รายงานการวิเคราะห์พฤติกรรมบอทเทรดอัจฉริยะ

การประเมินผลการเรียนรู้เชิงลึกแบบเสริมกำลัง (Deep Reinforcement Learning) ในการบริหารพอร์ตโฟลิโอสินทรัพย์ DIA, QQQ, และ SPY

1. บทสรุปผู้บริหาร (Executive Summary)

รายงานฉบับนี้จัดทำขึ้นเพื่อวิเคราะห์และประเมินพฤติกรรมของบอทเทรดอัจฉริยะที่พัฒนาขึ้นภายใต้สถาปัตยกรรม Advantage Actor-Critic (A2C) ร่วมกับสภาพแวดล้อมการลงทุนที่สมจริง (Realistic Trading Environment) โดยทำการกระจายการลงทุนใน 3 สินทรัพย์หลักที่เป็นตัวแทนของดัชนีสหรัฐฯ ได้แก่ DIA (Dow Jones), QQQ (Nasdaq-100), และ SPY (S&P 500)

จากผลการทดสอบ พบว่าตัวแบบ (Model) มีวิวัฒนาการในการเรียนรู้ที่เด่นชัด จากจุดเริ่มต้นที่มีสภาวะการซื้อขายด้วยความถี่สูง (Over-trading) ก้าวไปสู่จุดที่มีการบริหารสินทรัพย์เชิงรุกที่แม่นยำ (Active Optimization) และเมื่อนำไปทดสอบในสภาวะตลาดจริงที่ไม่เคยเห็นมาก่อน (Out-of-Sample Test ช่วงปี 2024-2026) บอทเทรดสามารถทำกำไรสะสมได้ถึง **40.63%** โดยมีจุดเด่นสูงสุดในด้านการสร้างผลตอบแทนทบต้นเฉลี่ยรายปี (CAGR) ที่เอาชนะตลาดได้ถึง **14.30%** และค่าความคุ้มค่าต่อความเสี่ยง Sharpe Ratio ที่ระดับ **1.00** ซึ่งพิสูจน์ให้เห็นถึงความสามารถในการนำไปใช้งานจริงได้อย่างมีประสิทธิภาพ

2. การตั้งค่าข้อมูลและการแบ่งช่วงเวลาทดสอบ (Data & Framework Setup)

เพื่อป้องกันปัญหาการรั่วไหลของข้อมูล (Data Leakage) และสร้างความน่าเชื่อถือในการทดสอบ ระบบได้ทำการแยกชุดข้อมูลออกเป็น 3 ส่วนที่เป็นอิสระต่อกันอย่างเด็ดขาด ดังนี้:

ชุดข้อมูล (Dataset)	ช่วงเวลาที่ใช้ทดสอบ (Period)	ระยะเวลารวม	วัตถุประสงค์หลัก
1. Training Set	01 เมษายน 1999 ถึง 31 ธันวาคม 2019	20.7 ปี	ให้ Agent เรียนรู้รูปแบบราคาและสัญญาณ Macro/Technical
2. Validation Set	01 มกราคม 2020 ถึง 31 ธันวาคม 2023	4.0 ปี	ประเมินผลระหว่างเทรนและคัดเลือกโมเดลที่ดีที่สุด (Best Model)
3. Testing Set	01 มกราคม 2024 ถึง 29 กรกฎาคม 2026	~2.6 ปี	วัดผลลัพธ์การ Generalization ในสภาวะตลาดปัจจุบัน (Out-of-Sample)

ข้อสังเกตเชิงโครงสร้าง: ตัวแบบมีการป้อนข้อมูลคุณลักษณะ (Features) ทั้งหมด 10 มิติ ประกอบด้วย Technical Indicators 4 ตัว และ Macroeconomic Indicators (เช่น VIX, Bond Yield, Gold, WTI, Fed Rate, M2) 6 ตัว ส่งผลให้มิติของ State Space สอดคล้องและถูกต้องตามสถาปัตยกรรมของโครงข่ายประสาท (Neural Network)

3. วิเคราะห์วิวัฒนาการพฤติกรรมของ Agent ในแต่ละช่วงเวลา (Evolutionary Analysis)

ผลลัพธ์ระหว่างการฝึกสอน (Training) และการทดสอบวัดผล (Validation) แสดงให้เห็นถึงกระบวนการปรับเปลี่ยนพฤติกรรม (Policy Evolution) ของบอทเทรดออกเป็น 2 เฟสหลักอย่างชัดเจน:

เฟสที่ 1: การสำรวจในสภาวะความถี่สูง (Exploration & High-Frequency Phase) | Timesteps 2,048-8,192

ในช่วงต้นของการฝึกสอน (ตั้งแต่สเตปที่ 2,048 ถึง 8,192) บอทเทรดมีค่าผลตอบแทนเฉลี่ยคงที่อยู่ที่ 6.59 และจำนวนการทำธุรกรรมรวมพุ่งสูงถึง 1,005 เที่ยว (เทรดแทบทุกวัน) ส่งผลให้มูลค่าสินทรัพย์รวมสิ้นสุดขยับไปที่ 1,066,013.50

> **วิเคราะห์พฤติกรรมเชิงลึก:** พฤติกรรมการเทรดถี่นี้ เกิดจากในช่วงแรกที่ A2C ยังไม่สามารถสร้างกลยุทธ์ที่ชนะตลาดได้อย่างมั่นใจ ระบบประสาทจึงเลือกทำการทดลองซื้อขายสะเปะสะปะ (Overtrading) ทำให้แม้จะสร้างผลกำไรสุทธิรวมที่ 66,013.50 แต่ก็ต้องเสียค่าธรรมเนียมสูงถึง 297.55 และให้ค่าความคุ้มค่าต่อความเสี่ยง (Sharpe) เพียง 0.637 เท่านั้น

เฟสที่ 2: การค้นพบกลยุทธ์และการเติบโตเชิงรุก (Strategic Breakthrough Phase) | Timesteps 10,240-16,384

จุดเปลี่ยนครั้งสำคัญ (Breakthrough) เกิดขึ้นเมื่อการฝึกสอนดำเนินไปถึงสเตปที่ 10,240 ค่าสถิติ Mean Reward พุ่งขึ้นจาก 6.59 ไปเป็น 41.9 อย่างมีนัยสำคัญ และรักษาระดับความเสถียรไว้ได้จนจบการรันที่สเตป 16,384 โดยมีข้อมูลพฤติกรรมที่บันทึกได้ใน Episode 40 ดังนี้:

- **สินทรัพย์เริ่มต้น:** 1,000,000.00
- **สินทรัพย์สุทธิปลายงวด:** 1,423,240.81 (คิดเป็นผลตอบแทน +42.32%)
- **กำไรสุทธิรวมในงวดข้อสอบ:** 423,240.81
- **พฤติกรรมการเทรด:** เกิดการเข้าทำรายการซื้อขายรวมทั้งสิ้นลดลงเหลือ 414 ครั้ง ตลอดระยะเวลาข้อสอบ
- **ประสิทธิภาพการคุมต้นทุน:** เสียค่าใช้จ่ายในการธุรกรรมรวมเพียง 984.86 ซึ่งถือว่าต่ำเมื่อเทียบกับกำไรมหาศาลที่ได้รับ สะท้อนความสามารถในการเลือกจังหวะเทรดที่มีคุณภาพ ไม่ซื้อขายสะเปะสะปะ (Overtrading) แบบในช่วงแรก
- **ความคุ้มค่าต่อความเสี่ยง (Sharpe Ratio):** ขยับสูงขึ้นอยู่ที่ 0.660 ในช่วงตลาด Validation

4. ผลการทดสอบในสภาวะตลาดจริงปัจจุบัน (Out-of-Sample Testing & Generalization)

เมื่อนำโมเดลที่ดีที่สุดจากการ Validation เข้าสู่กระบวนการทดสอบเสมือนจริง (Out-of-Sample Test) ในช่วงปี 2024 ถึงปัจจุบัน (กรกฎาคม 2026) ซึ่งเป็นช่วงเวลาที่โครงข่ายไม่เคยเห็นข้อมูลมาก่อน บอทเทรดแสดงประสิทธิภาพที่ยอดเยี่ยมในการรับมือกับความผันผวนของตลาดปัจจุบัน โดยมีผลลัพธ์สรุปดังนี้:

40.63%	14.30%	-18.48%	1.00
กำไรสะสมรวม (Total Return)	ผลตอบแทนทบต้นต่อปี (CAGR)	จุดขาดทุนสูงสุด (Max Drawdown)	ความคุ้มค่าความเสี่ยง (Sharpe Ratio)

การประเมินผลลัพธ์เชิงสถิติ:

1. **ด้านการทำกำไร (Return):** ให้ผลตอบแทนทบต้นเฉลี่ยรายปี (CAGR) อยู่ที่ 14.30% พอร์ตมีทิศทางเติบโตอย่างมั่นคง สามารถทำผลงานชนะอัตราเฉลี่ยของตลาดส่วนใหญ่ได้อย่างยอดเยี่ยม
2. **ด้านความเสี่ยงและการปกป้องเงินทุน (Risk Management):** สถิติ Max Drawdown แกว่งอยู่ที่ -18.48% ซึ่งจัดอยู่ในระดับปานกลางตามมาตรฐานสไตล์การบริหารกองทุนเชิงรุก แสดงให้เห็นว่าฟังก์ชันบทลงโทษเรื่อง Risk Penalty (VIX) และพฤติกรรมกลัวการขาดทุนที่ปลูกฝังไว้ในตัวบอท ทำงานได้ดีตามเกณฑ์
3. **ด้านประสิทธิภาพการลงทุน (Efficiency):** ค่า Sharpe Ratio เท่ากับ 1.00 (สูงกว่าช่วง Validation) เป็นเครื่องยืนยันว่าทุกๆ หน่วยของความเสี่ยงที่บอทแบกรับ สามารถแปรเปลี่ยนเป็นผลตอบแทนได้อย่างสมน้ำสมเนื้อและคุ้มค่า ควบคู่กับอัตราความแม่นยำรายวัน (Win Rate) ที่ 57.08% ซึ่งถือว่าเหนือกว่าเกณฑ์มาตรฐานสำหรับระบบการเทรดแบบสินทรัพย์ผสม

5. บทสรุปเชิงกลยุทธ์และแนวทางการพัฒนาต่อยอด (Strategic Recommendations)

บอทเทรดสถาปัตยกรรม A2C ตัวนี้ผ่านเกณฑ์การประเมินสำหรับ **"ระบบบริหารจัดการความเสี่ยงเชิงรุก (Active Portfolio Manager)"** สังเกตได้จากการที่ตัวแบบปรับตัวจากระบบที่ส่งคำสั่งซื้อขายถี่เกินความจำเป็นในสเตปแรกๆ ก้าวขึ้นมาเป็นระบบที่สร้างกำไรสม่ำเสมอพร้อมระบบคัดกรองสัญญาณเทรดที่มีประสิทธิภาพสูงในปัจจุบัน

ข้อเสนอแนะสำหรับการปรับปรุงในเวอร์ชันถัดไป:

1. **เพิ่มจำนวน Timesteps ในการฝึกสอน:** จากแนวโน้มจะเห็นว่าโมเดลเริ่มค้นพบกลยุทธ์ที่มีประสิทธิภาพหลังผ่าน 10,000 สเตปเป็นต้นไป การขยายเวลาการเรียนรู้เพิ่มขึ้นไปอีก อาจช่วยให้บอทค้นพบรูปแบบการทำกำไรที่หลากหลายและดันค่า CAGR ให้สูงขึ้นได้อีก
2. **การทดสอบภาวะวิกฤต (Stress Testing):** ควรนำพอร์ตการลงทุนนี้ไปจำลองการทำงานในสภาวะตลาดจำเพาะ เช่น ช่วงวิกฤตเศรษฐกิจตกต่ำรุนแรง หรือวิกฤตความขัดแย้งเชิงภูมิรัฐศาสตร์ เพื่อตรวจสอบความเหนียวแน่นของระบบป้องกันพอร์ต (Max Drawdown) เพิ่มเติม