated stock trading in quantitative finance*. Deep Reinforcement Learning Workshop, 34th Conference on Neural.

Liu, X. Y., Yang, H., Gao, J., & Wang, C. D. (2021, November). *FinRL: Deep reinforcement learning framework to automate trading in quantitative finance*. In Proceedings of the second ACM international conference on AI in finance (pp. 1-9).

Liu, Y., Zhang, J., & Yang, Q. (2020). Deep reinforcement learning in financial trading: A review. *Journal of Finance and Data Science, 6*(4), 242-260.

Thomas, G. W. (n.d.). (2025). *Notes on Markov Decision Processes (MDPs)*. Stanford University.

Xiong, Z., Liu, X. Y., Zhong, S., Walid, A., & Zhang, X. (2018). Practical deep reinforcement learning approach for stock trading. *arXiv preprint arXiv:1811.07522*.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright: © 2025 by the authors. This is a fully open-access article distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).