

โครงการคณิตศาสตร์

เรื่อง

การพัฒนาแบบจำลองเชิงปริมาณเพื่อสร้างผลตอบแทน
ที่เหนือกว่าตลาดด้วยกลยุทธ์แบบกองทุนป้องกันความเสี่ยง

Developing Quantitative Models to
Generate Alpha Using Hedge Fund Strategies

นำเสนอโดย

นางสาวณัฐริชา ทรงคำ รหัสนักศึกษา 6604105324

นางสาวธิดารัตน์ คัดจักร รหัสนักศึกษา 6604105326

รายงานฉบับนี้เป็นส่วนหนึ่งของรายวิชา 10305492 โครงการ

ภาคเรียนที่ 1 ปีการศึกษา 2569

สาขาวิชาคณิตศาสตร์ คณะวิทยาศาสตร์

มหาวิทยาลัยแม่โจ้

โครงการคณิตศาสตร์

เรื่อง

การพัฒนาแบบจำลองเชิงปริมาณเพื่อสร้างผลตอบแทน
ที่เหนือกว่าตลาดด้วยกลยุทธ์แบบกองทุนป้องกันความเสี่ยง

Developing Quantitative Models to
Generate Alpha Using Hedge Fund Strategies

นำเสนอโดย

นางสาวณัฐธิชา ทรงคำ รหัสนักศึกษา 6604105324

นางสาวธิดารัตน์ คิตจักร รหัสนักศึกษา 6604105326



..... อาจารย์ที่ปรึกษา

(อาจารย์ ดร.พศิน มรุพันธ์ธร)

ชื่อ-สกุล

นางสาวณัฐริชา ทรงคำ รหัสนักศึกษา 6604105324

นางสาวธิดารัตน์ คัดจักร รหัสนักศึกษา 6604105326

อาจารย์ที่ปรึกษา

อาจารย์ ดร.พศิน มรุตพันธุ์

หัวข้อโครงการ

การพัฒนาแบบจำลองเชิงปริมาณเพื่อสร้างผลตอบแทน
ที่เหนือกว่าตลาดด้วยกลยุทธ์แบบกองทุนป้องกันความเสี่ยง
Developing Quantitative Models to
Generate Alpha Using Hedge Fund Strategies

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาการสร้างและคัดกรองสูตรแอลฟา (*alpha formulas*) จากข้อมูลราคาและปริมาณการซื้อขายในรูปแบบ OHLCV ได้แก่ ราคาเปิด ราคาสูงสุด ราคาต่ำสุด ราคาปิด และปริมาณการซื้อขาย โดยใช้ข้อมูลของหลักทรัพย์ในดัชนี S&P 500 เป็นตัวแทนโดยประมาณของจักรวาลสินทรัพย์ในแพลตฟอร์ม WorldQuant BRAIN และครอบคลุมช่วงเวลาระหว่างปี ค.ศ. 2019 ถึง 2024 งานวิจัยได้พัฒนาสูตรแอลฟารวม 36 สูตร ครอบคลุมหลายหมวดหมู่ของแนวคิดเชิงการลงทุน เช่น โมเมนตัม การกลับเข้าสู่ค่าเฉลี่ย ความผันผวน ความสัมพันธ์ระหว่างราคาและปริมาณซื้อขาย และรูปแบบแท่งเทียน

ในด้านระเบียบวิธีวิจัย สูตรแอลฟาถูกนำไปประเมินเบื้องต้นจากความสามารถในการแยกความแตกต่างระหว่างหุ้นที่มีผลตอบแทนสูงและหุ้นที่มีผลตอบแทนต่ำ โดยอาศัยลักษณะการกระจายของค่าสัญญาณเป็นเกณฑ์ในการพิจารณา จากนั้นจึงคัดแยกสูตรออกเป็นสองกลุ่ม ได้แก่ กลุ่มสูตรที่สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ชัดเจน และกลุ่มสูตรที่แยกได้ไม่ชัดเจน ก่อนนำสูตรทั้งหมดไปทดสอบบนแพลตฟอร์ม WorldQuant BRAIN ภายใต้กรอบการจำลองพอร์ตแบบถือสถานะซื้อและขายชอร์ต พร้อมประเมินผลผ่านตัวชี้วัดสำคัญ ได้แก่ *Sharpe ratio*, *Turnover*, *Fitness*, *Returns*, *Drawdown* และ *Margin* นอกจากนี้ ยังใช้การทดสอบทางสถิติแบบ Welch's two-sample t-test เพื่อเปรียบเทียบค่าเฉลี่ยของตัวชี้วัดระหว่างสองกลุ่มสูตร

ผลการศึกษาพบว่า กลุ่มสูตรที่สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ชัดเจนมีผลการดำเนินงานเหนือกว่ากลุ่มที่แยกได้ไม่ชัดเจนอย่างมีนัยสำคัญทางสถิติในหลายมิติ โดยมีค่าเฉลี่ยของ *Fitness*, *Returns* และ *Margin* สูงกว่าอย่างมีนัยสำคัญที่ระดับ 0.01 และมีค่าเฉลี่ยของ *Turnover* ต่ำกว่าอย่างมีนัยสำคัญ ขณะที่ *Sharpe ratio* มีแนวโน้มสูงกว่าในระดับเกือบมีนัยสำคัญ และ *Drawdown* ไม่แตกต่างกันอย่างมีนัยสำคัญ ผลลัพธ์ดังกล่าวชี้ให้เห็นว่า ความสามารถในการจำแนกหุ้นในระดับข้อมูลสามารถใช้เป็นเกณฑ์คัดกรองเบื้องต้นสำหรับการพยากรณ์ประสิทธิภาพของสูตรแอลฟาในระดับพอร์ตการลงทุนได้

งานวิจัยนี้สนับสนุนแนวคิดว่า การวิเคราะห์คุณสมบัติของสัญญาณจากข้อมูล OHLCV ก่อนการทดสอบเต็มรูปแบบบนแพลตฟอร์มจริง เป็นแนวทางที่ช่วยเพิ่มประสิทธิภาพในการพัฒนาแอลฟา ลดภาระในการทดลองสูตรจำนวนมาก และช่วยเชื่อมโยงระหว่างการออกแบบสัญญาณเชิงปริมาณกับผลการดำเนินงานของพอร์ตในบริบทของการลงทุนเชิงระบบ

กิตติกรรมประกาศ

โครงการคณิตศาสตร์เรื่อง การพัฒนาแบบจำลองเชิงปริมาณเพื่อสร้างผลตอบแทนที่เหนือกว่าตลาดด้วยกลยุทธ์แบบกองทุนป้องกันความเสี่ยง สำเร็จลุล่วงได้ โดยได้รับความอนุเคราะห์เป็นอย่างดี และขอขอบพระคุณ อาจารย์ที่ปรึกษา อาจารย์ ดร.พศิน มรุพันธ์ธร ที่ได้ให้คำแนะนำ เสนอแนวทาง ตรวจสอบและแก้ไขข้อบกพร่องให้คณะผู้จัดทำโครงการโดยตลอด

ทางคณะผู้จัดทำโครงการหวังเป็นอย่างยิ่งว่า โครงการคณิตศาสตร์เรื่อง การพัฒนาแบบจำลองเชิงปริมาณเพื่อสร้างผลตอบแทนที่เหนือกว่าตลาดด้วยกลยุทธ์แบบกองทุนป้องกันความเสี่ยง ที่จัดทำขึ้น จะเป็นประโยชน์แก่ผู้สนใจที่จะนำไปศึกษาต่อ ในประเด็นที่ได้เสนอแนะไว้ในโครงการฉบับนี้ ให้บรรลุผลอย่างมีประสิทธิภาพต่อไป

Contents

1	บทนำ	1
1.1	ที่มาและความสำคัญของปัญหา	1
1.2	วัตถุประสงค์ของการศึกษา	2
1.3	ประโยชน์ที่คาดว่าจะได้รับ	3
1.4	ขอบเขตในการศึกษา	3
2	เอกสารที่เกี่ยวข้อง	4
2.1	แบบจำลองการกำหนดราคาสินทรัพย์ทุน (CAPM)	4
2.2	งานวิจัยเกี่ยวกับสัญญาณจากราคาและปริมาณซื้อขาย	5
2.3	กลยุทธ์การลงทุนแบบถือสถานะซื้อและสถานะขายชอร์ต (Long–Short)	5
2.4	ความสอดคล้องเชิงทฤษฎีของกลยุทธ์ Long–Short กับ CAPM	6
3	วิธีการใช้งานแพลตฟอร์ม WorldQuant BRAIN	6
4	แบบจำลองเชิงคณิตศาสตร์สำหรับกระบวนการลงทุนของ WorldQuant BRAIN	8
4.1	กระบวนการลงทุน	8
4.1.1	โครงสร้างข้อมูลตั้งต้น: Data Matrix	8
4.1.2	การแปลงข้อมูลเป็นสัญญาณ: Transformation	9
4.1.3	การจัดอันดับแบบภาคตัดขวาง: Ranking	9
4.1.4	การทำให้สัญญาณเป็นกลาง: Neutralization	10
4.1.5	การทำให้สัญญาณพร้อมใช้งานในพอร์ต: Normalization	10
4.1.6	การจัดสรรเงินลงทุนจริง: Capital Allocation	11
4.2	ตัวอย่างเชิงตัวเลขสำหรับพอร์ตการลงทุนขนาดเล็ก	12
4.3	การวัดผลการลงทุน	15
5	ตัวอย่างการสร้างสูตรแอลฟาจากสมมติฐาน	16
5.1	สัญญาณกลับทิศของการเปลี่ยนแปลงราคาแบบมีเงื่อนไขตามการเปลี่ยนแปลงของปริมาณซื้อขาย	17
5.2	สัญญาณความแข็งแรงของแท่งเทียนภายในช่วงวัน (Normalized Intraday Candle Strength)	19
5.3	สัญญาณกลับทิศของความสัมพันธ์ราคาเปิดกับปริมาณซื้อขายระยะสั้น (Contrarian Open–Volume Correlation, 10 วัน)	20
5.4	สัญญาณอันดับของส่วนต่างราคาเปิด–ปิดแบบกลับทิศ (Ranked Open–Close Deviation, Contrarian)	22

6	วิธีการดำเนินงาน	24
6.1	รวบรวมข้อมูล OHLCV	24
6.2	สร้างสูตรแอลฟา	25
6.3	ทดสอบความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำของสูตรแอลฟา	27
6.4	ทดสอบบน WorldQuant Brain Platform	28
6.5	ทดสอบความแตกต่างของค่าเฉลี่ยของตัววัดผล	29
7	ผลการดำเนินงาน	30
7.1	สูตรแอลฟาในทดลอง	30
7.1.1	สัญญาณกลับทิศของผลตอบแทนภายในวัน (Contrarian Intraday Return Signal)	30
7.1.2	สัญญาณความยาวไส้เทียนล่างแบบปรับขนาดด้วยราคาปิด (Normalized Lower Shadow Signal)	33
7.2	การวิเคราะห์แผนภาพกล่องของค่าสัญญาณแอลฟา	35
7.3	ผลการทดสอบทางสถิติ	37
8	สรุปและอภิปราย	39

1 บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในปัจจุบัน การลงทุนเชิงปริมาณ (*quantitative investment*) ได้กลายเป็นแนวทางสำคัญของอุตสาหกรรมการเงินสมัยใหม่ โดยเฉพาะในกองทุนประเภท *hedge fund* และสถาบันการเงินที่ใช้การตัดสินใจเชิงระบบแทนการพึ่งพาดุลยพินิจของมนุษย์ แนวทางดังกล่าวอาศัยการประยุกต์ใช้คณิตศาสตร์ สถิติ วิทยาการข้อมูล และการประมวลผลข้อมูลขนาดใหญ่ เพื่อสร้างกฎการลงทุนที่สามารถตรวจสอบ ทดสอบ และทำซ้ำได้ในเชิงประจักษ์ แทนที่จะอาศัยเพียงการคาดการณ์จากประสบการณ์หรือความเห็นส่วนบุคคล ทั้งนี้ แกนกลางของการลงทุนเชิงปริมาณคือการค้นหา *alpha* หรือสัญญาณการลงทุนที่สามารถใช้จำแนกหลักทรัพย์ซึ่งมีแนวโน้มให้ผลตอบแทนแตกต่างกัน และแปลงสัญญาณดังกล่าวให้เป็นพอร์ตลงทุนที่มีผลการดำเนินงานดีกว่ามาตรฐานภายใต้ระดับความเสี่ยงที่ยอมรับได้ [1]

ในบริบทนี้ แพลตฟอร์ม WorldQuant BRAIN เป็นหนึ่งในตัวอย่างสำคัญของโครงสร้างพื้นฐานสำหรับการวิจัยแอลฟาเชิงระบบ เนื่องจากเปิดโอกาสให้นักวิจัยสามารถเขียนนิพจน์ของแอลฟา ทดสอบย้อนหลังบนกลุ่มสินทรัพย์ภายใต้กรอบมาตรฐานเดียวกัน และประเมินผลผ่านตัวชี้วัดด้านผลตอบแทน ความเสี่ยง และประสิทธิภาพของการซื้อขาย การมีแพลตฟอร์มลักษณะนี้ทำให้การวิจัยเชิงปริมาณไม่ได้หยุดอยู่ที่ระดับการเสนอสมมติฐานทางทฤษฎี แต่สามารถก้าวไปสู่การตรวจสอบเชิงประจักษ์ได้อย่างเป็นระบบ อย่างไรก็ตาม แม้จะมีเครื่องมือสำหรับทดสอบแอลฟาอยู่แล้ว คำถามสำคัญเชิงวิจัยยังคงอยู่ที่ว่า *เราควรคัดเลือกหรือประเมินคุณภาพของสูตรแอลฟาอย่างไรก่อนนำไปทดสอบในระดับพอร์ตจริง*

โจทย์ดังกล่าวมีความสำคัญอย่างยิ่ง เพราะในทางปฏิบัติ สูตรแอลฟาสามารถถูกสร้างขึ้นได้เป็นจำนวนมากจากข้อมูลตลาดพื้นฐานเพียงไม่กี่ชนิด โดยเฉพาะข้อมูล OHLCV ซึ่งประกอบด้วยราคาเปิด ราคาสูงสุด ราคาต่ำสุด ราคาปิด และปริมาณการซื้อขาย แม้ข้อมูลชุดนี้จะเป็นข้อมูลพื้นฐานที่สุดของตลาด แต่กลับสามารถต่อยอดเป็นตัวแปรเชิงเทคนิคได้หลากหลาย เช่น ผลตอบแทนย้อนหลัง โมเมนตัม การกลับเข้าสู่ค่าเฉลี่ย ความผันผวน ช่วงการเคลื่อนไหวของราคา ความสัมพันธ์ระหว่างราคาและปริมาณซื้อขาย ตลอดจนรูปแบบแท่งเทียน ด้วยเหตุนี้ ความท้าทายของงานวิจัยจึงไม่ใช่เพียงการสร้างสูตรให้ได้จำนวนมาก แต่เป็นการตอบคำถามว่า สูตรประเภทใดมี “สาระเชิงข้อมูล” มากพอที่จะนำไปสู่ผลการดำเนินงานที่ดีเมื่อทดสอบจริงในแพลตฟอร์มการลงทุนเชิงระบบ

การศึกษานี้จึงตั้งอยู่บนแนวคิดสำคัญว่า สูตรแอลฟาที่ดีควรมีความสามารถในการแยกความแตกต่างระหว่างหุ้นที่ให้ผลตอบแทนสูงและหุ้นที่ให้ผลตอบแทนต่ำได้อย่างมีแบบแผน หากค่าของแอลฟากระจายต่างกันอย่างชัดเจนระหว่างสองกลุ่มดังกล่าว ก็อาจสะท้อนว่าสูตรนั้นมีพลังในการจำแนก (*discriminative power*) และมีโอกาสสูงที่จะนำไปสู่การจัดอันดับหุ้นและการสร้างพอร์ตที่มีประสิทธิภาพ ในทางกลับกัน หากค่าของแอลฟาระหว่างหุ้นสองกลุ่มมีการกระจายที่ทับซ้อนกันมาก สัญญาณดังกล่าวอาจไม่มีอำนาจในการจำแนกเพียงพอ และเมื่อนำไปสร้างพอร์ตจริงก็อาจไม่สามารถให้ผลลัพธ์ที่โดดเด่นได้ ดังนั้น งานวิจัยนี้จึงไม่ได้มอง “แอลฟา” เป็นเพียงสมการทางคณิตศาสตร์ แต่พิจารณาในฐานะเครื่องมือสำหรับคัดแยกข้อมูลเชิงผลตอบแทนของหลักทรัพย์อย่างเป็นระบบ

บนพื้นฐานของแนวคิดดังกล่าว งานวิจัยนี้ใช้ข้อมูล OHLCV ของหลักทรัพย์ในดัชนี S&P 500 เป็นตัวแทนโดยประมาณของกลุ่มสินทรัพย์ใน WorldQuant BRAIN และครอบคลุมช่วงเวลาระหว่างปี ค.ศ. 2019 ถึง 2024 เพื่อให้ได้ข้อมูลที่สะท้อนทั้งภาวะตลาดปกติ ช่วงความผันผวนสูง และช่วงฟื้นตัวของตลาด การกำหนดกรอบข้อมูลเช่น

นี้มีความสำคัญเชิงวิธีวิทยา เพราะช่วยให้ผลการทดลองไม่ยึดติดกับสภาพตลาดเพียงระยะใดระยะหนึ่ง และทำให้การประเมินความแข็งแกร่งของสูตรแอลฟามีความสมจริงมากขึ้น

ภายใต้ชุดข้อมูลดังกล่าว ผู้วิจัยได้พัฒนาสูตรแอลฟาจำนวน 36 สูตรจากข้อมูล OHLCV และจัดกลุ่มเป็นหมวดหมู่ตามลักษณะของแนวคิดที่สูตรต้องการสะท้อน ได้แก่ กลุ่มผลตอบแทนย้อนหลังและโมเมนตัม กลุ่มการกลับเข้าสู่ค่าเฉลี่ยและค่าเฉลี่ยเคลื่อนที่ กลุ่มความผันผวนและช่วงการเคลื่อนไหวของราคา กลุ่มความสัมพันธ์ระหว่างราคาและปริมาณซื้อขาย และกลุ่มรูปแบบแท่งเทียนกับการทะลุรอบราคา โครงสร้างดังกล่าวทำให้งานวิจัยสามารถเปรียบเทียบได้ทั้งในระดับสูตรรายตัว และระดับหมวดหมู่ของแนวคิดเชิงการลงทุน ว่ากลุ่มใดมีแนวโน้มสร้างสัญญาณที่มีคุณภาพมากกว่า

ขั้นตอนสำคัญของการศึกษาคือการแบ่งหุ้นออกเป็นกลุ่มผลตอบแทนสูงและกลุ่มผลตอบแทนต่ำ แล้วตรวจสอบว่าค่าของแอลฟาแต่ละสูตรมีการกระจายแตกต่างกันมากน้อยเพียงใดระหว่างสองกลุ่มนี้ โดยใช้การมองผ่านลักษณะของการกระจายข้อมูลในลักษณะแผนภาพกล่องเป็นเครื่องมือเบื้องต้นในการคัดกรองแนวคิด งานวิจัยนี้จึงให้ความสำคัญกับความแตกต่างของระดับช่วงตัว (*spread*) ของการกระจายค่าระหว่างสองกลุ่มหุ้น กล่าวคือ หากสูตรแอลฟาหนึ่งทำให้กลุ่มหุ้นผลตอบแทนสูงและต่ำแสดงรูปแบบการกระจายที่ต่างกันอย่างชัดเจน ก็อาจตีความได้ว่าสูตรนั้นมีความสามารถในการจำแนกที่ดีกว่า งานวิจัยจึงนำแนวคิดนี้ไปสู่สมมติฐานหลักว่า *สูตรแอลฟาที่สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ชัดเจนกว่า ควรมีแนวโน้มให้ผลการทดสอบบน WorldQuant BRAIN ดีกว่าสูตรที่แยกได้ไม่ชัดเจน*

ความสำคัญของโจทย์นี้ไม่ได้อยู่เพียงในเชิงเทคนิคของการสร้างสัญญาณ แต่ยังมีความหมายในเชิงระเบียบวิธีด้วย เพราะหากสามารถแสดงให้เห็นได้ว่าลักษณะการกระจายของค่าพีเจอร์หรือแอลฟาในระดับข้อมูลมีความสัมพันธ์กับผลการดำเนินงานของพอร์ตในแพลตฟอร์มจริง ก็จะได้กรอบคิดสำหรับการคัดกรองสูตรแอลฟาเบื้องต้นก่อนเข้าสู่ขั้นตอนการทดสอบเต็มรูปแบบบนแพลตฟอร์ม การมีขั้นตอนคัดกรองเช่นนี้ช่วยลดต้นทุนในการทดลองลดการพึ่งพาการลองผิดลองถูกแบบสุ่ม และช่วยให้การสร้างแอลฟาเป็นกระบวนการที่มีตรรกะเชิงวิทยาศาสตร์มากขึ้น

นอกจากนี้ ในเชิงผลลัพธ์ งานวิจัยยังประเมินประสิทธิภาพของสูตรแอลฟาบน WorldQuant BRAIN ผ่านตัวชี้วัดหลัก ได้แก่ *Sharpe ratio*, *Turnover*, *Fitness*, *Returns*, *Drawdown* และ *Margin* และนำผลของสองกลุ่มสูตรมาเปรียบเทียบกันด้วยการทดสอบ Welch's two-sample t-test เพื่อประเมินว่าความแตกต่างของค่าเฉลี่ยในแต่ละตัวชี้วัดมีนัยสำคัญทางสถิติหรือไม่ การออกแบบเช่นนี้ช่วยให้งานวิจัยไม่เพียงตรวจสอบในระดับคำอธิบายเชิงคุณภาพว่ากลุ่มสูตรใด “ดูเหมือน” ดีกว่า แต่ยังสามารถยกระดับข้อสรุปไปสู่การอนุมานเชิงสถิติได้อีกด้วย

ดังนั้น งานวิจัยนี้จึงมีความสำคัญในสองระดับพร้อมกัน ระดับแรกคือการพัฒนาองค์ความรู้เกี่ยวกับการสร้างและคัดกรองสูตรแอลฟาจากข้อมูล OHLCV ซึ่งเป็นข้อมูลพื้นฐานแต่มีศักยภาพสูงในการสร้างสัญญาณการลงทุน ระดับที่สองคือการเชื่อมโยงระหว่างคุณสมบัติของสัญญาณในระดับข้อมูลกับผลการดำเนินงานจริงของพอร์ตบนแพลตฟอร์ม WorldQuant BRAIN ซึ่งเป็นประเด็นที่มีคุณค่าทั้งในเชิงวิชาการและในเชิงปฏิบัติสำหรับผู้สนใจการวิจัยและพัฒนากลยุทธ์การลงทุนเชิงปริมาณ

1.2 วัตถุประสงค์ของการศึกษา

การศึกษานี้มีวัตถุประสงค์ดังต่อไปนี้

1. เพื่อพัฒนาสูตรแอลฟาจากข้อมูล OHLCV ของหลักทรัพย์ในดัชนี S&P 500 ซึ่งใช้เป็นตัวแทนโดยประมาณ

ของกลุ่มสินทรัพย์ในแพลตฟอร์ม WorldQuant BRAIN โดยครอบคลุมหลายหมวดหมู่ของแนวคิดเชิงการลงทุน เช่น โมเมนตัม การกลับเข้าสู่ค่าเฉลี่ย ความผันผวน ความสัมพันธ์ระหว่างราคาและปริมาณซื้อขาย และรูปแบบแท่งเทียน พร้อมทั้งศึกษาความสามารถของสูตรแอลฟาแต่ละกลุ่มในการแยกความแตกต่างระหว่างหุ้นผลตอบแทนสูงและหุ้นผลตอบแทนต่ำ

2. เพื่อทดสอบสูตรแอลฟาที่สร้างขึ้นบนแพลตฟอร์ม WorldQuant BRAIN และประเมินผลผ่านตัวชี้วัดมาตรฐาน รวมทั้งเปรียบเทียบผลการดำเนินงานเฉลี่ยระหว่างกลุ่มสูตรที่มีความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำได้ชัดเจนกับกลุ่มสูตรที่แยกได้ไม่ชัดเจน โดยใช้การทดสอบทางสถิติแบบ Welch's two-sample t-test เพื่อประเมินว่าคุณสมบัติของสูตรแอลฟาในระดับข้อมูลสามารถใช้เป็นเกณฑ์คัดกรองเบื้องต้นสำหรับการพยากรณ์ประสิทธิภาพของสูตรในระดับพอร์ตการลงทุนได้หรือไม่

1.3 ประโยชน์ที่คาดว่าจะได้รับ

การศึกษาครั้งนี้คาดว่าจะก่อให้เกิดประโยชน์ดังต่อไปนี้

1. ได้กรอบวิธีการในการสร้างและคัดกรองสูตรแอลฟาจากข้อมูล OHLCV อย่างเป็นระบบ โดยเชื่อมโยงข้อมูลตลาดพื้นฐานเข้ากับแนวคิดการลงทุนเชิงปริมาณ และใช้ความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำเป็นเกณฑ์เบื้องต้นในการประเมินคุณภาพของสูตรก่อนนำไปทดสอบบนแพลตฟอร์มจริง
2. ได้ข้อค้นพบเชิงประจักษ์เกี่ยวกับความสัมพันธ์ระหว่างคุณสมบัติของสัญญาณในระดับข้อมูลกับผลการดำเนินงานของพอร์ตบนแพลตฟอร์ม WorldQuant BRAIN ซึ่งช่วยเสริมองค์ความรู้เชิงประยุกต์สำหรับการวิจัยแอลฟา การพัฒนากลยุทธ์การลงทุนเชิงปริมาณ และการต่อยอดกระบวนการวิจัยที่ผสมผสานการวิเคราะห์ข้อมูล การทดสอบเชิงสถิติ และการประเมินผลเชิงการเงินเข้าด้วยกันอย่างเป็นระบบ

1.4 ขอบเขตในการศึกษา

การศึกษาครั้งนี้กำหนดขอบเขตไว้ดังต่อไปนี้

1. ใช้ข้อมูล OHLCV ของหลักทรัพย์ในดัชนี S&P 500 เป็นตัวแทนโดยประมาณของกลุ่มสินทรัพย์ในแพลตฟอร์ม WorldQuant BRAIN
2. ใช้ข้อมูลในช่วงปี ค.ศ. 2019 ถึง 2024 เป็นช่วงเวลาสำหรับการสร้างตัวแปร การทดสอบ และการประเมินผลของสูตรแอลฟา
3. มุ่งเน้นการสร้างสูตรแอลฟาจากข้อมูลราคาและปริมาณซื้อขายเท่านั้น โดยไม่รวมข้อมูลปัจจัยพื้นฐาน ข้อมูลมหภาค หรือข้อมูลทางเลือกประเภทอื่น
4. พิจารณาสูตรแอลฟาจำนวน 36 สูตร ซึ่งจัดออกเป็น 4 หมวดหมู่หลักตามลักษณะของแนวคิดเชิงเทคนิค และพฤติกรรมราคา
5. ใช้ความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำเป็นเกณฑ์คัดกรองเชิงโครงสร้างในระดับข้อมูล ก่อนนำสูตรไปทดสอบบนแพลตฟอร์ม WorldQuant BRAIN

6. ประเมินผลการดำเนินงานของสูตรผ่านตัวชี้วัด *Sharpe ratio*, *Turnover*, *Fitness*, *Returns*, *Drawdown* และ *Margin* และเปรียบเทียบค่าเฉลี่ยของสองกลุ่มสูตรด้วยการทดสอบ Welch's two-sample t-test

2 เอกสารที่เกี่ยวข้อง

2.1 แบบจำลองการกำหนดราคาสินทรัพย์ทุน (CAPM)

Capital Asset Pricing Model (CAPM) เป็นแบบจำลองเชิงดุลยภาพที่อธิบายความสัมพันธ์ระหว่างผลตอบแทนคาดหวังของสินทรัพย์กับความเสี่ยงเชิงระบบ (*systematic risk*) โดยมีรากฐานสำคัญจากทฤษฎีพอร์ตของ Markowitz ซึ่งเสนอว่าผู้ลงทุนควรพิจารณาผลตอบแทนคาดหวังควบคู่กับความแปรปรวนของพอร์ต และเลือกพอร์ตที่อยู่บนเส้นพรมแดนประสิทธิภาพ (*efficient frontier*) [2] ภายใต้สมมติฐานเรื่องตลาดทุนที่แข่งขันสมบูรณ์ ผู้ลงทุนสามารถกู้ยืมและให้กู้ที่อัตราปลอดความเสี่ยงได้ และมีความคาดหวังต่อผลตอบแทนและความเสี่ยงเหมือนกัน Sharpe, Lintner และ Mossin ได้พัฒนาแบบจำลอง CAPM เพื่อแสดงว่า ในดุลยภาพ ผลตอบแทนคาดหวังของสินทรัพย์จะขึ้นอยู่กับเฉพาะความเสี่ยงที่สัมพันธ์กับพอร์ตตลาดเท่านั้น [3, 4, 5, 6]

ในเชิงคณิตศาสตร์ CAPM เขียนได้เป็น

$$\mathbb{E}(R_i) = R_f + \beta_i (\mathbb{E}(R_m) - R_f), \quad \beta_i = \frac{\text{Cov}(R_i, R_m)}{\text{Var}(R_m)}, \quad (2.1)$$

โดยที่ $\mathbb{E}(R_i)$ คือผลตอบแทนคาดหวังของสินทรัพย์ i , R_f คืออัตราผลตอบแทนปราศจากความเสี่ยง, และ $\mathbb{E}(R_m)$ คือผลตอบแทนคาดหวังของพอร์ตตลาด สมการนี้แสดงว่า “ราคาของความเสี่ยง” ถูกกำหนดผ่าน *market risk premium* และสินทรัพย์แต่ละตัวได้รับผลตอบแทนชดเชยตามระดับของ β_i ซึ่งวัดการรวมเคลื่อนไหวกับตลาด [3, 7]

สาระสำคัญของ CAPM อยู่ที่การแยกความเสี่ยงออกเป็นสองส่วน ได้แก่ ความเสี่ยงเชิงระบบ และความเสี่ยงเฉพาะตัว (*idiosyncratic risk*) ซึ่งสามารถลดได้ด้วยการกระจายการลงทุน หากให้พอร์ตมีน้ำหนัก w และเวกเตอร์ผลตอบแทนเป็น R จะได้ว่า

$$R_p = w^\top R, \quad \text{Var}(R_p) = w^\top \Sigma w, \quad (2.2)$$

ซึ่งแสดงให้เห็นว่าความเสี่ยงของพอร์ตขึ้นกับโครงสร้างโควาเรียนซ์ระหว่างสินทรัพย์ ไม่ใช่เพียงความแปรปรวนรายตัวเท่านั้น [2] เมื่อพอร์ตมีการกระจายความเสี่ยงเพียงพอ ความเสี่ยงเฉพาะตัวจะถูกเฉลี่ยออกไป และตลาดจะชดเชยผลตอบแทนให้เฉพาะความเสี่ยงที่ไม่สามารถกระจายได้

ในเชิงประจักษ์ CAPM ยังเป็นพื้นฐานของการนิยาม “อัลฟา” ในการประเมินผลการลงทุน โดย Jensen เสนอว่า ผลตอบแทนส่วนเกินที่เกินกว่าระดับที่ควรได้ตาม CAPM สามารถใช้วัดความสามารถของผู้จัดการกองทุนหรือกลยุทธ์การลงทุนได้ [8] ซึ่งสามารถเขียนได้เป็น

$$\alpha_i = R_i - [R_f + \beta_i (R_m - R_f)]. \quad (2.3)$$

แนวคิดนี้มีความสำคัญอย่างยิ่งต่อการวิจัยแอลฟา เพราะทำให้ “ผลตอบแทนส่วนเกิน” ถูกตีความว่าเป็นผลตอบแทนที่ไม่สามารถอธิบายได้ด้วยความเสี่ยงตลาดเพียงอย่างเดียว

อย่างไรก็ดี งานวิจัยภายหลังพบว่าเบต้าเพียงตัวเดียวอาจไม่เพียงพอในการอธิบายความแตกต่างของผลตอบแทนข้ามหลักทรัพย์ Fama และ French เสนอแบบจำลองหลายปัจจัยซึ่งเพิ่มปัจจัยด้านขนาดบริษัทและมูลค่า

ทางบัญชีต่อมูลค่าตลาดเข้าไป [9] และ Carhart ขยายต่อด้วยปัจจัยโมเมนตัม [10] ดังนั้น ในบริบทของการวิจัยเชิงปริมาณสมัยใหม่ CAPM จึงยังคงมีบทบาทเป็นกรอบพื้นฐานในการคิดเรื่องผลตอบแทนที่สอดคล้องกับความเสี่ยง ขณะเดียวกันก็เป็นจุดตั้งต้นของแบบจำลองหลายปัจจัยที่ใช้ประเมินความสามารถของกลยุทธ์ long-short และสัญญาณแอลฟาในงานเชิงประจักษ์

2.2 งานวิจัยเกี่ยวกับสัญญาณจากราคาและปริมาณซื้อขาย

เนื่องจากการศึกษานี้สร้างสูตรแอลฟาจากข้อมูล OHLCV งานวิจัยที่ใช้ข้อมูลราคาและปริมาณซื้อขายโดยตรงจึงมีความเกี่ยวข้องอย่างมาก งานคลาสสิกของ Brock, Lakonishok และ LeBaron แสดงให้เห็นว่ากฎการซื้อขายเชิงเทคนิคอย่างเส้นค่าเฉลี่ยเคลื่อนที่และการทะลุรอบราคาให้ผลลัพธ์ที่แตกต่างจากสมมติฐานสุ่มเดินอย่างมีนัยสำคัญ [11] ต่อมา Lo, Mamaysky และ Wang ได้พยายามทำให้การวิเคราะห์ทางเทคนิคมีฐานเชิงวิชาการมากขึ้น ผ่านการนิยามรูปแบบกราฟราคาและขั้นตอนการอนุมาณเชิงสถิติอย่างเป็นระบบ [12] งานเหล่านี้สนับสนุนแนวคิดว่าข้อมูลราคาในอดีตอาจมีโครงสร้างบางอย่างที่สามารถแปลงเป็นสัญญาณเชิงปริมาณได้

นอกจากงานด้านเทคนิคโดยตรงแล้ว งานวิจัยด้านโมเมนตัมและการกลับตัวของราคายังมีความเกี่ยวข้องอย่างใกล้ชิด Jegadeesh และ Titman แสดงว่ากลยุทธ์ซื้อหุ้นผู้ชนะและขายหุ้นผู้แพ้ในอดีตให้ผลตอบแทนส่วนเกินในเชิงข้ามหลักทรัพย์ (*cross-sectional momentum*) [13] ขณะที่ Lo และ MacKinlay ชี้ว่ากำไรจากกลยุทธ์ contrarian บางส่วนอาจเกิดจากโครงสร้าง cross-autocovariance ของผลตอบแทน ไม่ได้เกิดจากการ overreaction เพียงอย่างเดียว [14] งานเหล่านี้มีนัยสำคัญต่อการสร้างสูตรแอลฟาในงานวิจัยนี้ เพราะหลายสูตรอาศัยแนวคิดเรื่องผลตอบแทนย้อนหลัง การสะสมผลตอบแทน และการเปลี่ยนแปลงของความผันผวนในช่วงเวลาต่าง ๆ

ในระดับที่กว้างขึ้น Asness, Moskowitz และ Pedersen แสดงให้เห็นว่าปัจจัย value และ momentum ให้ผลตอบแทนส่วนเกินอย่างสม่ำเสมอในหลายตลาดและหลายสินทรัพย์ [15] ขณะที่ Moskowitz, Ooi และ Pedersen แสดงหลักฐานของ *time-series momentum* ในสินทรัพย์หลายประเภท [16] งานเหล่านี้ช่วยยืนยันว่าการใช้สัญญาณที่มาจากอนุกรมเวลาของราคาและข้อมูลตลาดไม่ใช่เพียงเทคนิคเฉพาะกรณี แต่เป็นส่วนหนึ่งของงานวิจัยกว้างด้าน factor investing และ systematic trading

2.3 กลยุทธ์การลงทุนแบบถือสถานะซื้อและสถานะขายชอร์ต (Long-Short)

Long-Short Strategy เป็นกลยุทธ์การลงทุนเชิงรุกที่อนุญาตให้นักลงทุนถือสถานะซื้อ (Long position) และสถานะขายชอร์ต (Short position) พร้อมกันภายในพอร์ตเดียวกัน โดยมีเป้าหมายเพื่อสร้างผลตอบแทนจากส่วนต่างของผลการดำเนินงานระหว่างสินทรัพย์ทั้งสองฝั่ง มากกว่าการพึ่งพาทิศทางของตลาดโดยรวม [17, 18] ในเชิงคณิตศาสตร์ หากพอร์ตมีน้ำหนัก w_i กับสินทรัพย์ตัวที่ i จะได้ผลตอบแทนพอร์ตเป็น

$$R_p = \sum_{i=1}^n w_i R_i, \quad (2.4)$$

โดย $w_i > 0$ หมายถึงสถานะซื้อ และ $w_i < 0$ หมายถึงสถานะขายชอร์ต โครงสร้างดังกล่าวเปิดโอกาสให้สร้างพอร์ตแบบ self-financing ซึ่งเงินที่ได้จากฝั่ง short ถูกนำไปสนับสนุนฝั่ง long ได้

ในเชิงทฤษฎี กลยุทธ์ long-short มักถูกใช้เพื่อพยายามสกัดผลตอบแทนส่วนเกินที่ไม่อธิบายด้วยปัจจัยตลาด กล่าวคือ หากเลือกสินทรัพย์ฝั่ง long ที่คาดว่าจะให้ผลตอบแทนสูงกว่า และเลือกสินทรัพย์ฝั่ง short ที่คาดว่าจะให้ผลตอบแทนต่ำกว่า ผลตอบแทนของพอร์ตจะขึ้นกับ “ส่วนต่างเชิงสัมพัทธ์” ของสองฝั่งมากกว่าการขึ้นหรือ

ลงของตลาดโดยรวม [19, 20] แนวคิดนี้สอดคล้องอย่างยิ่งกับการสร้างแอลฟาในแพลตฟอร์มเชิงปริมาณ เช่น WorldQuant BRAIN ซึ่งแปลงสัญญาณเชิงภาคตัดขวางให้เป็นน้ำหนักพอร์ตแบบ long-short

งานวิจัยเชิงประจักษ์จำนวนมากสนับสนุนการใช้พอร์ต long-short ในฐานะเครื่องมือสกัดปัจจัยหรือความผิดพลาดในการกำหนดราคา ตัวอย่างสำคัญคือกลยุทธ์ momentum แบบซื้อผู้ขึ้นและขายผู้แพ้ [13] กลยุทธ์ pairs trading ซึ่งเป็นการ long และ short หุ้นที่มีความคล้ายคลึงกันในอดีต [21] และปัจจัย Betting Against Beta (BAB) ซึ่งสร้างพอร์ต long หุ้นเบต้าต่ำและ short หุ้นเบต้าสูง [22] งานวิจัยเหล่านี้แสดงให้เห็นว่า long-short ไม่ใช่เพียงโครงสร้างพอร์ตทางเทคนิค แต่เป็นแกนกลางของงานวิจัยปัจจัยและกลยุทธ์เชิงระบบในสมัยใหม่

อย่างไรก็ดี ในทางปฏิบัติกลยุทธ์ long-short ยังเผชิญข้อจำกัดสำคัญ เช่น ต้นทุนการซื้อขาย ต้นทุนการยืมหุ้น เพื่อ short ข้อจำกัดด้านสภาพคล่อง และข้อกำหนดด้านหลักประกัน ซึ่งอาจทำให้ผลตอบแทนที่เห็นในเชิงทฤษฎีลดลงเมื่อใช้งานจริง [17, 21] ด้วยเหตุนี้ งานวิจัยที่ประเมิน long-short alpha จึงมักต้องพิจารณาตัวชี้วัดอย่าง turnover, drawdown และ margin ควบคู่กับผลตอบแทนด้วย

2.4 ความสอดคล้องเชิงทฤษฎีของกลยุทธ์ Long-Short กับ CAPM

Long-Short Strategy มีความสอดคล้องกับ CAPM ในระดับพื้นฐาน เพราะทั้งสองต่างตั้งอยู่บนแนวคิดเรื่องการแยกความเสี่ยงเชิงระบบออกจากความเสี่ยงเฉพาะตัว หากนิยามผลตอบแทนพอร์ตในรูป

$$R_p = \alpha_p + \beta_p(R_m - R_f) + \varepsilon_p, \quad (2.5)$$

แล้วออกแบบน้ำหนักของพอร์ตให้มี $\beta_p \approx 0$ พอร์ตดังกล่าวจะมีลักษณะเป็น *market-neutral portfolio* ซึ่งผลตอบแทนจะพึ่งพา α_p เป็นหลักมากกว่าการเคลื่อนไหวของตลาด [3, 8] ในเชิงตีความ long-short จึงเป็นกลไกเชิงปฏิบัติที่พยายามกำจัด market exposure ออกจากพอร์ตเพื่อให้เหลือเฉพาะส่วนที่สะท้อนความสามารถในการคัดเลือกสินทรัพย์

อย่างไรก็ตาม งานวิจัยสมัยใหม่ชี้ว่าการทำให้พอร์ต “เป็นกลางต่อตลาด” เพียงอย่างเดียวอาจยังไม่เพียงพอ หากผลตอบแทนของกลยุทธ์ยังรับความเสี่ยงจากปัจจัยระบบอื่น เช่น ขนาด มูลค่า โมเมนตัม หรือความผันผวนต่ำ [9, 10, 22, 15] ดังนั้น ในเชิงทฤษฎี ความสอดคล้องของ long-short กับ CAPM ควรถูกมองเป็นกรอบตั้งต้น ขณะที่ในเชิงประจักษ์ การประเมินประสิทธิภาพของกลยุทธ์มักต้องอาศัยกรอบหลายปัจจัยร่วมด้วยเพื่อแยกได้ว่าผลตอบแทนของพอร์ตเกิดจาก alpha จริง หรือเป็นเพียงการรับ factor exposure ที่เป็นที่รู้จักอยู่แล้ว

ในบริบทของงานวิจัยนี้ การสร้างสูตรแอลฟาจากข้อมูล OHLCV แล้วนำไปใช้ในพอร์ต long-short จึงสามารถอธิบายได้ในสองระดับพร้อมกัน ระดับแรกคือระดับทฤษฎี ซึ่งอิงกับ CAPM และแนวคิด market neutrality ในการพยายามกำจัดความเสี่ยงตลาดออกจากพอร์ต ระดับที่สองคือระดับเชิงประจักษ์ ซึ่งอาศัยงานวิจัยด้าน momentum, reversal, volatility, price-volume relation และ technical analysis เพื่อสนับสนุนว่าข้อมูลราคาและปริมาณซื้อขายอาจมีสาระเชิงพยากรณ์ที่สามารถแปลงเป็นสัญญาณแอลฟาได้ [11, 12, 13, 14]

3 วิธีการใช้งานแพลตฟอร์ม WorldQuant BRAIN

แพลตฟอร์ม WorldQuant BRAIN สามารถมองได้ว่าเป็นสภาพแวดล้อมสำหรับการวิจัยเชิงปริมาณที่ออกแบบมาเพื่อให้ผู้ใช้พัฒนา ทดสอบ และประเมิน *alpha* อย่างเป็นระบบ โดยการใช้งานแพลตฟอร์มนี้ไม่ได้จำกัดอยู่เพียงการเขียนสมการหรือสร้างสัญญาณการลงทุนเท่านั้น แต่ครอบคลุมตั้งแต่การเรียนรู้แนวคิดพื้นฐานของระบบ การ

กำหนดสมมติฐานเชิงปริมาณ การสร้างนิพจน์ของ *alpha* การจำลองผลการลงทุนย้อนหลัง ตลอดจนการตีความผลลัพธ์และปรับปรุงแบบจำลองอย่างต่อเนื่อง กล่าวได้ว่า BRAIN ทำหน้าที่เป็นทั้งเครื่องมือในการสร้างแบบจำลองและการอภิปรายที่ช่วยให้ผู้ใช้เปลี่ยนแนวคิดเชิงเศรษฐศาสตร์หรือสถิติให้กลายเป็นกลยุทธ์ที่ตรวจสอบได้เชิงประจักษ์

การเริ่มต้นใช้งานแพลตฟอร์มมักเริ่มจากการสมัครบัญชีและเข้าสู่ระบบวิจัยของ BRAIN เมื่อเข้าสู่แพลตฟอร์มแล้ว ผู้ใช้ควรทำความเข้าใจก่อนว่าจุดมุ่งหมายหลักของระบบนี้คือการพัฒนาและทดสอบแนวคิดการลงทุนเชิงระบบ มิใช่การส่งคำสั่งซื้อขายจริงในตลาดการเงินโดยตรง ดังนั้นการใช้งาน BRAIN จึงควรถูกมองในฐานะกระบวนการวิจัยมากกว่าการเก็งกำไรแบบตูลายพนิง กล่าวคือ ผู้ใช้เสนอสมมติฐานในรูปของนิพจน์เชิงคณิตศาสตร์ ระบบจะแปลงนิพจน์ดังกล่าวให้เป็นเวกเตอร์น้ำหนักของพอร์ต และทำการประเมินผลผ่านการจำลองบนข้อมูลย้อนหลังภายใต้กรอบที่แพลตฟอร์มกำหนด

ในช่วงเริ่มต้น ผู้ใช้ควรศึกษาเนื้อหาจากส่วนการเรียนรู้ของแพลตฟอร์มเพื่อทำความเข้าใจนิยามของ *alpha* โครงสร้างข้อมูลที่ใช้ ตัวดำเนินการทางคณิตศาสตร์ที่ระบบรองรับ และเกณฑ์ที่ใช้ประเมินคุณภาพของกลยุทธ์ ขั้นตอนนี้มีความสำคัญอย่างยิ่ง เพราะแม้ว่าผู้ใช้จะมีพื้นฐานด้านคณิตศาสตร์หรือการเงินอยู่แล้ว การใช้งาน BRAIN อย่างมีประสิทธิภาพยังต้องอาศัยความเข้าใจตรรกะเฉพาะของแพลตฟอร์มด้วย กล่าวคือ ผู้ใช้จำเป็นต้องทราบว่านิพจน์ที่เขียนขึ้นจะถูกตีความอย่างไรในเชิงการสร้างพอร์ต มีข้อจำกัดอะไรในการจำลอง และผลลัพธ์ที่ได้ควรอ่านอย่างไรในบริบทของการวิจัยเชิงปริมาณ

เมื่อมีความเข้าใจพื้นฐานแล้ว ขั้นตอนถัดไปคือการสร้าง *alpha* ซึ่งเป็นหัวใจสำคัญของการใช้งานแพลตฟอร์มในทางปฏิบัติ ผู้ใช้จะต้องแปลงสมมติฐานการลงทุนให้อยู่ในรูปนิพจน์ที่อาศัยข้อมูลที่แพลตฟอร์มเปิดให้ใช้ เช่น ข้อมูลราคา ปริมาณซื้อขาย หรือข้อมูลเชิงพื้นฐานบางประเภท จากนั้นจึงใช้ตัวดำเนินการต่าง ๆ เพื่อสร้างสัญญาณเชิงภาคตัดขวางที่สามารถใช้เปรียบเทียบหลักทรัพย์ในกลุ่มการลงทุนเดียวกันได้ ขั้นตอนนี้ถือเป็นการเชื่อมโยงระหว่างความคิดเชิงทฤษฎีกับการทดสอบจริง เพราะนิพจน์ที่ถูกสร้างขึ้นจะเป็นตัวแทนอย่างเป็นทางการของแนวคิดที่ผู้ใช้ต้องการตรวจสอบ

หลังจากสร้าง *alpha* แล้ว ผู้ใช้จะนำสมการดังกล่าวเข้าสู่กระบวนการจำลองผลการลงทุนย้อนหลัง การจำลองนี้เป็นจุดสำคัญที่สุดของการใช้งาน BRAIN เพราะเป็นช่วงที่แพลตฟอร์มนำสัญญาณที่นิยามไว้ไปประยุกต์กับข้อมูลในอดีต แปลงสัญญาณให้เป็นน้ำหนักพอร์ต และคำนวณผลลัพธ์ของกลยุทธ์ตามช่วงเวลาที่กำหนด กระบวนการดังกล่าวช่วยให้ผู้ใช้ไม่เพียงเห็นว่าสัญญาณมีค่าบวกหรือลบเท่านั้น แต่ยังเห็นด้วยว่าสัญญาณนี้เมื่อนำไปสร้างเป็นพอร์ตจริงแล้วมีพฤติกรรมเช่นไรในเชิงผลตอบแทน ความเสี่ยง ความเสถียร และการหมุนเวียนของพอร์ต

ผลจากการจำลองจะถูกแสดงผ่านตัวชี้วัดและกราฟหลายประเภท ซึ่งผู้ใช้จำเป็นต้องอ่านอย่างรอบคอบ การใช้งานแพลตฟอร์มอย่างมีประสิทธิภาพจึงได้หมายถึงการรับสมการให้ได้จำนวนมากที่สุด แต่หมายถึงการตีความผลลัพธ์อย่างมีวินัย ผู้ใช้ควรพิจารณาว่ากลยุทธ์มีลักษณะการสร้างผลตอบแทนที่สม่ำเสมอหรือไม่ ความเสี่ยงอยู่ในระดับที่เหมาะสมหรือไม่ และผลลัพธ์ที่ได้มีความสมเหตุสมผลในเชิงเศรษฐศาสตร์หรือเป็นเพียงความบังเอิญจากข้อมูลในอดีตเท่านั้น กล่าวอีกนัยหนึ่ง แดชบอร์ดของแพลตฟอร์มควรถูกใช้ในฐานะเครื่องมือวินิจฉัยคุณภาพของ *alpha* มากกว่าจะเป็นเพียงหน้าจอรายงานผลลัพธ์เชิงตัวเลข

ลักษณะการใช้งานจริงของ BRAIN จึงเป็นกระบวนการวนซ้ำระหว่างการตั้งสมมติฐาน การเขียนนิพจน์ การจำลอง และการปรับปรุงแบบจำลอง ในหลายกรณี *alpha* รุ่นแรกที่สร้างขึ้นอาจยังไม่ให้ผลลัพธ์ที่น่าพอใจ ผู้ใช้จึงต้องกลับไปแก้ไขนิพจน์ เปลี่ยนตัวแปรนำเข้า ปรับโครงสร้างของสัญญาณ หรือทำให้แบบจำลองเรียบง่ายขึ้น เพื่อเพิ่มความเสถียร กระบวนการวนซ้ำเช่นนี้เป็นหัวใจของการวิจัยเชิงปริมาณ เพราะช่วยให้ผู้ใช้เรียนรู้จากผลการ

ทดลองแต่ละรอบ และค่อย ๆ พัฒนาแนวคิดให้มีความแข็งแกร่งมากขึ้นทั้งในเชิงสถิติและเชิงการใช้งานจริง

เมื่อผู้ใช้พัฒนา α จนได้ผลลัพธ์ที่อยู่ในเกณฑ์ที่ยอมรับได้ของแพลตฟอร์ม ก็สามารถดำเนินการส่งผลงานเข้าสู่ระบบประเมินของ BRAIN ได้ การส่ง α ในบริบทนี้ไม่ใช่เพียงการบันทึกสมการที่เขียนเสร็จแล้ว แต่เป็นการนำเสนอผลการวิจัยเชิงระบบที่ผ่านการออกแบบ ทดสอบ และปรับปรุงมาในระดับหนึ่งแล้ว กระบวนการนี้ช่วยให้ผู้ใช้มีเป้าหมายที่ชัดเจนในการพัฒนาแบบจำลอง และทำให้แพลตฟอร์มมีลักษณะเป็นระบบนิเวศของการเรียนรู้และการแข่งขันเชิงงานวิจัยมากกว่าการเป็นเพียงเครื่องคิดเลขทางการเงิน

ในระดับที่สูงขึ้น การใช้งานแพลตฟอร์มอาจขยายไปสู่ความสามารถเพิ่มเติม เช่น การเข้าถึงข้อมูลหรือภูมิภาคการลงทุนที่มากขึ้น การจำลองในช่วงเวลาที่ยาวขึ้น การเปรียบเทียบหลายแบบจำลองพร้อมกัน หรือการใช้เครื่องมือเสริมเพื่อเพิ่มประสิทธิภาพของกระบวนการวิจัย อย่างไรก็ตาม ไม่ว่าผู้ใช้จะอยู่ในระดับเริ่มต้นหรือระดับก้าวหน้า หลักการสำคัญของการใช้งาน BRAIN ยังคงเหมือนเดิม คือการใช้แพลตฟอร์มเป็นระบบสำหรับสร้าง ตรวจสอบ และปรับปรุงสมมติฐานการลงทุนอย่างเป็นวิธีการ

โดยสรุป วิธีการใช้งานแพลตฟอร์ม WorldQuant BRAIN ที่มีประสิทธิภาพควรเริ่มจากการทำความเข้าใจกรอบแนวคิดของระบบ ศึกษาโครงสร้างข้อมูลและตัวดำเนินการที่ใช้ได้ จากนั้นจึงสร้าง α จากสมมติฐานที่มีเหตุผลนำไปจำลองผลย้อนหลัง วิเคราะห์ผลลัพธ์อย่างรอบด้าน และปรับปรุงแบบจำลองอย่างต่อเนื่องก่อนการส่งผลงานเข้าสู่ระบบประเมิน หากมองในเชิงวิธีวิทยา BRAIN จึงไม่ใช่เพียงแพลตฟอร์มสำหรับเขียนสมการการลงทุน แต่เป็นสภาพแวดล้อมที่ช่วยฝึกให้ผู้ใช้คิดและทำงานแบบนักวิจัยเชิงปริมาณอย่างแท้จริง

4 แบบจำลองเชิงคณิตศาสตร์สำหรับกระบวนการลงทุนของ WorldQuant BRAIN

แพลตฟอร์ม WorldQuant BRAIN สามารถมองได้ว่าเป็นระบบที่นิยามแผนที่เชิงคณิตศาสตร์จากข้อมูลตลาดดิบไปสู่น้ำหนักการลงทุนของพอร์ต โดยให้ U_t แทนเซตของหลักทรัพย์ที่อยู่ในกลุ่มการลงทุน ณ เวลา t และให้ $N_t = |U_t|$ เป็นจำนวนหลักทรัพย์ที่สามารถลงทุนได้ในวันนั้น หากข้อมูลตลาดประกอบด้วยตัวแปรดิบจำนวน K ชนิด เช่น ราคา ปริมาณการซื้อขาย ผลตอบแทน หรือปัจจัยพื้นฐาน เราสามารถเขียนข้อมูลแต่ละชนิดในรูปเมทริกซ์

$$X^{(k)} \in \mathbb{R}^{T \times N}, \quad k = 1, 2, \dots, K,$$

โดยสมาชิก $x_{t,i}^{(k)}$ แทนค่าของตัวแปรชนิดที่ k สำหรับหลักทรัพย์ตัวที่ i ณ เวลา t จากนั้น α สามารถนิยามเป็นตัวดำเนินการ

$$\Phi : \{X^{(1)}, X^{(2)}, \dots, X^{(K)}\} \mapsto w_t \in \mathbb{R}^{N_t},$$

ซึ่งแปลงข้อมูลดิบให้กลายเป็นเวกเตอร์น้ำหนักการลงทุน w_t สำหรับใช้สร้างพอร์ตในแต่ละวัน ดังนั้นสาระสำคัญของระบบจึงไม่ใช่เพียงการพยากรณ์ผลตอบแทนของหลักทรัพย์รายตัว แต่เป็นการสร้างเวกเตอร์น้ำหนักที่กำหนดโครงสร้างของพอร์ตทั้งระบบอย่างเป็นรูปธรรม

4.1 กระบวนการลงทุน

4.1.1 โครงสร้างข้อมูลตั้งต้น: Data Matrix

ขั้นตอนแรกของกระบวนการคือการทำให้ข้อมูลตลาดอยู่ในรูปที่เหมาะสมต่อการประมวลผลเชิงเมทริกซ์และเชิงเวกเตอร์ สมมติให้ $x_{t,i}^{(k)}$ เป็นค่าของปัจจัยดิบชนิดที่ k สำหรับหลักทรัพย์ตัวที่ i ณ เวลา t ข้อมูลทั้งหมดที่สังเกตได้ใน

แต่ละวันจึงสามารถเขียนในรูปเวกเตอร์ภาคตัดขวาง

$$x_t^{(k)} = (x_{t,1}^{(k)}, x_{t,2}^{(k)}, \dots, x_{t,N_t}^{(k)})^\top \in \mathbb{R}^{N_t}.$$

หากพิจารณาต่อเนื่องตามเวลา ข้อมูลแต่ละชนิดจะรวมกันเป็นเมทริกซ์

$$X^{(k)} = \begin{bmatrix} x_{1,1}^{(k)} & x_{1,2}^{(k)} & \cdots & x_{1,N}^{(k)} \\ x_{2,1}^{(k)} & x_{2,2}^{(k)} & \cdots & x_{2,N}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T,1}^{(k)} & x_{T,2}^{(k)} & \cdots & x_{T,N}^{(k)} \end{bmatrix}.$$

รูปแบบนี้มีความสำคัญเพราะทำให้การสร้าง α สามารถใช้ทั้งโครงสร้างแบบอนุกรมเวลาและแบบภาคตัดขวาง พร้อมกัน กล่าวคือ ระบบสามารถอ้างอิงประวัติย้อนหลังของหลักทรัพย์แต่ละตัว และในเวลาเดียวกันก็สามารถเปรียบเทียบหลักทรัพย์ต่าง ๆ ภายในวันเดียวกันได้

4.1.2 การแปลงข้อมูลเป็นสัญญาณ: Transformation

เมื่อได้ข้อมูลดิบในรูปเมทริกซ์แล้ว ขั้นตอนถัดไปคือการสร้างสัญญาณก่อนการจัดอันดับ (*pre-signal*) ผ่านนิพจน์ของ α หากให้ $\mathcal{F}_{t,i}$ แทนเซตของข้อมูลทั้งหมดที่ใช้ได้สำหรับหลักทรัพย์ตัวที่ i ณ เวลา t เช่นหน้าตาข้อมูลย้อนหลังยาว L วัน เราสามารถเขียนสัญญาณเบื้องต้นได้เป็น

$$a_{t,i} = \Phi(\mathcal{F}_{t,i}),$$

โดยที่ Φ อาจเป็นฟังก์ชันเชิงเส้นหรือไม่เชิงเส้น เช่น ส่วนต่างของค่าเฉลี่ยเคลื่อนที่ อัตราการเปลี่ยนแปลงของราคา สัดส่วนทางบัญชี หรือการประกอบหลายตัวแปรเข้าด้วยกัน ตัวอย่างเชิงนามธรรมของสัญญาณชนิดหนึ่งคือ

$$a_{t,i} = f(x_{t,i}^{(1)}, x_{t-1,i}^{(1)}, \dots, x_{t-L+1,i}^{(1)}, x_{t,i}^{(2)}, \dots, x_{t,i}^{(K)}).$$

ในขั้นนี้ $a_{t,i}$ ยังเป็นเพียงคะแนนเชิงดิบซึ่งอาจมีหน่วยวัดและการกระจายตัวต่างกันมากระหว่างหลักทรัพย์หรือระหว่าง α คนละชนิด ดังนั้นจึงยังไม่เหมาะต่อการนำไปใช้เป็นน้ำหนักพอร์ตโดยตรง

4.1.3 การจัดอันดับแบบภาคตัดขวาง: Ranking

เพื่อให้สัญญาณมีความสามารถในการเปรียบเทียบกันภาคตัดขวาง ระบบมักใช้การจัดอันดับภายในวันเดียวกัน สมมติให้ $a_t = (a_{t,1}, \dots, a_{t,N_t})^\top$ เป็นเวกเตอร์สัญญาณดิบ ณ เวลา t แล้วนิยามอันดับของหลักทรัพย์ตัวที่ i เป็น $\text{rank}_t(a_{t,i})$ โดยมีค่าอยู่ระหว่าง 1 ถึง N_t จากนั้นสามารถแปลงอันดับดังกล่าวให้อยู่ในช่วง $[0, 1]$ ได้โดย

$$r_{t,i} = \frac{\text{rank}_t(a_{t,i}) - 1}{N_t - 1}.$$

หากต้องการให้สัญญาณมีศูนย์กลางอยู่ที่ศูนย์เพื่อสะดวกต่อการสร้างพอร์ตแบบ long-short อาจทำการเลื่อนสเกลเพิ่มเติมเป็น

$$\tilde{r}_{t,i} = 2r_{t,i} - 1,$$

ซึ่งทำให้ $\tilde{r}_{t,i} \in [-1, 1]$ การเปลี่ยนจาก $a_{t,i}$ ไปเป็น $r_{t,i}$ หรือ $\tilde{r}_{t,i}$ มีความหมายเชิงคณิตศาสตร์ว่า ระบบไม่ได้ใช้ขนาดสัมบูรณ์ของสัญญาณอีกต่อไป แต่ใช้ลำดับสัมพัทธ์ของหลักทรัพย์ภายในกลุ่มเดียวกันแทน วิธีนี้ช่วยลดอิทธิพลของ outlier และทำให้ปัจจัยที่มีหน่วยวัดต่างกันสามารถนำมาใช้ภายใต้กรอบเดียวกันได้

4.1.4 การทำให้สัญญาณเป็นกลาง: Neutralization

แม้ว่าสัญญาณหลังการจัดอันดับจะพร้อมสำหรับการเปรียบเทียบในภาคตัดขวางแล้ว แต่ก็ยังอาจปนเปื้อนด้วยองค์ประกอบร่วมของตลาด เช่น market beta, industry effect หรือ style exposure ดังนั้นจึงต้องมีขั้นตอนการทำให้เป็นกลาง (neutralization) ซึ่งในทางคณิตศาสตร์สามารถเขียนเป็นปัญหาการฉายเชิงเส้นได้

ให้ $z_t \in \mathbb{R}^{N_t}$ เป็นเวกเตอร์สัญญาณที่ต้องการทำให้เป็นกลาง และให้

$$B_t \in \mathbb{R}^{N_t \times p}$$

เป็นเมทริกซ์ความเสี่ยงหรือเมทริกซ์ของปัจจัยที่ไม่ต้องการให้สัญญาณมีความสัมพันธ์ด้วย เช่น คอลัมน์หนึ่งอาจเป็นเวกเตอร์หนึ่งทั้งหมดเพื่อแทน market mean ขณะที่คอลัมน์อื่นอาจเป็น dummy variables ของอุตสาหกรรม การทำให้สัญญาณเป็นกลางสามารถนิยามเป็นการดึงส่วนที่ตั้งฉากกับปริภูมิที่สร้างโดย B_t ออกมา ดังนี้

$$s_t = (I - P_{B_t})z_t, \quad P_{B_t} = B_t(B_t^\top B_t)^{-1}B_t^\top,$$

โดย P_{B_t} คือเมทริกซ์ฉายลงบนปริภูมิของปัจจัยใน B_t และ s_t คือเวกเตอร์สัญญาณหลัง neutralization กรณีพื้นฐานที่สุดเกิดขึ้นเมื่อ $B_t = \mathbf{1}$ ซึ่งเป็นเวกเตอร์หนึ่งทั้งหมด เราจะได้

$$s_{t,i} = z_{t,i} - \frac{1}{N_t} \sum_{j=1}^{N_t} z_{t,j},$$

ซึ่งก็คือการหักค่าเฉลี่ยภาคตัดขวางออกจากสัญญาณเดิมโดยตรง ผลลัพธ์คือ

$$\sum_{i=1}^{N_t} s_{t,i} = 0,$$

ทำให้สัญญาณมีลักษณะ dollar-neutral ในระดับโครงสร้าง และลดความเอนเอียงไปตามทิศทางตลาดโดยรวม

4.1.5 การทำให้สัญญาณพร้อมใช้งานในพอร์ต: Normalization

หลังจากได้สัญญาณที่ผ่านการ neutralize แล้ว ขั้นตอนถัดไปคือการแปลงสัญญาณดังกล่าวให้กลายเป็นน้ำหนักการลงทุนที่มีขนาดควบคุมได้ หากให้ $s_t = (s_{t,1}, \dots, s_{t,N_t})^\top$ เป็นสัญญาณหลัง neutralization เราสามารถนิยามน้ำหนักเบื้องต้นของพอร์ตเป็น

$$w_{t,i}^{(0)} = \frac{s_{t,i}}{\sum_{j=1}^{N_t} |s_{t,j}|}.$$

นิยามนี้ทำให้

$$\sum_{i=1}^{N_t} |w_{t,i}^{(0)}| = 1,$$

ซึ่งหมายความว่าพอร์ตถูกทำให้มี gross exposure เท่ากับหนึ่งหน่วย หากก่อนหน้านี้นี้สัญญาณถูกทำให้มีค่าเฉลี่ยเป็นศูนย์แล้ว จะได้เพิ่มเติมว่า

$$\sum_{i=1}^{N_t} w_{t,i}^{(0)} = 0,$$

ซึ่งตีความได้ว่าโครงสร้างพอร์ตมีความสมดุลระหว่างฝั่ง long และ short

ในทางปฏิบัติ บางระบบอาจเพิ่มข้อจำกัดเรื่องน้ำหนักสูงสุดต่อสินทรัพย์ เพื่อป้องกันการกระจุกตัว โดยนิยามการตัดค่าน้ำหนักเป็น

$$\bar{w}_{t,i} = \text{clip}(w_{t,i}^{(0)}, -\tau, \tau),$$

เมื่อ $\tau > 0$ เป็นเพดานน้ำหนัก จากนั้นจึง normalize ซ้ำอีกครั้งเป็น

$$w_{t,i} = \frac{\bar{w}_{t,i}}{\sum_{j=1}^{N_t} |\bar{w}_{t,j}|}.$$

ในเชิงคณิตศาสตร์ ขั้นตอนนี้ทำหน้าที่เปลี่ยนเวกเตอร์คะแนนให้กลายเป็นเวกเตอร์น้ำหนักที่สอดคล้องกับข้อจำกัดด้านการบริหารความเสี่ยง

4.1.6 การจัดสรรเงินลงทุนจริง: Capital Allocation

เมื่อได้เวกเตอร์น้ำหนักสุดท้าย w_t แล้ว จึงสามารถแปลงให้เป็นมูลค่าเงินลงทุนจริงได้โดยอาศัยขนาดพอร์ตทั้งหมด สมมติให้ $A_t > 0$ แทน book size หรือเงินทุนรวมที่ใช้ลงทุน ณ เวลา t ตำแหน่งเงินลงทุนของหลักทรัพย์ตัวที่ i สามารถเขียนได้เป็น

$$d_{t,i} = A_t w_{t,i},$$

โดย $d_{t,i} > 0$ หมายถึงสถานะ long และ $d_{t,i} < 0$ หมายถึงสถานะ short หากต้องการแปลงมูลค่าเงินลงทุนให้เป็นจำนวนหุ้น สมมติให้ $P_{t,i}$ เป็นราคาของหลักทรัพย์ตัวที่ i ณ เวลา t จะได้ว่า

$$q_{t,i} = \frac{d_{t,i}}{P_{t,i}},$$

โดย $q_{t,i}$ คือจำนวนหน่วยที่ต้องถือจริงในพอร์ต

โครงสร้างนี้แสดงให้เห็นอย่างชัดเจนว่า α ในระบบ WorldQuant BRAIN มิได้จบลงที่การสร้างคะแนนหรือการจัดอันดับ แต่สิ้นสุดที่การให้คำสั่งเชิงปริมาณที่นำไปใช้ซื้อขายได้จริง กล่าวคือ สัญญาณเชิงบวกจะถูกแปลงเป็นน้ำหนักบวกและนำไปสู่สถานะ long ขณะที่สัญญาณเชิงลบจะถูกแปลงเป็นน้ำหนักลบและนำไปสู่สถานะ short ดังนั้นทั้งกระบวนการสามารถสรุปเป็นลำดับการแปลงเชิงคณิตศาสตร์ได้ว่า

$$\{X^{(1)}, \dots, X^{(K)}\} \xrightarrow{\Phi} a_t \xrightarrow{\text{rank}} z_t \xrightarrow{\text{neutralize}} s_t \xrightarrow{\text{normalize}} w_t \xrightarrow{\text{allocate}} d_t.$$

หากพิจารณาในเชิงนามธรรมที่สุด กระบวนการของ WorldQuant BRAIN คือการสร้างตัวดำเนินการต่อเนื่องจากข้อมูลไปสู่พอร์ตลงทุน โดยเริ่มจากข้อมูลดิบในรูปเมทริกซ์ สร้างสัญญาณผ่านนิพจน์ของ α รับสัญญาณให้อยู่ในสเกลที่เปรียบเทียบกันได้ด้วยการจัดอันดับ กำจัดองค์ประกอบร่วมที่ไม่ต้องการด้วยการฉายเชิงเส้น แปลงสัญญาณที่ได้ให้เป็นน้ำหนักซึ่งสอดคล้องกับข้อจำกัดของพอร์ต และสุดท้ายเชื่อมโยงน้ำหนักเหล่านั้นเข้ากับเงินทุนจริงเพื่อให้เกิดตำแหน่งลงทุนที่นำไปดำเนินการได้จริง ด้วยเหตุนี้ คุณภาพของ α จึงขึ้นอยู่กับเพียงแต่ความสามารถในการสร้างสัญญาณที่ดีเท่านั้น แต่ยังคงขึ้นอยู่กับความเหมาะสมของโครงสร้างเชิงคณิตศาสตร์ที่ใช้เปลี่ยนสัญญาณดังกล่าวให้เป็นพอร์ตที่มีเสถียรภาพและควบคุมความเสี่ยงได้ด้วย

4.2 ตัวอย่างเชิงตัวเลขสำหรับพอร์ตการลงทุนขนาดเล็ก

เพื่อให้เห็นภาพของกระบวนการทั้งหมดอย่างเป็นรูปธรรม สมมติให้ ณ วัน t มีกลุ่มการลงทุนขนาดเล็กประกอบด้วยหลักทรัพย์ 4 ตัว คือ

$$u_t = \{A, B, C, D\}, \quad N_t = 4.$$

สมมติว่าใช้ข้อมูลเพียงชนิดเดียวคือราคาปิด (*close*) และกำหนด *alpha expression* เป็น

$$a_{t,i} = - \left(\frac{P_{t,i}}{P_{t-1,i}} - 1 \right),$$

ซึ่งเป็นสัญญาณแบบ *one-day reversal* กล่าวคือ หลักทรัพย์ที่ให้ผลตอบแทนสูงในวันก่อนหน้าจะได้สัญญาณต่ำกว่า และหลักทรัพย์ที่ให้ผลตอบแทนต่ำในวันก่อนหน้าจะได้สัญญาณสูงกว่า

ขั้นที่ 1: สร้าง Data Matrix สมมติให้ราคาปิดของหลักทรัพย์ทั้ง 4 ตัวในสองวันล่าสุดเป็นดังนี้

$$X^{(\text{close})} = \begin{bmatrix} 100 & 50 & 80 & 40 \\ 102 & 49 & 84 & 38 \end{bmatrix},$$

โดยแถวแรกแทนวัน $t - 1$ และแถวที่สองแทนวัน t ดังนั้นจะได้ว่า

$$P_{t-1} = (100, 50, 80, 40)^\top, \quad P_t = (102, 49, 84, 38)^\top.$$

ขั้นที่ 2: คำนวณผลตอบแทนรายวันและสัญญาณดิบ คำนวณผลตอบแทนหนึ่งวันย้อนหลังของแต่ละหลักทรัพย์

$$r_{t,i}^{(1)} = \frac{P_{t,i}}{P_{t-1,i}} - 1.$$

ดังนั้น

$$r_t^{(1)} = \left(\frac{102}{100} - 1, \frac{49}{50} - 1, \frac{84}{80} - 1, \frac{38}{40} - 1 \right)^\top = (0.02, -0.02, 0.05, -0.05)^\top.$$

จากนิพจน์ของ *alpha* จะได้สัญญาณดิบ

$$a_t = -r_t^{(1)} = (-0.02, 0.02, -0.05, 0.05)^\top.$$

ขั้นที่ 3: Ranking แบบภาคตัดขวาง นำค่าสัญญาณดิบมาเรียงอันดับจากน้อยไปมาก จะได้ลำดับดังนี้

$$C = -0.05 < A = -0.02 < B = 0.02 < D = 0.05.$$

ดังนั้นอันดับของแต่ละหลักทรัพย์คือ

$$\text{rank}_t(A) = 2, \quad \text{rank}_t(B) = 3, \quad \text{rank}_t(C) = 1, \quad \text{rank}_t(D) = 4.$$

แปลงอันดับให้อยู่ในช่วง $[0, 1]$ ด้วยสูตร

$$r_{t,i} = \frac{\text{rank}_t(a_{t,i}) - 1}{N_t - 1}.$$

เนื่องจาก $N_t = 4$ จะได้

$$r_t = \left(\frac{2-1}{3}, \frac{3-1}{3}, \frac{1-1}{3}, \frac{4-1}{3} \right)^\top = \left(\frac{1}{3}, \frac{2}{3}, 0, 1 \right)^\top.$$

จากนั้นเลื่อนสเกลให้อยู่ในช่วง $[-1, 1]$

$$z_{t,i} = 2r_{t,i} - 1,$$

จึงได้

$$z_t = \left(-\frac{1}{3}, \frac{1}{3}, -1, 1 \right)^\top.$$

ขั้นที่ 4: Neutralization ในกรณีพื้นฐานที่สุด ให้ทำ *mean neutralization* โดยหักค่าเฉลี่ยภาคตัดขวางออก

$$s_{t,i} = z_{t,i} - \frac{1}{N_t} \sum_{j=1}^{N_t} z_{t,j}.$$

สำหรับตัวอย่างนี้

$$\frac{1}{4} \left(-\frac{1}{3} + \frac{1}{3} - 1 + 1 \right) = 0,$$

ดังนั้น

$$s_t = z_t = \left(-\frac{1}{3}, \frac{1}{3}, -1, 1 \right)^\top.$$

กล่าวคือ ในตัวอย่างนี้การ neutralization ไม่เปลี่ยนค่า เพราะเวกเตอร์สัญญาณมีค่าเฉลี่ยเป็นศูนย์อยู่แล้ว

ขั้นที่ 5: Normalization นิยามน้ำหนักเบื้องต้นของพอร์ตเป็น

$$w_{t,i}^{(0)} = \frac{s_{t,i}}{\sum_{j=1}^{N_t} |s_{t,j}|}.$$

คำนวณผลรวมค่าสัมบูรณ์ก่อน

$$\sum_{j=1}^4 |s_{t,j}| = \left| -\frac{1}{3} \right| + \left| \frac{1}{3} \right| + |-1| + |1| = \frac{1}{3} + \frac{1}{3} + 1 + 1 = \frac{8}{3}.$$

ดังนั้นน้ำหนักเบื้องต้นคือ

$$w_t^{(0)} = \left(\frac{-1/3}{8/3}, \frac{1/3}{8/3}, \frac{-1}{8/3}, \frac{1}{8/3} \right)^\top = (-0.125, 0.125, -0.375, 0.375)^\top.$$

ขั้นที่ 6: Capital Allocation สมมติให้ขนาดพอร์ตเท่ากับ

$$A_t = 1,000,000.$$

มูลค่าเงินลงทุนของแต่ละหลักทรัพย์คือ

$$d_{t,i} = A_t w_{t,i}.$$

ดังนั้น

$$d_t = 1,000,000 \cdot (-0.1471, 0.1471, -0.3529, 0.3529)^\top$$

$$\approx (-147,058.82, 147,058.82, -352,941.18, 352,941.18)^\top.$$

กล่าวคือ ระบบถือสถานะ

$$\text{Short } A, \quad \text{Long } B, \quad \text{Short } C, \quad \text{Long } D.$$

ขั้นที่ 8: แปลงเป็นจำนวนหุ้น จากระาคาปิด ณ วัน t

$$P_t = (102, 49, 84, 38)^\top$$

จำนวนหุ้นที่ต้องถือคือ

$$q_{t,i} = \frac{d_{t,i}}{P_{t,i}}.$$

ดังนั้น

$$\begin{aligned} q_t &\approx \left(\frac{-147,058.82}{102}, \frac{147,058.82}{49}, \frac{-352,941.18}{84}, \frac{352,941.18}{38} \right)^\top \\ &\approx (-1441.75, 3001.20, -4201.68, 9287.93)^\top. \end{aligned}$$

ขั้นที่ 9: คำนวณผลตอบแทนพอร์ตในวันถัดไป สมมติว่าราคาปิดในวัน $t + 1$ เป็น

$$P_{t+1} = (101, 50, 82, 39)^\top.$$

ผลกำไรขาดทุนของแต่ละสถานะเท่ากับ

$$\Pi_{t+1,i} = q_{t,i}(P_{t+1,i} - P_{t,i}).$$

ดังนั้นจะได้

$$\Pi_{t+1,A} = (-1441.75)(101 - 102) = 1441.75,$$

$$\Pi_{t+1,B} = (3001.20)(50 - 49) = 3001.20,$$

$$\Pi_{t+1,C} = (-4201.68)(82 - 84) = 8403.36,$$

$$\Pi_{t+1,D} = (9287.93)(39 - 38) = 9287.93.$$

จึงได้กำไรรวมของพอร์ต

$$\Pi_{t+1} = 1441.75 + 3001.20 + 8403.36 + 9287.93 = 22,134.24.$$

เมื่อเทียบกับเงินทุนเริ่มต้น 1,000,000 จะได้ผลตอบแทนพอร์ตประมาณ

$$R_{p,t+1} = \frac{22,134.24}{1,000,000} = 0.02213 \approx 2.213\%.$$

ตัวอย่างนี้แสดงให้เห็นลำดับของกระบวนการทั้งหมดอย่างชัดเจน ตั้งแต่การเริ่มต้นด้วยข้อมูลราคาในรูปแบบทริกซ์ การสร้างสัญญาณดิบผ่าน *alpha expression* การจัดอันดับแบบภาคตัดขวาง การทำให้สัญญาณเป็นกลาง การ normalize และ truncate น้ำหนัก การจัดสรรเงินลงทุนจริง และสุดท้ายการคำนวณกำไรขาดทุนของพอร์ตในวันถัดไป ดังนั้นในเชิงปฏิบัติ *alpha* จึงทำหน้าที่เป็นตัวแปลงข้อมูลตลาดให้กลายเป็นคำสั่งเชิงปริมาณสำหรับการสร้างพอร์ตลงทุนอย่างเป็นระบบ

4.3 การวัดผลการลงทุน

ในแพลตฟอร์ม WorldQuant BRAIN ตัวชี้วัดในส่วน IS Summary (Aggregate Data) มักตีความตามกรอบการประเมินผลของพอร์ต ลอง-ชอร์ต โดยมีนิยามเชิงคณิตศาสตร์ดังนี้

ชาร์ป (Sharpe) เป็นอัตราส่วนผลตอบแทนต่อความผันผวนของ **กำไร-ขาดทุน** รายวัน (ปรับเป็นรายปี) ให้ PnL_t คือกำไร-ขาดทุน ณ วัน t , $\overline{PnL}$ คือค่าเฉลี่ยตัวอย่าง และ $s(PnL)$ คือส่วนเบี่ยงเบนมาตรฐานตัวอย่าง จะได้

$$Sharpe = \frac{\overline{PnL}}{s(PnL)} \sqrt{252}. \quad (4.1)$$

อัตราการหมุนเวียนพอร์ต (Turnover) วัดสัดส่วนการซื้อขายเทียบกับ **ขนาดพอร์ตเชิงบัญชี (Book Size)** โดยนิยามเชิงแนวคิดเป็นอัตราส่วนของ **มูลค่าซื้อขายรวม (ดอลลาร์)** ต่อ **ขนาดพอร์ตเชิงบัญชี** (มักรายงานเป็นค่าเฉลี่ยรายวัน) คือ

$$Turnover = \frac{\text{มูลค่าซื้อขายรวม (ดอลลาร์)}}{\text{ขนาดพอร์ตเชิงบัญชี}}. \quad (4.2)$$

ผลตอบแทนรวม (Returns) มักรายงานเป็นผลตอบแทนเทียบกับเงินทุนที่ใช้จริงของพอร์ตลอง-ชอร์ต โดยอาศัยว่าเงินลงทุนสุทธิโดยประมาณสัมพันธ์กับครึ่งหนึ่งของขนาดพอร์ตเชิงบัญชี จึงเขียนในรูป

$$Returns = \frac{\text{กำไร-ขาดทุนต่อปี (Annualized PnL)}}{\frac{1}{2} \text{ขนาดพอร์ตเชิงบัญชี}}. \quad (4.3)$$

คะแนนความฟิต (Fitness) เป็นคะแนนรวมที่ให้รางวัลกับความสม่ำเสมอและความสามารถในการทำกำไร (ผ่าน Sharpe และ Returns) และลงโทษการซื้อขายถี่ (ผ่าน Turnover) โดยใช้สูตร

$$Fitness = Sharpe \sqrt{\frac{|Returns|}{\max(Turnover, 0.125)}}. \quad (4.4)$$

การลดลงสูงสุด (Drawdown) วัดการขาดทุนจากจุดสูงสุดลงสู่จุดต่ำสุดที่รุนแรงที่สุดของเส้นมูลค่า/กำไรสะสม ให้ V_t เป็นมูลค่า (หรือความมั่งคั่งสะสม) และ $M_t = \max_{1 \leq s \leq t} V_s$ จะได้

$$DD_t = \frac{V_t - M_t}{M_t}, \quad \text{Drawdown} = \min_{1 \leq t \leq T} DD_t, \quad (4.5)$$

โดยหลายรายงานอาจแสดงเป็นค่าบวกด้วย $-Drawdown$

มาร์จิ้น (Margin) ตีความเป็น **กำไรต่อมูลค่าซื้อขาย** หรือประสิทธิภาพกำไรต่อกิจกรรมการซื้อขาย นิยามเป็น

$$Margin = \frac{PnL}{\text{มูลค่าซื้อขายรวม (ดอลลาร์)}}, \quad (4.6)$$

ซึ่งมักรายงานในหน่วย **จุดพื้นฐาน (basis points: bps)** เพื่อให้เปรียบเทียบได้ชัดเจนขึ้น

นอกจากการพิจารณาค่าตัวชี้วัดข้างต้นในเชิงตีความผลการลงทุนแล้ว แพลตฟอร์ม WorldQuant BRAIN ยังมี *submission criteria* สำหรับพิจารณาว่า alpha โดยอยู่ในสถานะที่สามารถส่งเข้า *alpha pool* ได้ โดยในทางปฏิบัติ เกณฑ์ดังกล่าวขึ้นอยู่กับ *Region* และ *Delay* ที่ใช้ในการจำลอง ดังนั้นเงื่อนไขเชิงตัวเลขอาจแตกต่างกันได้ตามบริบทของการทดสอบ

ในงานวิจัยนี้ เพื่อใช้เป็นกรอบอ้างอิงในการประเมินความพร้อมของสูตรแอลฟา จะใช้ชุดเกณฑ์ตัวอย่างที่มักอ้างถึงสำหรับกรณี **USA, Delay 1** ซึ่งประกอบด้วยเงื่อนไขสำคัญดังนี้

$$\text{Sharpe} \geq 1.25, \quad \text{Fitness} \geq 1.00, \quad 0.01 \leq \text{Turnover} \leq 0.70. \quad (4.7)$$

กล่าวคือ alpha ที่จะถือว่าอยู่ในระดับพร้อมสำหรับการส่งควรมีผลตอบแทนต่อความเสี่ยงอยู่ในระดับเพียงพอ มีคะแนน Fitness ผ่านเกณฑ์ขั้นต่ำ และไม่ซื้อขายบ่อยเกินไปหรือถี่เกินไปจนทำให้เกิดปัญหาเชิงคุณภาพของสัญญาณหรือเชิงต้นทุนธุรกรรม

นอกจากเกณฑ์ด้านผลการดำเนินงานแล้ว ยังมีเงื่อนไขด้านการกระจุกตัวของน้ำหนัก (*weight concentration*) เพื่อป้องกันไม่ให้ alpha อาศัยน้ำหนักเพียงไม่กี่หลักทรัพย์ในการขับเคลื่อนผลลัพธ์ โดยสามารถเขียนในรูปอย่างย่อได้เป็น

$$\max_i |w_{i,t}| < 0.10, \quad (4.8)$$

ซึ่งหมายความว่า น้ำหนักสัมบูรณ์สูงสุดของหลักทรัพย์รายตัวไม่ควรเกิน 10% ของพอร์ตในช่วงเวลาที่ประเมิน

ในด้านความทนทานของสัญญาณ (*robustness*) แพลตฟอร์มยังตรวจสอบ **Sub-universe Test** เพื่อประเมินว่า alpha ยังให้ผลลัพธ์ที่ดีเมื่อจำกัดจักรวาลการลงทุนให้แคบลงและมีสภาพคล่องสูงขึ้น โดยสามารถเขียนในรูปทั่วไปได้เป็น

$$\text{Sharpe}_{\text{sub}} \geq \eta_{\text{sub}}, \quad (4.9)$$

เมื่อ η_{sub} เป็นค่าขั้นต่ำที่แพลตฟอร์มกำหนดสำหรับจักรวาลย่อยนั้น ๆ ทั้งนี้ ในตัวอย่างเกณฑ์สาธารณะที่มักอ้างถึงสำหรับกรณี **USA, Delay 1** มีระบุค่าอ้างอิงประมาณ

$$\text{Sharpe}_{\text{sub}} > 0.26. \quad (4.10)$$

อีกเงื่อนไขสำคัญคือ **Self-Correlation Test** ซึ่งใช้จำกัดความซ้ำซ้อนของ alpha กับสัญญาณที่เคยส่งหรือมีอยู่แล้วในระบบ เพื่อป้องกันการส่งสูตรที่คล้ายคลึงกันมากเกินไป โดยอาจเขียนได้เป็น

$$\rho(\alpha_{\text{new}}, \alpha_{\text{existing}}) < 0.7, \quad (4.11)$$

ซึ่งหมายความว่า alpha ใหม่ควรมีความสัมพันธ์เชิงสัญญาณกับ alpha ที่มีอยู่เดิมต่ำกว่า 0.7 จึงจะถือว่ามีความแตกต่างเพียงพอในเชิงโครงสร้าง

ดังนั้น ในทางระเบียบวิธี งานวิจัยนี้จึงใช้ตัวชี้วัด *Sharpe, Turnover, Returns, Fitness, Drawdown* และ *Margin* สำหรับวิเคราะห์คุณภาพเชิงผลการลงทุนของสูตรแอลฟา และใช้เกณฑ์การผ่านของแพลตฟอร์มในฐานะกรอบอ้างอิงเพิ่มเติมเพื่อพิจารณาว่าสูตรใดมีศักยภาพเพียงพอสำหรับการนำไปสู่การ *submit* บน WorldQuant BRAIN ภายใต้เงื่อนไขของตลาดและการตั้งราคาที่ศึกษา

5 ตัวอย่างการสร้างสูตรแอลฟาจากสมมุติฐาน

บทนี้นำเสนอการทดสอบเชิงประจักษ์ของสัญญาณอัลฟาที่รายงานไว้ในงานวิจัยของ [1] โดยมีวัตถุประสงค์เพื่อประเมินความสัมพันธ์ระหว่างสัญญาณดังกล่าวกับผลตอบแทนล่วงหน้าในระยะสั้น ตลอดจนตรวจสอบความสม่ำเสมอของสมรรถนะของสัญญาณภายใต้กรอบการสร้างพอร์ตการลงทุนและการปรับความเสี่ยงตามแบบแผนที่

ยอมรับในงานวิจัยการเงินเชิงปริมาณ ทั้งนี้ การดำเนินการทดสอบจะยึดตามนิยามของตัวแปร ขั้นตอนการจัดหลักทรัพย์ และวิธีการประเมินผลตามที่งานวิจัยต้นฉบับกำหนดเป็นหลัก พร้อมทั้งระบุสมมติฐาน ขอบเขตข้อมูล และข้อจำกัดในการประยุกต์ใช้ให้ชัดเจน เพื่อให้ผลการศึกษาสามารถตรวจสอบได้และสามารถเปรียบเทียบกับหลักฐานเชิงประจักษ์เดิมได้อย่างเป็นระบบ

5.1 สัญญาณกลับทิศของการเปลี่ยนแปลงราคาแบบมีเงื่อนไขตามการเปลี่ยนแปลงของปริมาณซื้อขาย

$$(\text{sign}(\text{delta}(\text{volume}, 1)) * (-1 * \text{delta}(\text{close}, 1))) \quad (5.1)$$

ในส่วนนี้นำเสนอแนวคิดเชิงสมมติฐานของสัญญาณแบบ WorldQuant alpha ที่นิยามโดย

$$\alpha_{i,t} = \text{sign}(\Delta V_{i,t}) \cdot (-\Delta C_{i,t}), \quad \Delta X_{i,t} = X_{i,t} - X_{i,t-1}, \quad (5.2)$$

โดยที่ $V_{i,t}$ คือปริมาณการซื้อขาย (volume) ของหลักทรัพย์ i ณ วัน t และ $C_{i,t}$ คือราคาปิด (close) ณ วัน t . ฟังก์ชัน $\text{sign}(\cdot)$ ให้ค่า $+1$ เมื่อ $\Delta V_{i,t} > 0$, ให้ค่า -1 เมื่อ $\Delta V_{i,t} < 0$ และให้ค่า 0 เมื่อ $\Delta V_{i,t} = 0$.

การตีความสัญญาณ สมการข้างต้นสามารถตีความเป็นสัญญาณกลับทิศของการเปลี่ยนแปลงราคาแบบมีเงื่อนไขตามการเปลี่ยนแปลงของปริมาณซื้อขาย กล่าวคือ (1) เมื่อปริมาณซื้อขายเพิ่มขึ้น ($\Delta V_{i,t} > 0$) จะได้ $\alpha_{i,t} = -\Delta C_{i,t}$ ซึ่งมีลักษณะ *สวนทาง* กับการเปลี่ยนแปลงราคาล่าสุด (contrarian); (2) เมื่อปริมาณซื้อขายลดลง ($\Delta V_{i,t} < 0$) จะได้ $\alpha_{i,t} = +\Delta C_{i,t}$ ซึ่งมีลักษณะ *ไปตามทิศ* ของการเปลี่ยนแปลงราคาล่าสุด (momentum). ดังนั้นสัญญาณนี้แยกกลไกการคาดการณ์ผลตอบแทนออกเป็นสองระบอบ (regimes) ตามทิศทางของการเปลี่ยนแปลงสภาพคล่อง/ความหนาแน่นของการซื้อขายในระยะสั้น.

สมมติฐาน (Hypotheses) เพื่อให้การทดสอบเป็นเชิงวิชาการ สามารถตั้งสมมติฐานเชิงกลไกได้ดังนี้

H1 (Reversal ภายใต้ปริมาณซื้อขายเพิ่มขึ้น): เมื่อ $\Delta V_{i,t} > 0$ การเคลื่อนไหวของราคา $\Delta C_{i,t}$ มีแนวโน้มสะท้อนแรงซื้อขายเชิงสภาพคล่อง (liquidity-driven) หรือภาวะตอบสนองเกิน (overreaction) ในระยะสั้น ส่งผลให้คาดการณ์การกลับตัวของราคาในวันถัดไป

$$\mathbb{E}[R_{i,t+1} \mid \Delta V_{i,t} > 0] \propto -\Delta C_{i,t}. \quad (5.3)$$

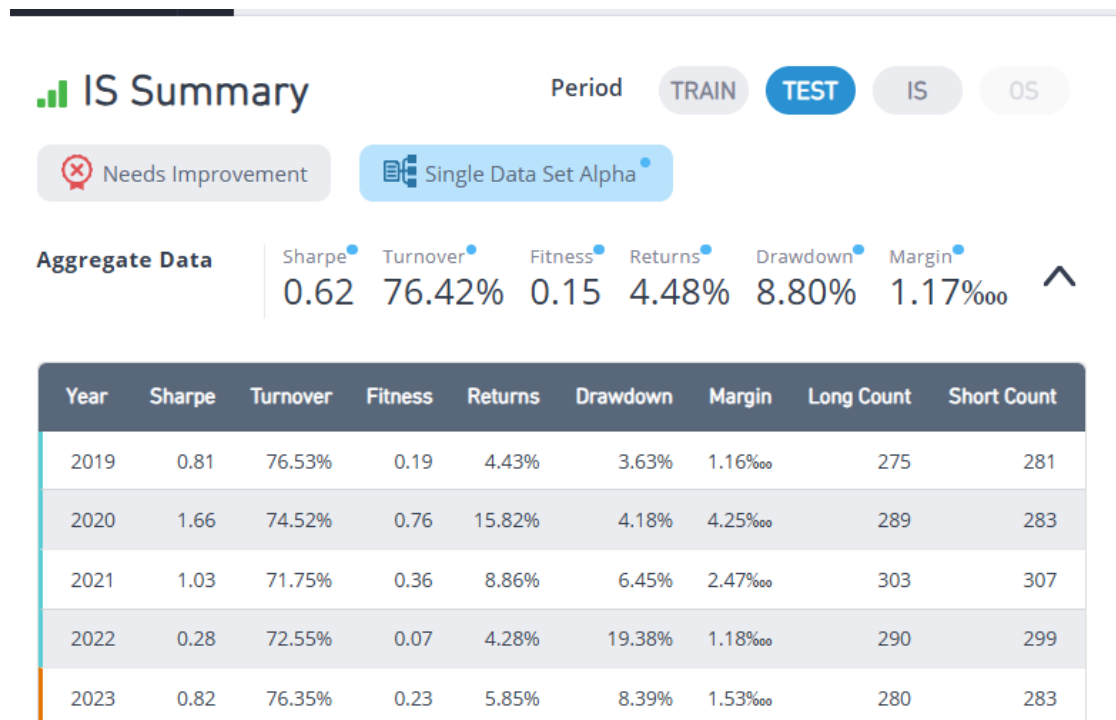
สรุปผลการทดสอบ (ช่วง TEST) จากผลสรุปการทดสอบนอกตัวอย่าง (out-of-sample; TEST) แพลตฟอร์มจัดระดับคุณภาพของสัญญาณเป็น *Needs Improvement* ซึ่งสะท้อนว่าแม้อัลฟาจะให้ผลตอบแทนเป็นบวก แต่สมรรถนะโดยรวมยังไม่อยู่ในระดับที่น่าพึงพอใจเมื่อพิจารณาควบคู่กับความเสี่ยงและภาระการซื้อขาย โดยผลรวมตลอดช่วงทดสอบมีค่า Sharpe เท่ากับ 0.62 ผลตอบแทนสะสม (Returns) เท่ากับ 4.48% และการลดลงสูงสุด (Drawdown) เท่ากับ 8.80% บ่งชี้ว่าผลตอบแทนต่อความเสี่ยงยังมีขนาดจำกัด และมีการถดถอยของมูลค่าพอร์ตที่มีนัยในเชิงเศรษฐศาสตร์เมื่อเทียบกับผลตอบแทนรวม

ภาระการซื้อขายและผลกระทบต่อต้นทุนธุรกรรม ในด้านความสามารถเชิงปฏิบัติ (implementability) พบว่าอัลฟามีอัตราการหมุนเวียนพอร์ต (Turnover) สูงถึง 76.42% ซึ่งหมายถึงความถี่ของการปรับพอร์ตอยู่ในระดับสูง และมีแนวโน้มทำให้ผลตอบแทนจริงถูกกลดทอนจากต้นทุนธุรกรรม (transaction costs) และผลกระทบจากการ

ซื้อขาย (market impact) นอกจากนี้ค่า Margin เท่ากับ 1.17‰ อยู่ในระดับต่ำ สื่อถึงกำไรต่อหน่วยการซื้อขายที่จำกัด และเพิ่มความเสี่ยงที่ผลตอบแทนเชิงทฤษฎีจะไม่คงอยู่หลังหักค่าใช้จ่ายในการดำเนินกลยุทธ์

ความสม่ำเสมอของผลลัพธ์รายปี เมื่อพิจารณาแยกตามปี (2019–2023) พบว่าผลการดำเนินงานมีความแปรผันข้ามระบอบตลาด (regime dependence) อย่างชัดเจน โดยปี 2020 ให้สมรรถนะดีที่สุด (Sharpe = 1.66, Returns = 15.82%, Drawdown = 4.18%, Margin = 4.25‰) ขณะที่ปี 2022 มีสมรรถนะอ่อนที่สุด (Sharpe = 0.28) และมี Drawdown สูงมากถึง 19.38% ซึ่งสะท้อนความเปราะบางของสัญญาณต่อการเปลี่ยนแปลงสถานะตลาด และชี้ว่าความได้เปรียบของสัญญาณอาจไม่เสถียรในทุกช่วงเวลา ทั้งนี้จำนวนหลักทรัพย์ฝั่ง Long และ Short ในแต่ละปีอยู่ในช่วงประมาณ 275–307 หลักทรัพย์ต่อฝั่ง แสดงว่าในเชิงจำนวนไม่ได้กระจุกตัวมากนัก อย่างไรก็ตาม ความไม่สม่ำเสมอของ Sharpe และระดับ Drawdown ที่สูงในบางปี ยังคงเป็นข้อจำกัดเชิงสาระสำคัญ

ข้อสรุปเชิงนโยบายและประเด็นเพื่อการปรับปรุง โดยสรุป อัลฟานี้ให้ผลตอบแทนรวมเป็นบวกในช่วง TEST แต่มีข้อจำกัดด้าน (i) ผลตอบแทนต่อความเสี่ยงที่ไม่สูง (Sharpe ต่ำ), (ii) ภาระการซื้อขายสูงร่วมกับ Margin ต่ำซึ่งลดทอนความเป็นไปได้เชิงปฏิบัติ และ (iii) ความไม่คงทนของผลลัพธ์เมื่อพิจารณาข้ามปี โดยเฉพาะภาวะ Drawdown ที่รุนแรงในบางช่วงเวลา ดังนั้น การพัฒนาต่อไปควรมุ่งลด Turnover และเพิ่มความทนทานของสัญญาณต่อ regime change รวมถึงประเมินผลหลังหักต้นทุนธุรกรรมอย่างเป็นระบบเพื่อยืนยันความมีนัยเชิงเศรษฐศาสตร์ของอัลฟา



รูปที่ 1: ผลการทดสอบ Alpha ใน WorldQuantBrain จากสมการ 5.1

5.2 สัญญาณความแข็งแกร่งของแท่งเทียนภายในช่วงวัน (Normalized Intraday Candle Strength)

$$\left((\text{close} - \text{open}) / ((\text{high} - \text{low}) + .001) \right) \quad (5.4)$$

ในส่วนนี้นำเสนอแนวคิดเชิงสมมติฐานของสัญญาณแบบ WorldQuant alpha ที่นิยามโดย

$$\alpha_{i,t} = \frac{C_{i,t} - O_{i,t}}{(H_{i,t} - L_{i,t}) + \varepsilon}, \quad \varepsilon = 0.001, \quad (5.5)$$

โดยที่ $O_{i,t}$ คือราคาเปิด (open), $H_{i,t}$ คือราคาสูงสุด (high), $L_{i,t}$ คือราคาต่ำสุด (low) และ $C_{i,t}$ คือราคาปิด (close) ของหลักทรัพย์ i ณ วัน t . ค่าคงที่ ε ถูกเพิ่มในตัวส่วนเพื่อหลีกเลี่ยงการหารด้วยศูนย์หรือค่าช่วงวัน (range) ที่เล็กมากจนทำให้สัญญาณผันผวนผิดปกติในเชิงตัวเลข

การตีความสัญญาณ สัญญาณนี้สามารถตีความเป็น ผลตอบแทนภายในวันแบบทำให้เป็นสัดส่วนด้วยช่วงการแกว่งของราคา (normalized intraday return) กล่าวคือ

- ตัวเศษ $C_{i,t} - O_{i,t}$ วัด ทิศทางและขนาด ของการเปลี่ยนแปลงภายในวัน: บวกเมื่อปิดสูงกว่าเปิด (แท่งเขียว), ลบเมื่อปิดต่ำกว่าเปิด (แท่งแดง)
- ตัวส่วน $H_{i,t} - L_{i,t}$ วัด ความผันผวนภายในวัน หรือช่วงการแกว่งของราคา

ดังนั้น $\alpha_{i,t}$ จะมีความหมายเป็น “ความแข็งแกร่งของแท่งเทียน” เมื่อเทียบกับช่วงการแกว่ง:

1. หาก $\alpha_{i,t}$ มีค่า ใกล้ +1 หมายถึงราคาปิดอยู่ใกล้ด้านบนของช่วงวันและปิดสูงกว่าเปิดอย่างมีนัยสำคัญเมื่อเทียบกับ range $\Rightarrow$ แรงซื้อภายในวันเด่นชัด
2. หาก $\alpha_{i,t}$ มีค่า ใกล้ -1 หมายถึงราคาปิดอยู่ใกล้ด้านล่างของช่วงวันและปิดต่ำกว่าเปิดอย่างมีนัยสำคัญเมื่อเทียบกับ range $\Rightarrow$ แรงขายภายในวันเด่นชัด
3. หาก $\alpha_{i,t}$ มีค่า ใกล้ 0 อาจเกิดจาก (i) ปิดใกล้เปิด (intraday net change ต่ำ) หรือ (ii) มี range สูงแต่ปิดกลับมาใกล้เปิด $\Rightarrow$ ความไม่ชัดเจนของทิศทาง (indecision)

ในเชิงขอบเขต เมื่อ $O_{i,t}, C_{i,t} \in [L_{i,t}, H_{i,t}]$ จะได้ว่าโดยทั่วไป

$$-1 \lesssim \alpha_{i,t} \lesssim 1 \quad (5.6)$$

(ค่าจริงอาจเบี่ยงเล็กน้อยจากการใช้ ε)

สมมติฐาน (Hypotheses) เพื่อทำให้การทดสอบเป็นเชิงวิชาการ สามารถตั้งสมมติฐานเชิงกลไกได้ดังนี้

H1 (Continuation เมื่อแท่งเทียนแข็งแกร่ง): เมื่อ $\alpha_{i,t}$ มีค่าสูง (แท่งเขียวแข็งแกร่งเมื่อเทียบกับ range) แสดงถึงแรงซื้อที่สามารถ “ปิดใกล้ปลายบน” ของช่วงวัน ซึ่งอาจสะท้อนแรงขับเคลื่อนของแนวโน้มระยะสั้น (short-term trend/price pressure) ทำให้คาดหมายผลตอบแทนวันถัดไปเป็นบวก

$$\mathbb{E}[R_{i,t+1} \mid \alpha_{i,t} \text{ สูง}] > 0. \quad (5.7)$$

ข้อสังเกตเชิงปฏิบัติ

- ค่าคงที่ $\varepsilon = 0.001$ ทำหน้าที่เป็น ตัวกันระเบิดเชิงตัวเลข (numerical stabilizer) โดยเฉพาะกรณี $H_{i,t} \approx L_{i,t}$ เช่น วันที่สภาพคล่องต่ำหรือหุ้นที่ไม่ค่อยเคลื่อนไหว
- สัญญาณนี้เป็น *scale-free* ในระดับหนึ่ง เพราะวัดความแข็งแกร่งของการปิดเทียบกับช่วงวัน จึงเหมาะกับการจัดอันดับ (ranking) ข้ามหุ้นหลายราคา

สรุปเชิงนัย โดยสรุป Alpha#101 วัด “ความแรงของการเคลื่อนไหวภายในวัน” ผ่านผลต่างปิด-เปิด แล้วทำให้เป็นสัดส่วนด้วยช่วงสูง-ต่ำของวัน จึงตีความได้ว่าเป็นตัวชี้วัดแรงซื้อ/แรงขายที่เกิดขึ้นจริงภายในกรอบความผันผวนของวันนั้น ๆ และสามารถนำไปทดสอบได้ทั้งมุมมอง *momentum* (H1) และ *reversal เมื่อสุดโต่ง* (H2) ภายใต้ข้อจำกัดด้านความไวต่อ regime ของ range และสัญญาณรบกวนในช่วงที่ความผันผวนต่ำ (H3)



รูปที่ 2: ผลการทดสอบ Alpha ใน WorldQuantBrain จากสมการ 5.4

5.3 สัญญาณกลับทิศของความสัมพันธ์ราคาเปิดกับปริมาณซื้อขายระยะสั้น (Contrarian Open-Volume Correlation, 10 วัน)

$$(-1 * \text{correlation}(\text{open}, \text{volume}, 10)) \quad (5.8)$$

ในส่วนนี้นำเสนอแนวคิดเชิงสมมติฐานของสัญญาณแบบ WorldQuant alpha ที่นิยามโดย

$$\alpha_{i,t} = -\rho_{i,t}^{(10)}(O, V), \quad \rho_{i,t}^{(10)}(O, V) = \text{corr}(O_{i,t-k}, V_{i,t-k})_{k=0}^9, \quad (5.9)$$

โดยที่ $O_{i,t}$ คือราคาเปิด (open) และ $V_{i,t}$ คือปริมาณการซื้อขาย (volume) ของหลักทรัพย์ i ณ วัน t . สัญลักษณ์ $\text{corr}(\cdot, \cdot)$ หมายถึงค่าสหสัมพันธ์เพียร์สัน (Pearson correlation) ที่คำนวณบนหน้าต่างเวลา 10 วันล่าสุด และมีค่าช่วง $[-1, 1]$.

การตีความสัญญาณ แกนกลางของสัญญาณคือการวัดว่า “ในช่วง 10 วันที่ผ่านมา ราคาเปิดกับปริมาณซื้อขายเคลื่อนไหวไปด้วยกันแค่ไหน” แล้ว *กลับเครื่องหมาย* ด้วยการคูณ -1 ดังนั้น

1. หาก $\rho_{i,t}^{(10)}(O, V) > 0$ (ราคาเปิดสูงขึ้นมักมาพร้อมปริมาณซื้อขายสูงขึ้น) จะได้ $\alpha_{i,t} < 0$
ซึ่งตีความว่า ความสัมพันธ์เชิงบวกดังกล่าวอาจสะท้อนแรงไต่ราคา/ความร้อนแรง (crowding, attention) ที่มีโอกาสถดถอยในอนาคต (contrarian ต่อแรงซื้อขาย)
2. หาก $\rho_{i,t}^{(10)}(O, V) < 0$ (ราคาเปิดสูงขึ้นแต่वलุ่มลดลง หรือราคาเปิดลดลงแต่वलุ่มเพิ่มขึ้น) จะได้ $\alpha_{i,t} > 0$
ซึ่งตีความว่า ตลาดอาจอยู่ในภาวะ “ไม่สอดคล้อง” ระหว่างราคาและสภาพคล่อง (price-liquidity divergence) และการกลับเครื่องหมายทำให้สัญญาณเอนเอียงไปทางการคาดหวังการปรับกลับ (reversion) หรือการแก้ความไม่สมดุลในระยะสั้น

กล่าวโดยสรุป สัญญาณนี้เป็น *สัญญาณกลับทิศของความสัมพันธ์ราคา-वलุ่ม* (contrarian to open-volume comovement): เมื่อราคาเปิดและवलุ่ม “ไปด้วยกันมาก” สัญญาณจะลดลง และเมื่อ “สวนกัน” สัญญาณจะเพิ่มขึ้น

สมมติฐาน (Hypotheses) เพื่อให้การทดสอบเป็นเชิงวิชาการ สามารถตั้งสมมติฐานเชิงกลไกได้ดังนี้

H1 (Overheating/Crowding เมื่อสหสัมพันธ์บวกสูง): เมื่อ $\rho_{i,t}^{(10)}(O, V)$ สูงมาก การที่ราคาเปิดขยับขึ้นพร้อมवलุ่มที่เพิ่มขึ้นอาจสะท้อนแรงซื้อแบบไต่ราคา (demand shock) หรือความสนใจเกินปกติ (attention-driven trading) ซึ่งมีแนวโน้มเกิดภาวะตอบสนองเกิน (overreaction) และตามด้วยการย่อตัว

$$\mathbb{E}[R_{i,t+1} \mid \rho_{i,t}^{(10)}(O, V) \text{ สูง}] \propto -\rho_{i,t}^{(10)}(O, V). \quad (5.10)$$

ข้อสังเกตเชิงปฏิบัติ

- การใช้ `correlation( , , 10)` ทำให้สัญญาณเป็นแบบ “สถิติหน้าต่างสั้น” ซึ่งมักตอบสนองเร็ว แต่มีโอกาสเกิด noise สูง
- การใช้ราคา `open` ทำให้สัญญาณไวต่อ *overnight information* และ *opening auction* ซึ่งอาจเพิ่มทั้งโอกาสและความเสี่ยงจาก gap
- การกลับเครื่องหมาย (-1) เป็นการกำหนดให้สัญญาณมีธรรมชาติ *contrarian* ต่อการเคลื่อนไหวร่วมกันของราคาเปิดกับवलุ่ม

สรุปเชิงนัย โดยสรุป Alpha#6 วัดความสัมพันธ์ระยะสั้น 10 วันระหว่างราคาเปิดและปริมาณซื้อขาย แล้วกลับทิศเพื่อให้เป็นสัญญาณเชิงสวนทาง (contrarian) ต่อภาวะที่ราคาและवलุ่มเคลื่อนไหวไปด้วยกันอย่างเด่นชัด โดยตั้งอยู่บนแนวคิดที่ว่า “ความร้อนแรง/การไต่ราคา” มักตามด้วยการถดถอยของราคา ขณะที่ “ความไม่สอดคล้องราคา-สภาพคล่อง” มักนำไปสู่การปรับกลับในระยะสั้น



รูปที่ 3: ผลการทดสอบ Alpha ใน WorldQuantBrain จากสมการ 5.8

5.4 สัญญาณอันดับของส่วนต่างราคาเปิด-ปิดแบบกลับทิศ (Ranked Open–Close Deviation, Contrarian)

$$\text{rank}((-1 * ((1 - (\text{open} / \text{close})))1))) \quad (5.11)$$

ในส่วนนี้นำเสนอแนวคิดเชิงสมมติฐานของสัญญาณแบบ WorldQuant alpha ที่นิยามโดย

$$\alpha_{i,t} = \text{rank}\left(-\left(1 - \frac{O_{i,t}}{C_{i,t}}\right)\right), \quad (5.12)$$

โดยที่ $O_{i,t}$ คือราคาเปิด (open) และ $C_{i,t}$ คือราคาปิด (close) ของหลักทรัพย์ i ณ วัน t . ทั้งนี้เลขชี้กำลัง $(\cdot)^1$ ไม่มีผลต่อค่าเชิงตัวเลข จึงเท่ากับนิพจน์เดิม

การจัดรูปสมการเพื่อความเข้าใจ สามารถจัดรูปนิพจน์ภายใน **rank** ได้ดังนี้

$$-\left(1 - \frac{O_{i,t}}{C_{i,t}}\right) = -1 + \frac{O_{i,t}}{C_{i,t}} = \frac{O_{i,t}}{C_{i,t}} - 1 = \frac{O_{i,t} - C_{i,t}}{C_{i,t}}. \quad (5.13)$$

ดังนั้นสัญญาณเทียบเท่ากับ

$$\alpha_{i,t} = \text{rank}\left(\frac{O_{i,t} - C_{i,t}}{C_{i,t}}\right), \quad (5.14)$$

ซึ่งตีความได้ว่าเป็น ผลต่างเปิด-ปิดแบบทำให้เป็นสัดส่วนด้วยราคาปิด แล้วนำไป จัดอันดับข้ามหุ้น ณ วันเดียวกัน

การตีความสัญญาณ นิพจน์ $\frac{O_{i,t} - C_{i,t}}{C_{i,t}}$ มีสัญญาณสัมพันธ์กับการเปลี่ยนแปลงภายในวันดังนี้

- ถ้า $C_{i,t} > O_{i,t}$ (วันบวก/แท่งเขียว) $\Rightarrow O_{i,t} - C_{i,t} < 0$ ทำให้ค่าภายใน **rank** เป็นลบ
- ถ้า $C_{i,t} < O_{i,t}$ (วันลบ/แท่งแดง) $\Rightarrow O_{i,t} - C_{i,t} > 0$ ทำให้ค่าภายใน **rank** เป็นบวก

เมื่อใช้ $\text{rank}(\cdot)$ สัญญาณจะกลายเป็น สัญญาณแบบสัมพันธ์ (cross-sectional): หุ้นที่มีค่า $\frac{O-C}{C}$ สูง (มักเป็นหุ้นที่ “ลงแรง” ภายในวัน) จะได้ **อันดับสูง**; หุ้นที่มีค่า ต่ำ (มักเป็นหุ้นที่ “ขึ้นแรง” ภายในวัน) จะได้ **อันดับต่ำ**. ในเชิงกลยุทธ์ หากนำอันดับไปใช้สร้างพอร์ตแบบ long-short โดยทั่วไปจะตีความเป็นแนว *contrarian* ระยะสั้น คือ

1. หุ่นที่ลงแรงในวัน t (อันดับสูง) มีโอกาส *ดิ่งกลับ* ใน $t + 1 \Rightarrow$ พิจารณาฝั่ง Long
2. หุ่นที่ขึ้นแรงในวัน t (อันดับต่ำ) มีโอกาส *ย่อตัว* ใน $t + 1 \Rightarrow$ พิจารณาฝั่ง Short

ทั้งนี้ทิศทาง Long/Short จริงขึ้นกับ convention ของแพลตฟอร์ม (เช่น long rank สูงหรือ long rank ต่ำ) แต่โครงสร้างสัญญาณนี้ชี้ไปทาง “วัดการลงแรงเมื่อเทียบกับราคาปิด” เป็นหลัก

สมมติฐาน (Hypotheses) เพื่อให้การทดสอบเป็นเชิงวิชาการ สามารถตั้งสมมติฐานเชิงกลไกได้ดังนี้

H1 (Reversal จากแรงขายเกินภายในวัน): เมื่อ $\frac{O_{i,t}-C_{i,t}}{C_{i,t}}$ สูงมาก (ราคาปิดต่ำกว่าเปิดมาก) อาจสะท้อนแรงขายเชิงสภาพคล่องหรือภาวะตอบสนองเกิน (overreaction) ในระยะสั้น ทำให้คาดการณ์การกลับตัวในวันถัดไป

$$\mathbb{E}[R_{i,t+1} \mid \alpha_{i,t} \text{ สูง}] > 0. \quad (5.15)$$

ข้อสังเกตเชิงปฏิบัติ

- **rank** ลดผลกระทบจาก outliers และทำให้สัญญาณเหมาะสมกับพอร์ตแบบ cross-sectional long-short
- นิพจน์นี้ใช้ข้อมูลวันเดียว (ไม่ต้องพึ่ง lookback) จึงอาจมี Turnover สูงเมื่อใช้งานจริง หากรีบาลานซ์ทุกวัน
- กรณี $C_{i,t}$ ใกล้ศูนย์ (แทบไม่เกิดในหุ่นทั่วไป) อาจทำให้สัดส่วนผิดปกติ แต่ในทางปฏิบัติราคาหุ้นมีค่าบวกและไม่เล็กมาก

สรุปเชิงนัย โดยสรุป Alpha#33 เป็นสัญญาณที่วัด “การลง/ขึ้นภายในวัน” ผ่านสัดส่วน $\frac{O-C}{C}$ แล้วจัดอันดับข้ามหุ่น เพื่อจับความเป็นไปได้ของ *mean-reversion ระยะสั้น* ในเชิง cross-sectional โดยเฉพาะการคาดหวังการดิ่งกลับของหุ่นที่ปิดต่ำกว่าเปิดมากเมื่อเทียบกับหุ่นอื่น ๆ ในวันเดียวกัน



รูปที่ 4: ผลการทดสอบ Alpha ใน WorldQuantBrain จากสมการ 5.11

6 วิธีการดำเนินงาน

งานวิจัยนี้ดำเนินการตามลำดับขั้นตอนของการพัฒนากลยุทธ์เชิงปริมาณ โดยเริ่มจากการรวบรวมข้อมูลตลาดในรูปแบบอนุกรมเวลา จากนั้นจึงนำข้อมูลดังกล่าวมาจัดโครงสร้างและแปลงให้อยู่ในรูปแบบที่เหมาะสมต่อการสร้างตัวแปรเชิงเทคนิคและสูตรแอลฟา ขั้นตอนทั้งหมดถูกออกแบบให้สามารถทำซ้ำได้อย่างเป็นระบบ เพื่อให้ผลการทดลองมีความโปร่งใส ตรวจสอบได้ และสามารถนำไปต่อยอดในงานวิจัยเชิงประยุกต์ได้ในภายหลัง

6.1 รวบรวมข้อมูล OHLCV

ขั้นตอนแรกของงานวิจัยคือการรวบรวมข้อมูลตลาดในรูปแบบ OHLCV ซึ่งประกอบด้วยราคาเปิด (*Open*) ราคาสูงสุด (*High*) ราคาต่ำสุด (*Low*) ราคาปิด (*Close*) และปริมาณการซื้อขาย (*Volume*) ของหลักทรัพย์แต่ละตัว ข้อมูลประเภทนี้เป็นข้อมูลพื้นฐานที่มีความสำคัญอย่างยิ่งสำหรับการวิเคราะห์เชิงปริมาณ เนื่องจากสามารถใช้เป็นรากฐานในการคำนวณผลตอบแทน ความผันผวน โมเมนตัม สัญญาณกลับตัว ตลอดจนรูปแบบแท่งเทียนและตัวชี้วัดทางเทคนิคอื่น ๆ ได้อย่างครอบคลุม

ในการศึกษาครั้งนี้ ผู้วิจัยเลือกใช้ข้อมูล OHLCV ของหลักทรัพย์ที่อยู่ในดัชนี S&P 500 เป็นตัวแทนโดยประมาณ (*proxy*) ของกลุ่มสินทรัพย์ที่ใช้ในแพลตฟอร์ม WorldQuant BRAIN การเลือกใช้ S&P 500 มีเหตุผลสำคัญในด้านความหลากหลายของอุตสาหกรรม ขนาดของกลุ่มการลงทุน และสภาพคล่องของหลักทรัพย์ ซึ่งทำให้ชุดข้อมูลดังกล่าวมีความเหมาะสมสำหรับใช้เป็นฐานในการศึกษาพฤติกรรมของสูตรแอลฟาในบริบทของการลงทุนเชิงระบบ แม้ว่ากลุ่มสินทรัพย์จริงในแพลตฟอร์ม WorldQuant BRAIN อาจมีขอบเขตและเงื่อนไขเฉพาะที่แตกต่างออกไป แต่การใช้ S&P 500 ในงานวิจัยนี้ช่วยให้สามารถสร้างกรอบการทดลองที่มีความสมเหตุสมผล โปร่งใส และสามารถทำซ้ำได้

ช่วงเวลาของข้อมูลที่ใช้ในการศึกษาครอบคลุมตั้งแต่ปี ค.ศ. 2019 ถึง 2024 โดยกำหนดให้เป็นช่วงสำหรับการฝึกและการทดสอบแบบจำลอง (*training and testing period*) บนแพลตฟอร์ม ช่วงเวลาดังกล่าวครอบคลุมทั้งภาวะตลาดปกติ ภาวะความผันผวนสูง และช่วงการฟื้นตัวของตลาด จึงเหมาะสมสำหรับใช้ประเมินความเสถียรและความทนทานของสูตรแอลฟาในสภาวะตลาดที่แตกต่างกัน การกำหนดกรอบเวลานี้ช่วยให้การประเมินผลไม่จำกัดอยู่เพียงช่วงตลาดใดช่วงหนึ่ง และเพิ่มความน่าเชื่อถือของข้อสรุปที่ได้จากการทดลอง

ในงานวิจัยนี้ การดึงข้อมูล OHLCV ดำเนินการด้วยโปรแกรมภาษา Python โดยพัฒนาเป็นสคริปต์สำหรับเชื่อมต่อแหล่งข้อมูล จัดรูปแบบข้อมูล และบันทึกผลลัพธ์ให้อยู่ในรูปแบบที่พร้อมใช้งานในขั้นตอนถัดไป โค้ดโปรแกรมดังกล่าวถูกจัดเก็บไว้บน GitHub เพื่อสนับสนุนหลักการท้าววิจัยที่สามารถตรวจสอบและทำซ้ำได้ (*reproducibility*) การใช้ Python ในขั้นตอนนี้มีข้อได้เปรียบสำคัญในด้านความยืดหยุ่นของการประมวลผลข้อมูล ความสามารถในการจัดการข้อมูลอนุกรมเวลาขนาดใหญ่ และความสะดวกในการเชื่อมโยงกับไลบรารีด้านการวิเคราะห์เชิงการเงินและสถิติ

หลังจากดึงข้อมูลแล้ว ข้อมูลจะถูกตรวจสอบความสมบูรณ์และความสอดคล้องของอนุกรมเวลา เช่น การตรวจสอบค่าว่าง ความต่อเนื่องของวันซื้อขาย และความสมเหตุสมผลของช่วงราคาในแต่ละวัน ขั้นตอนนี้มีความสำคัญเนื่องจากคุณภาพของข้อมูลต้นทางส่งผลโดยตรงต่อความน่าเชื่อถือของสูตรแอลฟาและผลการทดสอบในลำดับถัดไป หากข้อมูลมีความคลาดเคลื่อน เช่น ราคาขาดช่วง ปริมาณซื้อขายผิดปกติ หรือการจัดเรียงวันเวลาไม่ถูกต้อง อาจทำให้สัญญาณที่สร้างขึ้นมีอคติและนำไปสู่ข้อสรุปที่คลาดเคลื่อนได้

ภายหลังการตรวจสอบ ข้อมูล OHLCV จะถูกจัดเก็บในโครงสร้างแบบตารางเวลา (*time-series table*) เพื่อ

ให้ง่ายต่อการเรียกใช้งานและการสร้างตัวแปรอนุพันธ์ เช่น ผลตอบแทนย้อนหลัง ค่าเฉลี่ยเคลื่อนที่ ส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่ ช่วงการเคลื่อนไหวของราคา และตัววัดเชิงเทคนิคอื่น ๆ กล่าวได้ว่า ขั้นตอนการรวบรวมข้อมูล OHLCV ไม่ได้เป็นเพียงการดาวน์โหลดข้อมูลเท่านั้น แต่เป็นกระบวนการเตรียมข้อมูลให้พร้อมสำหรับการสร้างแบบจำลองเชิงปริมาณอย่างมีระบบ

6.2 สร้างสูตรแอลฟา

หลังจากเตรียมข้อมูล OHLCV เรียบร้อยแล้ว ขั้นตอนถัดไปของงานวิจัยคือการสร้างสูตรแอลฟา (*alpha formulas*) เพื่อใช้เป็นสัญญาณเชิงปริมาณสำหรับอธิบายความแตกต่างของลักษณะราคาและปริมาณการซื้อขายระหว่างหลักทรัพย์แต่ละตัว ในการศึกษาครั้งนี้ ผู้วิจัยออกแบบชุดสูตรแอลฟาโดยอาศัยข้อมูล OHLCV และจัดโครงสร้างสูตรให้ครอบคลุมมิติสำคัญของพฤติกรรมตลาดที่สามารถสะท้อนออกมาในข้อมูลราคาได้โดยตรง ทั้งนี้ สูตรที่ใช้ในการวิเคราะห์ประกอบด้วยทั้งหมด 81 สูตร แบ่งออกเป็น 9 หมวดหมู่หลัก โดยในแต่ละหมวดหมู่ประกอบด้วย 9 สูตร เพื่อให้สามารถเปรียบเทียบผลลัพธ์ได้อย่างเป็นระบบทั้งในระดับรายสูตรและระดับกลุ่มแนวคิด

หมวดหมู่แรกเป็นสูตรที่สะท้อนผลตอบแทนย้อนหลังและโมเมนตัม (*Returns and Momentum*) ซึ่งครอบคลุมทั้งผลตอบแทนย้อนหลังในหลายช่วงเวลาและผลรวมสะสมของผลตอบแทนในหน้าต่างเวลาที่แตกต่างกัน สูตรในกลุ่มนี้มีจุดมุ่งหมายเพื่อจับแรงส่งของราคาและลักษณะการเคลื่อนไหวสะสมของหลักทรัพย์ในระยะสั้นถึงระยะกลาง โดยตั้งอยู่บนแนวคิดว่าหุ้นที่มีพฤติกรรมราคาแข็งแกร่งอย่างต่อเนื่องอาจมีลักษณะแตกต่างจากหุ้นที่มีผลตอบแทนอ่อนแออย่างเป็นระบบ

หมวดหมู่ที่สองเป็นสูตรที่เกี่ยวข้องกับการกลับเข้าสู่ค่าเฉลี่ย ค่าเฉลี่ยเคลื่อนที่ ช่องว่างของราคา และผลตอบแทนภายในวัน (*Mean Reversion, Moving Average, Gaps, and Intraday Effects*) สูตรในกลุ่มนี้ถูกออกแบบเพื่อสะท้อนแนวคิดที่ว่าราคาอาจเบี่ยงเบนออกจากระดับสมดุลในระยะสั้นก่อนจะกลับเข้าสู่ค่าเฉลี่ยในเวลาต่อมา ขณะเดียวกันยังครอบคลุมพฤติกรรมของช่องว่างระหว่างราคาเปิดกับราคาปิดวันก่อนหน้า ตลอดจนการเคลื่อนไหวของราคาในระหว่างวัน ซึ่งอาจแฝงข้อมูลเกี่ยวกับแรงซื้อขายระยะสั้นและพฤติกรรมของนักลงทุนในตลาด

หมวดหมู่ที่สามเป็นสูตรที่มุ่งวัดความผันผวน ช่วงการเคลื่อนไหวของราคา ค่าช่วงราคาจริง และการจัดอันดับเชิงเวลา (*Volatility, Range, True Range, and Ranks*) สูตรในกลุ่มนี้ครอบคลุมตัวแปร เช่น ส่วนเบี่ยงเบนมาตรฐานของผลตอบแทนย้อนหลังหลายช่วงเวลา ช่วงราคาสูงสุด-ต่ำสุดเมื่อเทียบกับราคาปิด ค่า *true range* และการจัดอันดับของช่วงราคาในหน้าต่างเวลา จุดมุ่งหมายของกลุ่มนี้คือการตรวจสอบว่าความผันผวนและขนาดของการแกว่งตัวของราคาในอดีตสามารถใช้เป็นสัญญาณจำแนกหลักทรัพย์ได้หรือไม่ โดยตั้งอยู่บนแนวคิดที่ว่าหุ้นที่มีโครงสร้างความผันผวนแตกต่างกันอาจมีแนวโน้มผลตอบแทนที่แตกต่างกันตามไปด้วย

หมวดหมู่ที่สี่เป็นสูตรที่เกี่ยวข้องกับการทะลุกรอบราคา ตัวชี้วัด stochastic และรูปแบบแท่งเทียนเชิงแพตเทิร์น (*Breakouts, Stochastic, and Candlestick Patterns*) สูตรในกลุ่มนี้ครอบคลุมทั้งการวัดตำแหน่งของราคาปิดเมื่อเทียบกับกรอบราคาย้อนหลัง การวัดระยะห่างจากจุดสูงสุดย้อนหลัง ตลอดจนการเข้ารหัสรูปแบบแท่งเทียน เช่น *bullish engulfing*, *bearish engulfing*, *hammer* และ *shooting star* กลุ่มนี้มีจุดมุ่งหมายเพื่อทดสอบว่ารูปแบบเชิงโครงสร้างของราคาและพฤติกรรมการทะลุกรอบในระยะสั้นสามารถแปรสภาพเป็นสัญญาณเชิงปริมาณที่ใช้แยกหุ้นผลตอบแทนสูงและต่ำได้เพียงใด

หมวดหมู่ที่ห้าเป็นสูตรที่มุ่งวัดแนวโน้ม ความชันของราคา และตำแหน่งของราคาปัจจุบันเมื่อเทียบกับกรอบ

ราคาย้อนหลัง (*Trend, Slope, and Price Position*) สูตรในกลุ่มนี้ใช้แนวคิดว่าหลักทรัพย์ที่มีแนวโน้มราคาชัดเจน หรือมีราคาปัจจุบันอยู่ในตำแหน่งที่โดดเด่นเมื่อเทียบกับช่วงราคาย้อนหลัง อาจมีคุณลักษณะที่แตกต่างจากหลักทรัพย์ที่เคลื่อนไหวแบบไร้ทิศทางหรืออยู่ในกรอบราคาที่อ่อนแอกว่า กลุ่มนี้จึงช่วยตรวจสอบว่าข้อมูลด้านแนวโน้มและโครงสร้างตำแหน่งของราคามีศักยภาพในการใช้เป็นสัญญาณเชิงคัดเลือกลักษณะหรือไม

หมวดหมู่ที่หกเป็นสูตรที่เกี่ยวข้องกับปริมาณการซื้อขายในเชิงมูลค่าและแรงกดดันของกระแสเงิน (*Dollar Volume and Flow Pressure*) โดยมุ่งเน้นการวัดมูลค่าการซื้อขาย (*dollar volume*) เมื่อเทียบกับค่าเฉลี่ยย้อนหลัง ความสัมพันธ์ระหว่างราคาปิดกับปริมาณซื้อขาย การไหลของเงินแบบมีทิศทาง (*signed dollar flow*) ตลอดจนการผสมผสานระหว่างผลตอบแทนและปริมาณซื้อขาย กลุ่มนี้ตั้งอยู่บนแนวคิดว่าการเคลื่อนไหวของราคาที่เกิดขึ้นพร้อมกับปริมาณการซื้อขายหรือมูลค่าการซื้อขายที่ผิดปกติอาจสะท้อนแรงซื้อขายที่มีนัยสำคัญมากกว่าการเปลี่ยนแปลงของราคาที่เกิดขึ้นในภาวะปริมาณซื้อขายต่ำ

หมวดหมู่ที่เจ็ดเป็นสูตรที่เกี่ยวข้องกับพลวัตของราคาเปิดและช่องว่างระหว่างวัน (*Overnight and Gap Dynamics*) โดยมุ่งเน้นการวัดความสัมพันธ์ระหว่างราคาเปิดของวันปัจจุบันกับราคาปิด ราคาสูงสุด หรือราคาต่ำสุดของวันก่อนหน้า ตลอดจนการวัดความผันผวนและการจัดอันดับของพฤติกรรมช่องว่างราคาในช่วงเวลาย้อนหลัง สูตรในกลุ่มนี้มีพื้นฐานอยู่บนแนวคิดว่าการเคลื่อนไหวข้ามคืนและช่องว่างของราคาเปิดอาจสะท้อนข้อมูลใหม่ที่ตลาดรับรู้ระหว่างช่วงปิดทำการ และอาจมีความสัมพันธ์กับคุณภาพของผลตอบแทนของหลักทรัพย์ในเชิงภาคตัดขวาง

หมวดหมู่ที่แปดเป็นสูตรที่มุ่งวัดระบอบความผันผวนและการบีบตัวของช่วงราคา (*Volatility Regime and Compression*) โดยครอบคลุมตัวแปรที่เกี่ยวข้องกับความผันผวนของผลตอบแทนในหลายช่วงเวลา อัตราส่วนระหว่างความผันผวนระยะสั้นกับระยะยาว ความกว้างของช่วงราคาสูงสุด-ต่ำสุด ตลอดจนแรงกดดันจากการทะลุกรอบราคา สูตรในกลุ่มนี้ตั้งอยู่บนแนวคิดว่ารูปแบบของความผันผวนและการหดหรือขยายตัวของช่วงราคาอาจสะท้อนการเปลี่ยนแปลงของสถานะตลาดและระดับความไม่แน่นอน ซึ่งสามารถใช้เป็นสัญญาณจำแนกหลักทรัพย์ได้

หมวดหมู่ที่เก้าเป็นสูตรที่เกี่ยวข้องกับโครงสร้างของแท่งเทียน (*Candlestick Body and Wick Structure*) โดยมุ่งเน้นการวัดขนาดของตัวแท่ง สัดส่วนของไส้เทียนบนและล่าง ตำแหน่งของราคาเปิดและราคาปิดภายในกรอบสูงสุด-ต่ำสุดของวัน ตลอดจนรูปแบบแท่งเทียนเฉพาะ เช่น *marubozu* และ *doji* กลุ่มนี้มีพื้นฐานอยู่บนแนวคิดว่ารายละเอียดภายในแท่งราคาแต่ละวันอาจสะท้อนสมดุลของแรงซื้อและแรงขาย รวมถึงพฤติกรรมของตลาดในระดับจุลภาคที่ไม่ปรากฏชัดหากพิจารณาเพียงผลตอบแทนสุทธิของวัน

การจัดสูตรแอลฟาออกเป็น 9 หมวดหมู่ดังกล่าวมีประโยชน์ในเชิงระเบียบวิธี เนื่องจากช่วยให้การวิเคราะห์ผลลัพธ์สามารถดำเนินไปได้อย่างมีโครงสร้างมากขึ้น กล่าวคือ งานวิจัยไม่ได้มุ่งพิจารณาเพียงว่าสูตรใดสูตรหนึ่งให้ผลดีหรือไม่เท่านั้น แต่ยังสามารถวิเคราะห์ต่อได้ว่ากลุ่มแนวคิดใด เช่น ผลตอบถอยหลังและโมเมนตัม การกลับเข้าสู่ค่าเฉลี่ย ความผันผวนและช่วงราคา การทะลุกรอบและรูปแบบแท่งเทียน แนวโน้มของราคา กระแสเงินจากปริมาณซื้อขาย พลวัตของช่องว่างระหว่างวัน ระบอบความผันผวน หรือโครงสร้างของแท่งเทียน มีแนวโน้มสร้างสัญญาณที่สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ดีกว่าในบริบทของข้อมูลที่ศึกษา ด้วยเหตุนี้ ขั้นตอนการสร้างสูตรแอลฟาจึงมิใช่เพียงการสร้างตัวแปรจำนวนมากจากข้อมูล OHLCV แต่เป็นการออกแบบกรอบการทดลองที่เปิดโอกาสให้ตีความผลได้อย่างเป็นระบบทั้งในเชิงข้อมูล เชิงสถิติ และเชิงการลงทุน

6.3 ทดสอบความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำของสูตรแอลฟา

หลังจากสร้างสูตรแอลฟาทั้ง 36 สูตรแล้ว งานวิจัยดำเนินการทดสอบความสามารถของแต่ละสูตรในการแยกความแตกต่างระหว่างหุ้นที่มีผลตอบแทนเฉลี่ยสูงกับหุ้นที่มีผลตอบแทนเฉลี่ยต่ำ แนวคิดหลักของขั้นตอนนี้ตั้งอยู่บนสมมติฐานว่า หากสูตรแอลฟาหนึ่งมีสาระเชิงข้อมูล (*informative content*) จริง สูตรดังกล่าวควรให้ค่าหรือรูปแบบของสัญญาณที่แตกต่างกันอย่างเป็นระบบระหว่างหุ้นที่ในอดีตให้ผลตอบแทนสูงและหุ้นที่ให้ผลตอบแทนต่ำ กล่าวอีกนัยหนึ่ง หากสัญญาณไม่สามารถจำแนกสองกลุ่มนี้ได้ ก็ยากที่จะคาดหวังว่าสัญญาณดังกล่าวจะนำไปสู่ผลการดำเนินงานของพอร์ตที่โดดเด่นในขั้นตอนการลงทุนจริง

ในทางปฏิบัติ งานวิจัยใช้ไฟล์จัดอันดับหุ้นตามผลตอบแทนเฉลี่ยย้อนหลังตั้งแต่ปี 2019 เป็นต้นมา เพื่อแบ่งหลักทรัพย์ออกเป็นสองกลุ่ม ได้แก่ กลุ่มหุ้นผลตอบแทนต่ำจำนวน 250 ตัวล่างสุด และกลุ่มหุ้นผลตอบแทนสูงจำนวน 250 ตัวบนสุด จากนั้นจึงคำนวณค่าของสูตรแอลฟาแต่ละสูตรในระดับรายวันสำหรับหุ้นแต่ละตัว และสรุปเป็นค่าเฉลี่ยของแอลฟารายหลักทรัพย์ (*per-ticker mean(alpha)*) ตลอดช่วงเวลาการศึกษา วิธีการนี้ทำให้แต่ละหุ้นถูกแทนด้วยค่ากลางของสัญญาณในแต่ละสูตร และเปิดโอกาสให้เปรียบเทียบการกระจายของค่าดังกล่าวระหว่างกลุ่มหุ้นผลตอบแทนสูงและต่ำได้อย่างชัดเจน

เกณฑ์การตีความในงานวิจัยนี้อาศัยตรรกะที่ว่า สูตรแอลฟาที่ดีควรสามารถสร้างการแยก (*separation*) ระหว่างสองกลุ่มได้อย่างมีความหมาย หากค่าของแอลฟาในกลุ่มหุ้นผลตอบแทนสูงและกลุ่มหุ้นผลตอบแทนต่ำมีการกระจายแตกต่างกันอย่างเด่นชัด เช่น มีค่ากลางต่างกันหรือมีทิศทางที่สอดคล้องกับสมมติฐานทางการเงิน ก็สามารถมองได้ว่าสูตรดังกล่าวมีศักยภาพในการแยกหุ้นที่ “ดี” ออกจากหุ้นที่ “ด้อย” ในเชิงผลตอบแทน ในทางกลับกัน หากการกระจายของค่าทั้งสองกลุ่มทับซ้อนกันมาก ก็สะท้อนว่าสูตรนั้นอาจมีอำนาจจำแนกต่ำ

จากมุมมองเชิงสถิติ การแยกกลุ่มได้ดีหมายความว่าสัญญาณของแอลฟามีความสัมพันธ์กับโครงสร้างของผลตอบแทนในข้อมูล กล่าวคือ ค่าของแอลฟาไม่ได้สุ่มกระจายโดยไม่มีรูปแบบ แต่มีแนวโน้มปรากฏแตกต่างกันระหว่างหุ้นสองกลุ่มที่นิยามจากผลตอบแทนจริง ส่วนในมุมมองเชิงการเงิน หากแอลฟาสามารถจำแนกหุ้นผลตอบแทนสูงและต่ำได้อย่างสม่ำเสมอ ก็ย่อมมีความเป็นไปได้มากขึ้นที่เมื่อนำแอลฟานั้นไปใช้จัดอันดับหลักทรัพย์และสร้างพอร์ตลงทุน จะสามารถนำไปสู่ตัวชี้วัดผลการดำเนินงานที่ดีกว่า เช่น ผลตอบแทนเฉลี่ยที่สูงกว่า ค่า Sharpe ratio ที่ดีขึ้น หรือค่า Fitness ที่เด่นกว่าเมื่อเทียบกับสูตรที่ไม่สามารถแยกกลุ่มหุ้นได้อย่างชัดเจน

อย่างไรก็ดี ในงานวิจัยนี้ การแยกกลุ่มหุ้นผลตอบแทนสูงและต่ำไม่ได้ถูกตีความว่าเป็นหลักฐานเพียงพอของความสามารถในการทำกำไรโดยตรง แต่ถูกใช้เป็นขั้นตอนคัดกรองเชิงโครงสร้าง (*screening mechanism*) กล่าวคือ สูตรที่แสดงความสามารถในการจำแนกกลุ่มหุ้นได้ดีจะถูกละทิ้งไปเป็นผู้สมัครที่มีศักยภาพสูงกว่าในการนำไปสู่ผลการดำเนินงานที่ดี ขณะที่สูตรซึ่งไม่สามารถแยกสองกลุ่มได้อย่างมีรูปแบบมักมีแนวโน้มให้สัญญาณที่อ่อนหรือไม่มีความหมายเชิงพยากรณ์เพียงพอ

ดังนั้น ขั้นตอนการทดสอบความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำจึงทำหน้าที่เป็นสะพานเชื่อมระหว่างการออกแบบสูตรแอลฟากับการประเมินผลเชิงพอร์ต โดยช่วยให้ผู้วิจัยสามารถพิจารณาได้ตั้งแต่ระดับคุณสมบัติของสัญญาณว่า สูตรใดมีความน่าจะเป็นในการสร้างผลลัพธ์การลงทุนที่ดีมากกว่า ภายใต้ตรรกะพื้นฐานที่ว่า แอลฟาซึ่งสามารถแยกหุ้นที่มีลักษณะผลตอบแทนแตกต่างกันได้ ย่อมมีโอกาสสูงกว่าที่จะสะท้อนออกมาเป็น performance metrics ที่ดีกว่าในขั้นตอนของการจำลองและการสร้างพอร์ตลงทุน

6.4 ทดสอบบน WorldQuant Brain Platform

ภายหลังจากการรวบรวมข้อมูล OHLCV การสร้างสูตรแอลฟา และการตรวจสอบความสามารถของสัญญาณในการแยกหุ้นผลตอบแทนสูงและต่ำแล้ว ขั้นตอนถัดไปของงานวิจัยคือการนำสูตรแอลฟาที่ได้ไปทดสอบบนแพลตฟอร์ม WorldQuant BRAIN เพื่อประเมินศักยภาพของสูตรในบริบทของการสร้างพอร์ตการลงทุนเชิงระบบ การทดสอบในขั้นตอนนี้มีความสำคัญอย่างยิ่ง เนื่องจากแม้สูตรแอลฟาจะมีคุณสมบัติในการจำแนกหลักทรัพย์ได้ดีในระดับสัญญาณ แต่การนำไปใช้จริงในเชิงพอร์ตยังต้องพิจารณากระบวนการจัดอันดับ การปรับให้อยู่ในรูปแบบที่เปรียบเทียบกันได้ การทำให้สัญญาณเป็นกลาง การกำหนดน้ำหนักลงทุน และผลลัพธ์สุดท้ายในรูปของผลตอบแทนและความเสี่ยงของพอร์ตโดยรวม

ในการศึกษานี้ ผู้วิจัยนำสูตรแอลฟาที่สร้างขึ้นจากข้อมูล OHLCV ของหลักทรัพย์ในดัชนี S&P 500 ซึ่งใช้เป็นตัวแทนโดยประมาณของกลุ่มสินทรัพย์ใน WorldQuant BRAIN ไปเขียนให้อยู่ในรูปแบบที่แพลตฟอร์มสามารถประมวลผลได้ จากนั้นจึงดำเนินการจำลองผลการลงทุนย้อนหลัง (*backtesting/simulation*) บนแพลตฟอร์มในช่วงเวลาที่กำหนด โดยครอบคลุมข้อมูลระหว่างปี 2019 ถึง 2024 ซึ่งทำหน้าที่เป็นทั้งช่วงสำหรับการเรียนรู้รูปแบบของข้อมูลและช่วงสำหรับการทดสอบความทนทานของสูตรภายใต้สภาวะตลาดที่แตกต่างกัน

กระบวนการทดสอบบนแพลตฟอร์ม WorldQuant BRAIN มิได้เป็นเพียงการคำนวณค่าของสูตรในแต่ละวันเท่านั้น แต่เป็นการนำสูตรแอลฟาเข้าสู่กลไกการจำลองพอร์ตอย่างครบถ้วน กล่าวคือ เมื่อระบบได้รับนิพจน์ของแอลฟาแล้ว แพลตฟอร์มจะคำนวณค่าสัญญาณของหลักทรัพย์แต่ละตัวในวันซื้อขาย จากนั้นทำการจัดอันดับเชิงภาคตัดขวาง ปรับสัญญาณให้อยู่ในสเกลที่สามารถเปรียบเทียบกันได้ และแปลงสัญญาณดังกล่าวให้เป็นน้ำหนักการลงทุนสำหรับพอร์ตแบบ long-short ภายใต้ข้อกำหนดของแพลตฟอร์ม กระบวนการนี้ทำให้การประเมินสูตรแอลฟาอยู่ในกรอบที่มีมาตรฐานเดียวกัน และช่วยลดความคลาดเคลื่อนที่อาจเกิดจากการออกแบบวิธีคำนวณพอร์ตที่แตกต่างกันเองในแต่ละสูตร

การประเมินผลของสูตรแอลฟาบนแพลตฟอร์มอาศัยตัวชี้วัดผลการดำเนินงานหลายประเภทเพื่อสะท้อนคุณภาพของสูตรในหลายมิติพร้อมกัน ตัวชี้วัดหลักที่ใช้ในการศึกษานี้ประกอบด้วย *Sharpe ratio*, *Turnover*, *Fitness*, *Returns*, *Drawdown* และ *Margin* โดยแต่ละตัวชี้วัดมีบทบาทในการอธิบายมิติที่แตกต่างกันของกลยุทธ์ กล่าวคือ *Sharpe ratio* ใช้สะท้อนผลตอบแทนต่อหน่วยความเสี่ยง *Returns* ใช้วัดความสามารถในการสร้างผลตอบแทนโดยรวม *Drawdown* ใช้แสดงระดับการขาดทุนสะสมสูงสุด *Turnover* ใช้บ่งชี้ความถี่ในการเปลี่ยนพอร์ต และ *Margin* ใช้วัดความสามารถในการสร้างผลตอบแทนต่อหน่วยของการซื้อขาย ขณะที่ *Fitness* ทำหน้าที่เป็นตัวชี้วัดเชิงรวมที่สะท้อนคุณภาพของสูตรในลักษณะที่สัมพันธ์ทั้งกับผลตอบแทน ความเสี่ยง และความคุ้มค่าของการนำไปใช้จริง

ในเชิงระเบียบวิธี การใช้แพลตฟอร์ม WorldQuant BRAIN เป็นสภาพแวดล้อมในการทดสอบมีข้อได้เปรียบสำคัญ เนื่องจากช่วยให้สูตรแอลฟาทุกสูตรถูกประเมินภายใต้เงื่อนไขการจำลองเดียวกัน ผู้วิจัยจึงสามารถเปรียบเทียบประสิทธิภาพของสูตรในแต่ละหมวดหมู่ได้อย่างเป็นธรรม ไม่ว่าจะเป็นกลุ่ม *momentum*, *mean reversion*, *volatility*, *volume-related factors* หรือ *candlestick patterns* การเปรียบเทียบเช่นนี้ทำให้สามารถตอบคำถามวิจัยได้ชัดเจนยิ่งขึ้นว่า แนวคิดของสูตรประเภทใดมีแนวโน้มสร้างสัญญาณที่สามารถนำไปใช้ในเชิงพอร์ตได้ดีกว่าในช่วงเวลาที่ศึกษา

นอกจากนี้ การทดสอบบนแพลตฟอร์มยังช่วยทำหน้าที่เป็นขั้นตอนตรวจสอบความแข็งแกร่งของสูตรแอลฟาในระดับการนำไปใช้งานจริง กล่าวคือ สูตรบางตัวอาจดูมีความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำได้

ในระดับสัญญาณ แต่เมื่อนำเข้าสู่กระบวนการสร้างพอร์ต อาจให้ผลลัพธ์ที่ไม่ดีนักเนื่องจากมี turnover สูงเกินไป มี drawdown สูง หรือไม่สามารถรักษาคุณภาพของผลตอบแทนในระดับพอร์ตได้ ในทางกลับกัน สูตรที่สามารถรักษาความสามารถเชิงสัญญาณไว้ได้เมื่อผ่านกระบวนการจำลองพอร์ตบน WorldQuant BRAIN ย่อมมีความน่าเชื่อถือมากกว่าในฐานะผู้สมัครสำหรับการใช้งานจริง

กล่าวโดยสรุป ขั้นตอนการทดสอบบน WorldQuant BRAIN Platform เป็นกระบวนการที่เชื่อมโยงระหว่าง การออกแบบสูตรแอลฟาในเชิงแนวคิดกับการประเมินศักยภาพของสูตรในเชิงปฏิบัติ โดยช่วยให้ผู้วิจัยสามารถ พิจารณาได้ว่าสูตรใดไม่เพียงมีความสามารถในการแยกความแตกต่างของหลักทรัพย์ในระดับข้อมูล แต่ยังสามารถ แปรสภาพเป็นพอร์ตลงทุนที่ให้ผลลัพธ์ดีภายใต้เกณฑ์มาตรฐานของแพลตฟอร์มได้จริง ด้วยเหตุนี้ ผลการทดสอบ บน WorldQuant BRAIN จึงเป็นองค์ประกอบสำคัญในการตัดสินคุณภาพสุดท้ายของสูตรแอลฟาในงานวิจัยนี้

6.5 ทดสอบความแตกต่างของค่าเฉลี่ยของตัววัดผล

ภายหลังจากการนำสูตรแอลฟาแต่ละสูตรไปทดสอบบนแพลตฟอร์ม WorldQuant BRAIN แล้ว งานวิจัยนี้ดำเนินการ เปรียบเทียบผลการดำเนินงานของสูตรแอลฟาสองกลุ่มในระดับเชิงสถิติ โดยมีวัตถุประสงค์เพื่อประเมินว่า กลุ่ม สูตรที่มาจากกรณีศึกษาภาพกล่องมีระดับของช่วงตัวต่างกัน และกลุ่มสูตรที่มาจากกรณีที่แผนภาพกล่องมีระดับ ของช่วงตัวไม่ต่างกัน มีผลการดำเนินงานเฉลี่ยแตกต่างกันหรือไม่ในมิติของตัวชี้วัดสำคัญของพอร์ตลงทุน

ในการศึกษานี้ ตัวชี้วัดที่ใช้เปรียบเทียบประกอบด้วย *Sharpe ratio*, *Turnover*, *Fitness*, *Returns*, *Drawdown* และ *Margin* ซึ่งเป็นตัวชี้วัดหลักที่สะท้อนคุณภาพของสูตรแอลฟาในหลายมิติ กล่าวคือ *Sharpe ratio* ใช้พิจารณาผลตอบแทนต่อหน่วยความเสี่ยง *Returns* ใช้วัดระดับผลตอบแทนเฉลี่ย *Drawdown* ใช้สะท้อนความ เสี่ยงจากการขาดทุนสะสมสูงสุด *Turnover* ใช้วัดความถี่ในการปรับเปลี่ยนพอร์ต *Margin* ใช้วัดประสิทธิภาพของ ผลตอบแทนต่อการซื้อขาย และ *Fitness* ใช้เป็นตัววัดเชิงรวมที่สะท้อนคุณภาพของสูตรภายใต้กรอบการประเมิน ของแพลตฟอร์ม

ในระดับการออกแบบการทดสอบ ให้กำหนดให้กลุ่มสูตรจากตารางที่ 1 เป็นกลุ่มที่หนึ่ง และกลุ่มสูตรจาก ตารางที่ 2 เป็นกลุ่มที่สอง โดยแต่ละกลุ่มมีจำนวนสูตรเท่ากันจำนวน 17 สูตร สำหรับตัวชี้วัดแต่ละตัว m งานวิจัย กำหนดสมมติฐานทางสถิติในรูปแบบสองด้าน (*two-sided test*) ดังนี้

$$H_{0,m} : \mu_{1,m} = \mu_{2,m}$$

$$H_{1,m} : \mu_{1,m} \neq \mu_{2,m}$$

โดยที่ $\mu_{1,m}$ และ $\mu_{2,m}$ แทนค่าเฉลี่ยของตัวชี้วัด m สำหรับกลุ่มสูตรที่หนึ่งและกลุ่มสูตรที่สองตามลำดับ สมมติฐานว่างจึงหมายถึงกรณีที่สองกลุ่มไม่มีความแตกต่างกันในเชิงค่าเฉลี่ยของตัวชี้วัดนั้น ขณะที่สมมติฐานทางเลือกหมายถึงถึงกรณีที่ทั้งสองกลุ่มมีผลการดำเนินงานเฉลี่ยแตกต่างกัน

เพื่อทดสอบสมมติฐานดังกล่าว งานวิจัยเลือกใช้ *Welch's two-sample t-test* สำหรับแต่ละตัวชี้วัด เนื่องจากการทดสอบรูปแบบนี้ไม่จำเป็นต้องสมมติว่าความแปรปรวนของทั้งสองกลุ่มเท่ากัน และมีความเหมาะสมกับข้อมูลเชิงประจักษ์ที่ได้จากสูตรแอลฟาคนละชุดซึ่งอาจมีการกระจายตัวแตกต่างกัน การเลือกใช้ *Welch's t-test* จึงช่วยลดข้อจำกัดของการใช้ *t-test* แบบดั้งเดิมที่กำหนดสมมติฐานเรื่องความแปรปรวนเท่ากันไว้ล่วงหน้า และทำให้การเปรียบเทียบมีความยืดหยุ่นมากขึ้นในเชิงวิธีวิทยา

ในเชิงการคำนวณ การทดสอบจะเปรียบเทียบค่าเฉลี่ยของตัวชี้วัดแต่ละตัวระหว่างสองกลุ่ม พร้อมรายงานค่าสถิติ t , จำนวนองศาเสรี (*degrees of freedom*), และค่า p -value สำหรับใช้ประกอบการตัดสินใจทางสถิติ โดย

งานวิจัยกำหนดระดับนัยสำคัญไว้ที่ 0.10, 0.05 และ 0.01 เพื่อใช้พิจารณาความเข้มแข็งของหลักฐานในการปฏิเสธสมมติฐานว่าง หากค่า p -value ต่ำกว่าระดับนัยสำคัญที่กำหนด จะถือว่าหลักฐานเชิงสถิติเพียงพอที่จะสรุปได้ว่าค่าเฉลี่ยของตัวชี้วัดนั้นแตกต่างกันระหว่างสองกลุ่มสูตร

นอกจากการพิจารณาในเชิงสถิติแล้ว งานวิจัยยังตีความผลในเชิงการเงินควบคู่กันไปด้วย กล่าวคือ สำหรับ *Sharpe ratio*, *Fitness*, *Returns* และ *Margin* ค่าที่สูงกว่าจะสะท้อนประสิทธิภาพของสูตรที่ดีกว่า ขณะที่สำหรับ *Turnover* และ *Drawdown* ค่าที่ต่ำกว่ามักถูกตีความว่าเป็นคุณลักษณะที่พึงประสงค์มากกว่า เนื่องจากสัมพันธ์กับต้นทุนการซื้อขายที่ต่ำกว่าและความเสี่ยงด้านการขาดทุนสะสมที่ต่ำกว่า ดังนั้น แม้การทดสอบทางสถิติจะมีหน้าที่ชี้ว่าความแตกต่างระหว่างสองกลุ่มมีนัยสำคัญหรือไม่ แต่ข้อสรุปเชิงงานวิจัยจะอาศัยทั้งทิศทางของส่วนต่างค่าเฉลี่ยและความหมายทางเศรษฐศาสตร์ของตัวชี้วัดนั้นร่วมกัน

ด้วยกรอบวิธีการดังกล่าว การทดสอบความแตกต่างของค่าเฉลี่ยของตัววัดผลจึงทำหน้าที่เป็นขั้นตอนสำคัญในการยกระดับการวิเคราะห์จากการเปรียบเทียบเชิงพรรณนาไปสู่การเปรียบเทียบเชิงอนุมาน กล่าวคือ ไม่เพียงตรวจสอบว่ากลุ่มสูตรหนึ่งดูเหมือนให้ผลดีกว่าอีกกลุ่มจากค่าเฉลี่ยที่สังเกตได้เท่านั้น แต่ยังประเมินด้วยว่าความแตกต่างดังกล่าวมีน้ำหนักเชิงสถิติมากพอที่จะสนับสนุนข้อสรุปทางการวิจัยหรือไม่

7 ผลการดำเนินงาน

7.1 สูตรแอลฟาในทดลอง

ให้ $O_{i,t}$, $H_{i,t}$, $L_{i,t}$, $C_{i,t}$ แทนราคาเปิด ราคาสูงสุด ราคาต่ำสุด และราคาปิดของหลักทรัพย์ i ณ เวลา t ตามลำดับ และให้

$$MA_n(C_{i,t}) = \frac{1}{n} \sum_{k=0}^{n-1} C_{i,t-k}, \quad STD_n(C_{i,t}) = \sqrt{\frac{1}{n-1} \sum_{k=0}^{n-1} (C_{i,t-k} - MA_n(C_{i,t}))^2}.$$

หากต้องการให้สูตรอยู่ในรูปแบบเดียวกับที่ใช้บน WorldQuant BRAIN สามารถครอบด้วย $\text{rank}(\cdot)$ ได้ในภายหลังตามข้อกำหนดของการจำลอง

7.1.1 สัญญาณกลับทิศของผลตอบแทนภายในวัน (Contrarian Intraday Return Signal)

สมมติฐาน (Hypotheses) เพื่ออธิบายกลไกของสัญญาณ สามารถตั้งสมมติฐานเชิงพฤติกรรมของตลาดได้ดังนี้

H1 (Intraday Overselling and Short-Term Rebound): เมื่อราคาปิดของหลักทรัพย์ต่ำกว่าราคาเปิดอย่างมีนัยสำคัญภายในวันเดียว การปรับตัวลงดังกล่าวอาจสะท้อนแรงขายที่มากเกินไปในระยะสั้น (*intraday overselling*) ซึ่งอาจทำให้ราคามีแนวโน้มฟื้นตัวหรือดีดกลับในช่วงเวลาถัดไป

$$\mathbb{E}[R_{i,t+1} \mid r_{i,t}^{\text{intra}} < 0] \propto -r_{i,t}^{\text{intra}}. \quad (7.1)$$

หรือเขียนในรูปที่สอดคล้องกับนิยามของอัลฟาโดยตรงได้เป็น

$$\mathbb{E}[R_{i,t+1} \mid \alpha_{i,t}] \propto \alpha_{i,t}, \quad \alpha_{i,t} = -r_{i,t}^{\text{intra}}. \quad (7.2)$$

การตีความสัญญาณ สัญญาณนี้อาศัยแนวคิดของ *ผลตอบแทนภายในวัน (intraday return)* ซึ่งคำนวณจากความแตกต่างระหว่างราคาปิดและราคาเปิดของหลักทรัพย์ในวันเดียวกัน กล่าวคือเปรียบเทียบการเปลี่ยนแปลงของราคาจากช่วงเปิดตลาดไปจนถึงช่วงปิดตลาด

หากราคาปิดต่ำกว่าราคาเปิดอย่างมาก แสดงว่าราคาปรับตัวลดลงตลอดวันและผลตอบแทนภายในวันจะเป็นลบ อย่างไรก็ตาม เมื่อนำผลตอบแทนภายในวันไปคูณด้วยเครื่องหมายลบ สัญญาณอัลฟาจะให้ค่าบวก ซึ่งสะท้อนแนวคิด *contrarian* ว่าการอ่อนตัวแรงภายในวันอาจเป็นการขายมากเกินไปในระยะสั้น และราคาอาจมีโอกาสดีกลับในวันถัดไป

ในทางกลับกัน หากราคาปิดสูงกว่าราคาเปิดอย่างมาก ผลตอบแทนภายในวันจะเป็นบวก และเมื่อคูณด้วยเครื่องหมายลบแล้ว สัญญาณจะให้ค่าลบ ซึ่งสะท้อนว่าการปรับตัวขึ้นแรงภายในวันอาจลดโอกาสของผลตอบแทนบวกในช่วงเวลาถัดไป

ดังนั้นสัญญาณนี้สะท้อนแนวคิดของ *mean-reversion* ระยะสั้น โดยการใช้การเปลี่ยนแปลงของราคาภายในวันเป็นตัววัดความสุดโต่งของการเคลื่อนไหวของราคา



รูปที่ 5: ตัวอย่างแท่งเทียนสำหรับอธิบายผลตอบแทนภายในวัน โดยกรณีที่ราคาปิดต่ำกว่าราคาเปิดจะทำให้ผลตอบแทนภายในวันเป็นลบ และสัญญาณอัลฟามีค่าเป็นบวก

รูปที่ 5 ใช้ประกอบการตีความสัญญาณผลตอบแทนภายในวัน โดยหัวใจสำคัญของสัญญาณนี้อยู่ที่ความสัมพันธ์ระหว่างราคาเปิด $O_{i,t}$ และราคาปิด $C_{i,t}$ หากในวันหนึ่งราคาปิดต่ำกว่าราคาเปิด จะได้ว่า

$$C_{i,t} < O_{i,t} \quad \Rightarrow \quad r_{i,t}^{\text{intra}} = \frac{C_{i,t} - O_{i,t}}{O_{i,t}} < 0,$$

ซึ่งหมายถึงหลักทรัพย์ให้ผลตอบแทนภายในวันเป็นลบ และเมื่อนำไปคูณด้วยเครื่องหมายลบ จะได้สัญญาณอัลฟาเป็นบวก กล่าวคือ

$$\alpha_{i,t} = -r_{i,t}^{\text{intra}} > 0.$$

ในเชิงเศรษฐศาสตร์การเงิน ภาพลักษณะดังกล่าวสะท้อนว่าหุ้นอาจถูกขายลงแรงเกินไปภายในวัน และมีความเป็นไปได้ที่จะเกิดการดีดกลับในวันถัดไป ดังนั้นค่าอัลฟาที่สูงจึงสอดคล้องกับมุมมองแบบ *contrarian* ที่มองว่าการอ่อนตัวแรงระหว่างวันอาจสร้างโอกาสของผลตอบแทนบวกในช่วงเวลาถัดไป

สูตรทางคณิตศาสตร์ ให้กำหนดผลตอบแทนภายในวันของหลักทรัพย์ i ณ เวลา t เป็น

$$r_{i,t}^{\text{intra}} = \frac{C_{i,t} - O_{i,t}}{O_{i,t}}, \quad (7.3)$$

โดยที่ $C_{i,t}$ คือราคาปิดของหลักทรัพย์ i ณ วัน t และ $O_{i,t}$ คือราคาเปิดของหลักทรัพย์ i ณ วันเดียวกัน

สูตรสัญญาณอัลฟา ดังนั้นสัญญาณอัลฟาที่ใช้ในกลยุทธ์สามารถเขียนได้เป็น

$$\alpha_{i,t} = - \left(\frac{C_{i,t} - O_{i,t}}{O_{i,t}} \right) \quad (7.4)$$

ซึ่งสามารถเขียนในรูปแบบภาษาของแพลตฟอร์ม WorldQuant ได้เป็น

$$-((\text{close} - \text{open}) / \text{open}) \quad (7.5)$$



รูปที่ 6: ผลการทดสอบ Alpha ใน WorldQuantBrain จากสมการ 7.5

สรุปผลการทดสอบ (ช่วง IS) จากผลการทดสอบในช่วงข้อมูลตัวอย่าง (in-sample; IS) แพลตฟอร์มจัดระดับคุณภาพของสัญญาณเป็น *Needs Improvement* แม้ว่าสัญญาณจะให้ผลตอบแทนรวมเป็นบวก โดยมีค่า Sharpe เท่ากับ 1.05 ผลตอบแทนสะสม (Returns) เท่ากับ 17.71% และการลดลงสูงสุดของพอร์ต (Drawdown) เท่ากับ 17.95% สะท้อนว่าผลตอบแทนต่อความเสี่ยงยังอยู่ในระดับปานกลางและมีช่วงเวลาที่พอร์ตเผชิญการถดถอยของมูลค่าค่อนข้างสูง

ภาระการซื้อขาย อัลฟามีอัตราการหมุนเวียนพอร์ต (Turnover) สูงที่ 81.66% ซึ่งอาจเพิ่มต้นทุนธุรกรรมและลดทอนผลตอบแทนจริง ขณะที่ค่า Margin เท่ากับ 4.34‰ แสดงว่ากำไรต่อหน่วยการซื้อขายยังอยู่ในระดับจำกัด

ความสม่ำเสมอของผลลัพธ์ ผลลัพธ์รายปีมีความแปรผัน โดยปี 2020 ให้สมรรถนะดีที่สุด (Sharpe = 2.29, Returns = 32.16%) ขณะที่ปี 2021 มี Drawdown สูงสุดถึง 17.95% แสดงว่าสัญญาณมีความไวต่อการ

เปลี่ยนแปลงของสภาวะตลาด แม้จำนวนหุ้นฝั่ง Long และ Short ในแต่ละปีจะอยู่ในช่วงประมาณ 277-307 หลัก
ทรัพย์สินต่อฝั่ง

7.1.2 สัญญาณความยาวไส้เทียนล่างแบบปรับขนาดด้วยราคาปิด (Normalized Lower Shadow Signal)

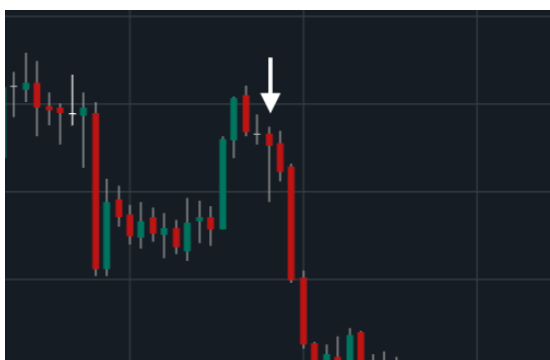
สมมติฐาน (Hypotheses) H1 (Intraday Rebound as Temporary Recovery): เมื่อราคาต่ำสุดของวันต่ำกว่าราคาเปิดหรือราคาปิดอย่างมาก แสดงว่าระหว่างวันเกิดแรงขายค่อนข้างรุนแรง แม้ว่าราคาจะสามารถฟื้นตัวกลับขึ้นมาได้ก่อนปิดตลาด แต่การฟื้นตัวดังกล่าวอาจเป็นเพียงการตีกลับชั่วคราวในระยะสั้น และยังสะท้อนความเปราะบางของราคาอยู่ ดังนั้นหลักทรัพย์ที่มีไส้เทียนล่างยาวมากอาจมีแนวโน้มให้ผลตอบแทนต่ำลงในเวลาถัดไป

$$\mathbb{E}[R_{i,t+1} \mid (\min(O_{i,t}, C_{i,t}) - L_{i,t}) \text{ สูง}] < 0 \quad (7.6)$$

การตีความสัญญาณ สัญญาณนี้วัด ความยาวของไส้เทียนด้านล่าง ของแท่งราคาภายในวัน โดยพิจารณาความแตกต่างระหว่างราคาต่ำสุดของวัน ($L_{i,t}$) กับค่าที่ต่ำที่สุดระหว่างราคาเปิดและราคาปิด ($\min(O_{i,t}, C_{i,t})$) ซึ่งเป็นตำแหน่งของตัวแท่งเทียนจริง (real body) ส่วนล่าง

หากค่าของ $\min(O_{i,t}, C_{i,t}) - L_{i,t}$ มีขนาดใหญ่ แสดงว่าราคาเคยลดลงอย่างมากในระหว่างวัน แต่สามารถฟื้นตัวกลับขึ้นมาได้ก่อนปิดตลาด สถานการณ์นี้สะท้อนว่ามีแรงซื้อกลับในช่วงราคาต่ำ อย่างไรก็ตาม ในงานวิจัยนี้ สัญญาณถูกนิยามให้มีเครื่องหมายลบ เพื่อสะท้อนมุมมองแบบ *contrarian* ว่าการฟื้นตัวดังกล่าวอาจไม่ยั่งยืน และอาจตามมาด้วยการอ่อนตัวของราคาในวันถัดไป

เพื่อให้สามารถเปรียบเทียบสัญญาณระหว่างหลักทรัพย์ที่มีระดับราคาต่างกันได้ จึงนำค่าดังกล่าวไปหารด้วยราคาปิด $C_{i,t}$ ทำให้สัญญาณอยู่ในรูปของสัดส่วน (normalized measure)



รูปที่ 7: ตัวอย่างแท่งเทียนที่ใช้ประกอบการอธิบายสัญญาณไส้เทียนล่างแบบปรับขนาดด้วยราคาปิด โดยลูกศรชี้แท่งเทียนที่พิจารณาเป็นตัวอย่าง

รูปที่ 7 แสดงตัวอย่างเชิงภาพของแท่งเทียนที่ใช้ในการตีความสัญญาณนี้ โดยพิจารณาแท่งเทียนที่ลูกศรชี้เป็นหลัก ในการอ่านสัญญาณ ไส้เทียนล่างจะวัดจากจุดต่ำสุดของวัน $L_{i,t}$ ขึ้นมาจนถึงขอบล่างของตัวแท่งจริง ซึ่งมีค่าเท่ากับ $\min(O_{i,t}, C_{i,t})$ หากระยะดังกล่าวมีขนาดมาก แสดงว่าราคาปรับตัวลงลึกในระหว่างวันก่อนจะฟื้นขึ้นมาใกล้ระดับเปิดหรือปิด ภาพลักษณะเช่นนี้สะท้อนว่าแม้จะมีแรงซื้อกลับเข้ามาในช่วงราคาต่ำ แต่ตลาดยังคงมีความ

ผันผวนและแรงขายกดดันอยู่พอสมควร ดังนั้นในรอบของงานวิจัยนี้จึงตีความไส้เทียนล่างที่ยาวเป็นสัญญาณเชิงกลับทิศ กล่าวคือ การฟื้นตัวภายในวันอาจเป็นเพียงการฟื้นตัวชั่วคราว และราคามีโอกาสอ่อนตัวลงต่อในวันถัดไป

สูตรทางคณิตศาสตร์ ให้กำหนด $O_{i,t}$ เป็นราคาเปิด $C_{i,t}$ เป็นราคาปิด และ $L_{i,t}$ เป็นราคาต่ำสุดของหลักทรัพย์ i ณ วัน t โดยนิยามความยาวไส้เทียนล่างที่ถูกปรับขนาดด้วยราคาปิดเป็น

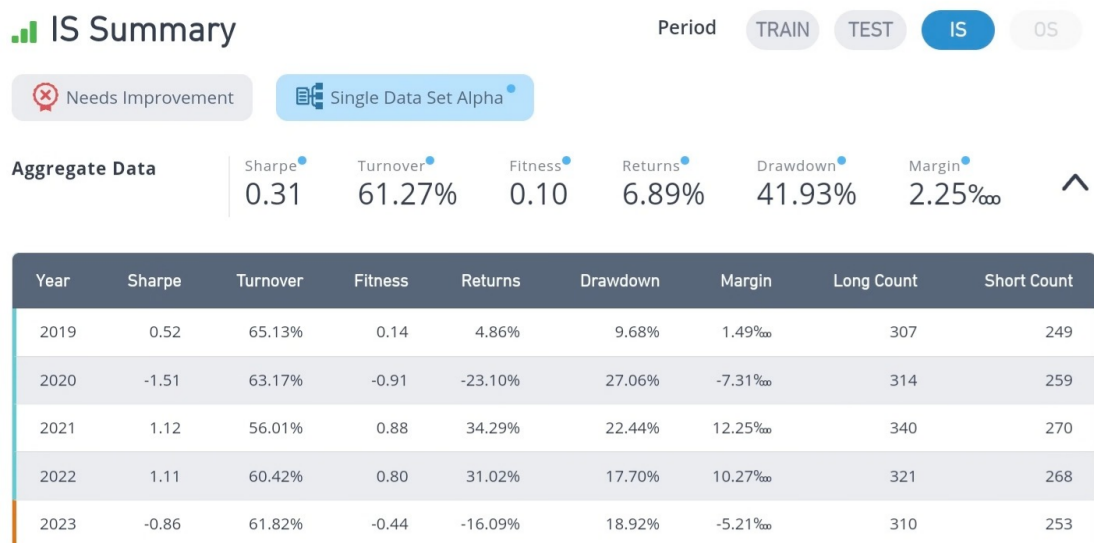
$$S_{i,t} = \frac{\min(O_{i,t}, C_{i,t}) - L_{i,t}}{C_{i,t}} \quad (7.7)$$

สูตรสัญญาณอัลฟา ดังนั้นสัญญาณอัลฟาสามารถเขียนได้เป็น

$$\alpha_{i,t} = -\frac{\min(O_{i,t}, C_{i,t}) - L_{i,t}}{C_{i,t}} \quad (7.8)$$

5 และในรูปแบบภาษาของแพลตฟอร์ม WorldQuant เขียนได้เป็น

$$-(\min(\text{open}, \text{close}) - \text{low}) / \text{close} \quad (7.9)$$



รูปที่ 8: ผลการทดสอบ Alpha ใน WorldQuantBrain จากสมการ 7.9

สรุปผลการทดสอบ (ช่วง IS) จากผลการทดสอบในช่วงข้อมูลตัวอย่าง (in-sample; IS) แพลตฟอร์มจัดระดับคุณภาพของสัญญาณเป็น *Needs Improvement* โดยมีค่า Sharpe เท่ากับ 0.31 ผลตอบแทนสะสม (Returns) เท่ากับ 6.89% และการลดลงสูงสุดของพอร์ต (Drawdown) สูงถึง 41.93% สะท้อนว่าผลตอบแทนต่อความเสี่ยงอยู่ในระดับต่ำและพอร์ตมีความผันผวนค่อนข้างสูง

ภาระการซื้อขาย อัลฟามีอัตราการหมุนเวียนพอร์ต (Turnover) 61.27% ซึ่งถือว่าค่อนข้างสูงและอาจเพิ่มต้นทุนธุรกรรม ขณะที่ค่า Margin เท่ากับ 2.25‰ แสดงว่ากำไรต่อหน่วยการซื้อขายยังอยู่ในระดับจำกัด

ความสม่ำเสมอของผลลัพธ์ ผลลัพธ์รายปีมีความผันผวนสูง โดยปี 2021 และ 2022 ให้ผลตอบแทนค่อนข้างดี (Returns 34.29% และ 31.02%) ขณะที่ปี 2020 และ 2023 ให้ผลตอบแทนติดลบ ส่งผลให้ประสิทธิภาพของสัญญาณไม่สม่ำเสมอในแต่ละช่วงเวลา

7.2 การวิเคราะห์แผนภาพกล่องของค่าสัญญาณแอลฟา

เพื่อประเมินความสามารถของสูตรแอลฟาในการแยกหุ้นผลตอบแทนสูงและหุ้นผลตอบแทนต่ำ งานวิจัยนี้ใช้แผนภาพกล่อง (boxplot) ของค่าเฉลี่ยแอลฟารายหลักทรัพย์ (*per-ticker mean(alpha)*) ตลอดช่วงเวลาตั้งแต่ปี 2019 เป็นต้นมา โดยเปรียบเทียบระหว่างกลุ่มหุ้นผลตอบแทนต่ำ 250 ตัวล่างสุด (*bad*) และกลุ่มหุ้นผลตอบแทนสูง 250 ตัวบนสุด (*good*) หลักการตีความสำคัญคือ หากสูตรแอลฟาสามารถจำแนกหุ้นได้ดี การกระจายของค่าระหว่างสองกลุ่มควรแยกออกจากกันอย่างชัดเจน ทั้งในแง่ของค่ามัธยฐาน ช่วงควอไทล์ และระดับโดยรวมของการกระจาย

จากรูปที่ 9 ซึ่งแสดงสูตรในกลุ่มผลตอบแทนย้อนหลังและโมเมนตัม พบว่าเกือบทุกสูตรในกลุ่มนี้มีการแยกตัวระหว่างกลุ่ม *bad* และ *good* อย่างชัดเจน โดยกลุ่มหุ้นผลตอบแทนสูงมีค่ามัธยฐานของแอลฟาสูงกว่ากลุ่มหุ้นผลตอบแทนต่ำอย่างสม่ำเสมอ ทั้งในสูตรผลตอบแทนย้อนหลังระยะ 1, 2, 5, 10 และ 20 วัน รวมถึงสูตรที่ใช้ผลรวมสะสมของผลตอบแทน (*mom_sumret_5*, *mom_sumret_10*, *mom_sumret_20*, *mom_sumret_60*) ลักษณะดังกล่าวสะท้อนว่าพีเจอรในกลุ่มนี้มีอำนาจจำแนกสูงในเชิงภาคตัดขวาง กล่าวคือ หุ้นที่ให้ผลตอบแทนดีในภาพรวมมักมีค่าของสัญญาณโมเมนตัมหรือผลตอบแทนสะสมอยู่ในระดับสูงกว่าหุ้นกลุ่มผลตอบแทนต่ำอย่างเป็นระบบ จึงเป็นหลักฐานสนับสนุนว่ากลุ่มสูตรประเภทนี้มีศักยภาพสูงในการถูกนำไปใช้สร้างแอลฟาเชิงปฏิบัติ

รูปที่ 10 แสดงผลของกลุ่มสูตรที่เกี่ยวข้องกับการกลับเข้าสู่ค่าเฉลี่ย ค่าเฉลี่ยเคลื่อนที่ ช่องว่างของราคา และผลตอบแทนภายในวัน พบว่าในสูตรประเภท *reversal* เช่น *rev_z_5*, *rev_z_10*, *rev_z_20* และ *rev_z_60* กลุ่มหุ้นผลตอบแทนสูงมีค่าของสัญญาณต่ำกว่ากลุ่มหุ้นผลตอบแทนต่ำอย่างชัดเจน ซึ่งตีความได้ว่าสูตรดังกล่าวสามารถตรวจจับลักษณะการเบี่ยงเบนจากค่าเฉลี่ยที่แตกต่างกันระหว่างสองกลุ่มหุ้นได้ นอกจากนี้ สูตรในกลุ่ม *rev_ma_ratio* ก็แสดงรูปแบบการแยกที่ชัดเจนเช่นกัน โดยกลุ่ม *good* มีค่าเฉลี่ยแอลฟาต่ำกว่ากลุ่ม *bad* อย่างต่อเนื่อง ในขณะที่สูตร *gap_1d* และ *intraday_ret* กลับให้ค่าของกลุ่ม *good* สูงกว่าอย่างเด่นชัด สะท้อนว่าตัวแปรที่เกี่ยวข้องกับช่องว่างของราคาเปิดและผลตอบแทนภายในวันอาจมีความสัมพันธ์เชิงบวกกับความสามารถในการแยกหุ้นที่มีผลตอบแทนดีในระยะยาว อย่างไรก็ตาม แม้กลุ่มนี้จะแสดงการแยกตัวที่ชัดในหลายสูตร แต่ทิศทางของการแยกไม่ได้เหมือนกันทุกสูตร จึงต้องตีความตามนิยามของสัญญาณแต่ละตัวโดยเฉพาะ

รูปที่ 11 แสดงกลุ่มสูตรที่เกี่ยวข้องกับความผันผวน ช่วงการเคลื่อนไหวของราคา ค่าช่วงราคาจริง และการจัดอันดับเชิงเวลา (*volatility / range / true range / ranks*) โดยผลลัพธ์ชี้ว่าหลายสูตรในกลุ่มนี้สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ค่อนข้างชัดเจน โดยเฉพาะ *vol_ret_5*, *vol_ret_10*, *vol_ret_20*, *vol_ret_60*, *true_range_pct*, *atr_14* และ *range_pct* ซึ่งกลุ่มหุ้นผลตอบแทนสูงมีค่ามัธยฐานของสัญญาณสูงกว่ากลุ่มหุ้นผลตอบแทนต่ำอย่างสม่ำเสมอ สะท้อนว่าหุ้นที่ให้ผลตอบแทนดีมักมีระดับความผันผวนหรือขนาดการแกว่งตัวของราคาที่เด่นกว่าหุ้นกลุ่มอ่อนแอในหลายมิติ อย่างไรก็ตาม สูตร *ts_rank_range_20* และ *vol_logvol_20* มีการทับซ้อนของการกระจายค่อนข้างมาก จึงอาจตีความได้ว่าตัวแปรในกลุ่มนี้ไม่ได้มีอำนาจจำแนกเท่ากันทุกสูตร แต่โดยภาพรวมยังสนับสนุนแนวคิดที่ว่าความผันผวนและช่วงการเคลื่อนไหวของราคาเป็นแหล่งข้อมูลสำคัญสำหรับการจำแนกหุ้นในเชิงภาคตัดขวาง

รูปที่ 12 แสดงกลุ่มสูตรที่เกี่ยวข้องกับการทะลุกรอบราคา ตัวชี้วัด *stochastic* และรูปแบบแท่งเทียนเชิงแพด

เทิร์น (*breakouts / stochastic / candlestick patterns*) โดยผลลัพธ์สะท้อนว่ากลุ่มสูตรประเภท *stochastic* เช่น *stoch_k_10* และ *stoch_k_20* มีการแยกตัวระหว่างกลุ่ม *good* และ *bad* ได้ค่อนข้างชัด กล่าวคือ กลุ่มหุ้นผลตอบแทนสูงมักมีค่ามัธยฐานสูงกว่า ซึ่งสอดคล้องกับแนวคิดที่ว่าหุ้นที่ปิดใกล้ด้านบนของกรอบราคาย้อนหลังมักมีลักษณะความแข็งแกร่งเชิงราคา ในส่วนของสูตร *breakout_high_10* และ *breakout_high_20* แม้ค่าของสัญญาณจะเป็นลบทั้งสองกลุ่ม แต่กลุ่ม *good* มีแนวโน้มอยู่ใกล้จุดสูงสุดย้อนหลังมากกว่า ขณะที่สูตรเชิงรูปแบบแท่งเทียน เช่น *bullish_engulfing*, *bearish_engulfing*, *hammer*, *shooting_star* และ *engulfing_score* แสดงการแยกตัวที่อ่อนกว่าหรือมีการทับซ้อนของควอไทล์ค่อนข้างมาก จึงอาจตีความได้ว่าตัวแปรที่วัดตำแหน่งหรือการทะลุกรอบของราคาให้ข้อมูลจำแนกได้ชัดกว่าตัวแปรที่อาศัยรูปแบบแท่งเทียนเชิงเหตุการณ์แบบไม่ต่อเนื่อง

รูปที่ 13 แสดงกลุ่มสูตรที่สะท้อนแนวโน้ม ความชันของราคา และตำแหน่งของราคาปัจจุบันเมื่อเทียบกับกรอบราคาในอดีต ซึ่งเป็นหนึ่งในกลุ่มที่แสดงการแยกตัวระหว่างสองกลุ่มหุ้นได้ชัดเจนที่สุด สูตรอย่าง *trend_ma_gap_5_20*, *trend_ma_gap_10_60*, *trend_slope_5* และ *trend_slope_20* แสดงให้เห็นว่ากลุ่มหุ้นผลตอบแทนสูงมีระดับของสัญญาณสูงกว่ากลุ่มหุ้นผลตอบแทนต่ำอย่างต่อเนื่อง สะท้อนว่าหุ้นที่ให้ผลตอบแทนดีมักอยู่ในสถานะที่ราคาเคลื่อนไหวตามแนวโน้มขาขึ้นชัดเจนกว่า นอกจากนี้ สูตรที่วัดตำแหน่งของราคาในกรอบสูงสุดและต่ำสุดย้อนหลัง เช่น *pos_in_range_20*, *pos_in_range_60*, *dist_from_low_20* และ *ts_rank_close_20* ก็แสดงรูปแบบสอดคล้องกัน โดยหุ้นกลุ่ม *good* มักมีราคาปัจจุบันอยู่ใกล้บริเวณด้านบนของกรอบราคาย้อนหลังมากกว่า ส่วน *dist_from_high_20* แม้จะมีค่าเป็นลบทั้งสองกลุ่ม แต่กลุ่ม *good* มีแนวโน้มอยู่ใกล้จุดสูงสุดย้อนหลังมากกว่าเช่นกัน ภาพรวมของกลุ่มนี้จึงสนับสนุนแนวคิดที่ว่า *information content* ของแนวโน้มและตำแหน่งราคามีความสัมพันธ์กับผลตอบแทนข้ามหลักทรัพย์อย่างมีนัยสำคัญ

รูปที่ 14 แสดงกลุ่มสูตรที่เกี่ยวข้องกับมูลค่าการซื้อขายและแรงกดดันของกระแสเงิน (*dollar volume / flow pressure*) ซึ่งเชื่อมโยงพฤติกรรมของราคากับปริมาณซื้อขายในเชิงมูลค่า ผลลัพธ์แสดงให้เห็นว่าสูตร *dollar_vol_ratio_20*, *dollar_vol_ratio_60*, *signed_dollar_flow_10*, *signed_dollar_flow_20* และ *absret_x_volratio20* สามารถแยกกลุ่มหุ้นผลตอบแทนสูงและต่ำได้ค่อนข้างชัด โดยกลุ่มหุ้นผลตอบแทนสูงมีค่ามัธยฐานสูงกว่าอย่างต่อเนื่อง สะท้อนว่าการเคลื่อนไหวของราคาที่เกิดขึ้นร่วมกับมูลค่าการซื้อขายสูงหรือแรงซื้อขายที่มีทิศทางชัดเจนอาจเป็นสัญญาณของหุ้นที่มีคุณภาพเชิงผลตอบแทนดีกว่า นอกจากนี้ สูตร *corr_close_volume_20* และ *corr_close_volume_60* ยังแสดงความแตกต่างเชิงระดับระหว่างสองกลุ่ม แม้จะมีการทับซ้อนอยู่บ้าง ขณะที่ *volume_shock_z_20* ให้ภาพการแยกที่อ่อนกว่าเมื่อเทียบกับสูตรอื่นในหมวดเดียวกัน ดังนั้นโดยภาพรวม กลุ่มตัวแปรที่ผสมข้อมูลราคากับปริมาณซื้อขายในเชิงมูลค่าสามารถให้ข้อมูลจำแนกที่มีนัยสำคัญมากกว่าการวัดความผิดปกติของปริมาณซื้อขายเพียงลำพัง

รูปที่ 15 แสดงกลุ่มสูตรที่เกี่ยวข้องกับพลวัตของช่องว่างระหว่างวัน (*overnight / gap dynamics*) โดยผลลัพธ์มีความน่าสนใจในเชิงโครงสร้าง กล่าวคือ สูตร *overnight_gap*, *open_to_prev_low*, *close_to_prev_low*, *mean_gap_5* และ *std_gap_20* ให้ค่าของกลุ่มหุ้นผลตอบแทนสูงสูงกว่ากลุ่มหุ้นผลตอบแทนต่ำค่อนข้างชัดเจน ซึ่งสะท้อนว่าพฤติกรรมของราคาเปิดและราคาปิดเมื่อเทียบกับจุดต่ำของวันก่อนหน้าอาจแฝงข้อมูลเกี่ยวกับคุณภาพของหุ้นในเชิงผลตอบแทนได้ในทางกลับกัน สูตร *open_to_prev_high* และ *close_to_prev_high* มีค่าติดลบมากกว่าในกลุ่ม *good* ซึ่งสอดคล้องกับนิยามของสัญญาณที่วัดระยะห่างจากจุดสูงสุดก่อนหน้า ส่วน *gap_rank_20* มีการทับซ้อนของการกระจายค่อนข้างมากระหว่างสองกลุ่ม และ *gap_reversal_score* แสดงการแยกที่อ่อนกว่าสูตรอื่นในหมวดเดียวกัน จึงอาจตีความได้ว่าสูตรในกลุ่ม *gap dynamics* ไม่ได้มีคุณภาพ

สม่ำเสมอทุกตัว แต่บางสูตรมีศักยภาพชัดเจนในการแยกหุ้นผลตอบแทนสูงและต่ำ

รูปที่ 16 แสดงกลุ่มสูตรที่เกี่ยวข้องกับระบอบความผันผวนและการบีบตัวของช่วงราคา (*volatility regime / compression*) ซึ่งมุ่งวัดทั้งความผันผวนของผลตอบแทน อัตราส่วนความผันผวนระยะสั้นต่อระยะยาว ความกว้างของช่วงราคาสูงสุด-ต่ำสุด และแรงกดดันจากการทะลุรอบราคา ผลลัพธ์ชี้ว่าบางสูตรในกลุ่มนี้ เช่น `ret_to_range_20` และ `range_breakout_pressure_20` สามารถแยกกลุ่ม *good* และ *bad* ได้ค่อนข้างชัด โดยกลุ่มหุ้นผลตอบแทนสูงมีค่ามัธยฐานสูงกว่าอย่างต่อเนื่อง ขณะที่สูตรอย่าง `vol_ratio_20_60` และ `vol_ratio_5_20` แม้ความแตกต่างของค่ากลางจะไม่กว้างมาก แต่ยังคงมีทิศทางของการแยกที่สม่ำเสมอในระดับหนึ่ง ในทางกลับกัน สูตรบางตัว เช่น `range_ratio_5_20` และ `atr_ratio_14_60` มีช่วงควอไทล์ที่ทับซ้อนกันค่อนข้างมาก จึงสะท้อนว่าอำนาจจำแนกของกลุ่มความผันผวนไม่ได้สม่ำเสมอทุกสูตร โดยภาพรวม กลุ่มนี้แสดงให้เห็นว่าความผันผวนและการบีบตัวของช่วงราคาอาจมีข้อมูลเชิงพยากรณ์ แต่คุณภาพของสัญญาณขึ้นอยู่กับรูปแบบการนิยามตัวแปรเป็นสำคัญ

รูปที่ 17 แสดงกลุ่มสูตรที่อาศัยโครงสร้างของแท่งเทียน (*candlestick body / wick structure*) เช่น ขนาดของตัวแท่ง สัดส่วนของไส้เทียนบนและล่าง ตำแหน่งของราคาเปิดและราคาปิดภายในกรอบสูงสุด-ต่ำสุดของวัน และรูปแบบแท่งเทียนเฉพาะ เช่น *marubozu* และ *doji* ผลลัพธ์แสดงว่าสูตรในกลุ่มนี้สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ในระดับปานกลางถึงค่อนข้างดี โดยเฉพาะ `body_pct`, `upper_wick_pct`, `lower_wick_pct`, `body_to_range` และ `close_location` ซึ่งกลุ่ม *good* มีค่ามัธยฐานสูงกว่ากลุ่ม *bad* อย่างสม่ำเสมอ สะท้อนว่าหุ้นที่ให้ผลตอบแทนดีมักมีลักษณะการปิดใกล้จุดสูงของวัน มีตัวแท่งที่เด่นกว่า และมีโครงสร้างแท่งที่สะท้อนแรงซื้อภายในวันมากกว่า อย่างไรก็ตาม สูตรบางตัว เช่น `wick_imbalance`, `marubozu` และ `doji` แสดงการแยกที่อ่อนกว่าเมื่อเทียบกับสูตรเชิงต่อเนื่อง จึงอาจตีความได้ว่าตัวแปรเชิงรูปแบบแท่งเทียนแบบไบนารีมีพลังในการจำแนกน้อยกว่าตัวแปรที่วัดโครงสร้างของแท่งเทียนในเชิงปริมาณโดยตรง

เมื่อพิจารณาร่วมกันทั้งหกรูป พบว่ากลุ่มสูตรที่มีการแยกตัวของแผนภาพกล่องชัดเจนและสม่ำเสมอมากที่สุด ได้แก่ กลุ่มผลตอบแทนย้อนหลัง/โมเมนตัม และกลุ่มแนวโน้ม/ตำแหน่งราคา ขณะที่กลุ่มการกลับเข้าสู่ค่าเฉลี่ยและกลุ่มช่องว่างระหว่างวันแม้จะแสดงความสามารถในการแยกเช่นกัน แต่ความชัดเจนของผลขึ้นอยู่กับนิยามของสูตรแต่ละตัวมากกว่า ส่วนกลุ่มความผันผวนและการบีบตัวของช่วงราคาให้ผลแบบผสม กล่าวคือบางสูตรมีศักยภาพชัดเจน แต่บางสูตรยังมีการทับซ้อนของการกระจายสูง ขณะที่กลุ่มโครงสร้างแท่งเทียนให้ผลในระดับปานกลางถึงค่อนข้างดี โดยเฉพาะตัวแปรเชิงต่อเนื่องที่วัดโครงสร้างของแท่งโดยตรง ข้อสังเกตเหล่านี้สอดคล้องกับผลการทดสอบบนแพลตฟอร์ม WorldQuant BRAIN และผลการเปรียบเทียบค่าเฉลี่ยของตัวชี้วัด ซึ่งชี้ว่ากลุ่มสูตรที่มีการแยกตัวชัดเจนในระดับข้อมูลมักให้ผลการดำเนินงานที่ดีกว่าในระดับพอร์ตลงทุนจริง

กล่าวโดยสรุป การใช้แผนภาพกล่องของค่าเฉลี่ยแอลฟารายหลักทรัพย์สินเป็นเครื่องมือวิเคราะห์เบื้องต้นช่วยให้สามารถมองเห็นความแตกต่างของโครงสร้างสัญญาณระหว่างหุ้นผลตอบแทนสูงและต่ำได้อย่างเป็นรูปธรรม และช่วยทำหน้าที่เป็นกลไกคัดกรองสูตรแอลฟาก่อนนำไปทดสอบเชิงพอร์ตบนแพลตฟอร์ม WorldQuant BRAIN ได้อย่างมีประสิทธิภาพ

7.3 ผลการทดสอบทางสถิติ

ตาราง 1 แสดงผลการทดสอบ Welch's two-sample t-test เพื่อเปรียบเทียบค่าเฉลี่ยของตัวชี้วัดสำคัญระหว่างกลุ่มสูตรจากตารางที่ 1 และตารางที่ 2 โดยใช้ Welch's t-test เนื่องจากวิธีดังกล่าวเหมาะสมกับกรณีที่มีความแปรปรวนของสองกลุ่มอาจไม่เท่ากัน และมีความยืดหยุ่นมากกว่าการทดสอบ t-test แบบสมมติให้ความแปรปรวน

ตารางที่ 1: ผลการทดสอบ Welch's two-sample t-test เพื่อเปรียบเทียบค่าเฉลี่ยของแต่ละตัวชี้วัดระหว่างตารางที่ 1 และตารางที่ 2

ตัวชี้วัด	ค่าเฉลี่ย T1	ค่าเฉลี่ย T2	ส่วนต่าง (T1-T2)	t-stat	df	p-value	นัยสำคัญ
Sharpe	0.705	0.439	0.266	2.048	25.299	0.0510	*
Turnover (%)	30.622	56.502	-25.880	-3.321	30.580	0.0023	***
Fitness	0.502	0.141	0.361	4.800	29.821	0.0000	***
Returns (%)	13.345	3.391	9.954	6.505	27.822	0.0000	***
Drawdown (%)	24.212	25.439	-1.226	-0.208	17.074	0.8379	n.s.
Margin (%)	11.252	1.292	9.960	5.846	17.387	0.0000	***

หมายเหตุ: ใช้ Welch's two-sample t-test เพื่อเปรียบเทียบค่าเฉลี่ยของแต่ละตัวชี้วัดระหว่างกลุ่มสูตรจากตารางที่ 1 และตารางที่ 2 โดยแต่ละตารางมีจำนวนตัวอย่าง 17 สูตร ระดับนัยสำคัญกำหนดเป็น $*p < 0.10$, $**p < 0.05$, และ $***p < 0.01$; ส่วน n.s. หมายถึงไม่มีนัยสำคัญทางสถิติ

เท่ากัน ผลที่ได้จึงสะท้อนการเปรียบเทียบค่าเฉลี่ยของแต่ละตัวชี้วัดภายใต้เงื่อนไขที่ค่อนข้างสมจริงสำหรับข้อมูลเชิงประจักษ์ลักษณะนี้

ในเชิงสถิติ ค่า p -value ทำหน้าที่เป็นเกณฑ์หลักในการประเมินว่าความแตกต่างของค่าเฉลี่ยระหว่างสองกลุ่มมีนัยสำคัญหรือไม่ หากค่า p -value ต่ำกว่าระดับนัยสำคัญที่กำหนด จะสามารถปฏิเสธสมมติฐานว่างที่ว่าค่าเฉลี่ยของสองกลุ่มไม่แตกต่างกันได้ สำหรับผลในตารางนี้พบว่า *Turnover*, *Fitness*, *Returns* และ *Margin* มีค่า p -value ต่ำกว่า 0.01 ทั้งหมด จึงสรุปได้ว่าตัวชี้วัดทั้งสี่มีความแตกต่างระหว่างตารางที่ 1 และตารางที่ 2 อย่างมีนัยสำคัญทางสถิติในระดับสูง กล่าวคือ ความแตกต่างที่สังเกตได้ไม่น่าจะเกิดจากความผันผวนแบบสุ่มของตัวอย่างเพียงอย่างเดียว

สำหรับ *Sharpe ratio* พบว่า ค่าเฉลี่ยของตารางที่ 1 สูงกว่าตารางที่ 2 อยู่ที่ 0.266 และมีค่า $p = 0.0510$ ซึ่งอยู่ใกล้เคียงกับเกณฑ์ 0.05 มาก ผลลัพธ์นี้จึงสามารถตีความได้ว่า มีหลักฐานเชิงสถิติในระดับอ่อนว่าตารางที่ 1 มีประสิทธิภาพด้านผลตอบแทนต่อความเสี่ยงดีกว่าตารางที่ 2 โดยหากใช้เกณฑ์นัยสำคัญที่ระดับ 10% จะถือว่าเป็นนัยสำคัญ แต่หากใช้เกณฑ์ที่เข้มงวดกว่าคือ 5% จะยังไม่เพียงพอที่จะสรุปอย่างเด็ดขาด ดังนั้นผลในส่วนของ *Sharpe* จึงควรถูกมองว่าเป็น “แนวโน้มเชิงบวก” มากกว่าข้อสรุปที่หนักแน่นเท่ากับตัวชี้วัดอื่น

ในทางตรงกันข้าม *Drawdown* มีค่า $p = 0.8379$ ซึ่งสูงมากเมื่อเทียบกับเกณฑ์นัยสำคัญทั่วไป แสดงว่าไม่พบหลักฐานเชิงสถิติที่ว่าค่าเฉลี่ยของ *drawdown* ระหว่างสองกลุ่มแตกต่างกัน กล่าวอีกนัยหนึ่ง แม้ว่ากลุ่มสูตรจากตารางที่ 1 จะให้ผลตอบแทนและตัวชี้วัดด้านประสิทธิภาพสูงกว่าอย่างชัดเจน แต่ในมิติของการขาดทุนสะสมสูงสุดนั้น ทั้งสองกลุ่มยังไม่แตกต่างกันอย่างมีนัยสำคัญในเชิงสถิติ

เมื่อพิจารณาในเชิงการเงิน ผลลัพธ์ดังกล่าวมีความหมายค่อนข้างชัดเจน กลุ่มสูตรในตารางที่ 1 ให้ *Returns* เฉลี่ยสูงกว่าตารางที่ 2 ถึง 9.954 จุดเปอร์เซ็นต์ และมีความแตกต่างอย่างมีนัยสำคัญสูงมาก นั้นหมายความว่า หากมองในมิติของความสามารถในการสร้างผลตอบแทน กลุ่มสูตรจากตารางที่ 1 แสดงศักยภาพเหนือกว่าอย่างเด่นชัด นอกจากนี้ *Fitness* ที่สูงกว่าของตารางที่ 1 ยังชี้ว่าผลตอบแทนดังกล่าวไม่ได้เกิดขึ้นเพียงจากการรับความเสี่ยงมากขึ้นหรือการซื้อขายถี่ขึ้นเท่านั้น แต่เป็นผลตอบแทนที่มีคุณภาพเชิงโครงสร้างดีกว่าเมื่อประเมินภายใต้กรอบของแพลตฟอร์ม

ผลของ *Turnover* มีนัยสำคัญมากในเชิงปฏิบัติ เพราะตารางที่ 1 มีค่าเฉลี่ยต่ำกว่าตารางที่ 2 อย่างมีนัยสำคัญทางสถิติ โดยต่ำกว่าประมาณ 25.88 จุดเปอร์เซ็นต์ การที่กลยุทธ์มี *turnover* ต่ำกว่าหมายถึงการปรับเปลี่ยนพอร์ตน้อยกว่า ซึ่งโดยทั่วไปสัมพันธ์กับต้นทุนธุรกรรมที่ต่ำกว่า ความเสี่ยงจาก *slippage* ที่ต่ำกว่า และความสามารถในการนำกลยุทธ์ไปใช้จริงได้มากกว่า ดังนั้นผลข้อนี้จึงถือเป็นข้อได้เปรียบสำคัญของกลุ่มสูตรจากตารางที่ 1 ไม่ใช่เพียงในเชิงสถิติ แต่รวมถึงในมิติของการใช้งานจริงในพอร์ตลงทุน

ในทำนองเดียวกัน *Margin* ที่สูงกว่าของตารางที่ 1 อย่างมีนัยสำคัญสะท้อนว่ากลุ่มสูตรดังกล่าวสามารถสร้างผลตอบแทนต่อหน่วยของการซื้อขายได้ดีกว่า ซึ่งเป็นสัญญาณที่ดีในเชิงเศรษฐศาสตร์การลงทุน เพราะหมายความว่ากลยุทธ์สามารถเปลี่ยนการเคลื่อนไหวของราคาหรือสัญญาณให้กลายเป็นกำไรได้อย่างมีประสิทธิภาพมากกว่า เมื่อพิจารณาพร้อมกับ *turnover* ที่ต่ำกว่า ยิ่งตอกย้ำว่ากลุ่มสูตรในตารางที่ 1 มีความคุ้มค่าในการเทรดสูงกว่าตารางที่ 2

แม้ว่า *Drawdown* จะไม่แตกต่างกันอย่างมีนัยสำคัญ แต่ข้อสรุปนี้ก็มีความสำคัญเช่นกัน เพราะมันบ่งชี้ว่าความเหนือกว่าของตารางที่ 1 ในด้านผลตอบแทนไม่ได้มาพร้อมกับหลักฐานว่ามีความเสี่ยงด้านการขาดทุนสูงสุดแตกต่างจากตารางที่ 2 อย่างชัดเจน ในเชิงการเงิน นี่อาจตีความได้ว่ากลุ่มสูตรจากตารางที่ 1 สามารถให้ผลตอบแทนที่สูงกว่าโดยไม่ต้องแลกกับ *drawdown* ที่สูงขึ้นอย่างมีนัยสำคัญ ซึ่งถือเป็นคุณลักษณะที่น่าสนใจสำหรับการคัดเลือกกลยุทธ์

โดยสรุป ทั้งในเชิงสถิติและเชิงการเงิน ผลการทดสอบสนับสนุนข้อสรุปว่ากลุ่มสูตรจากตารางที่ 1 มีคุณภาพโดยรวมเหนือกว่ากลุ่มสูตรจากตารางที่ 2 กล่าวคือ กลุ่มสูตรดังกล่าวให้ผลตอบแทนเฉลี่ยสูงกว่า มี *fitness* สูงกว่า มี *margin* สูงกว่า และมี *turnover* ต่ำกว่าอย่างมีนัยสำคัญทางสถิติ ขณะที่ความแตกต่างด้าน *drawdown* ไม่ชัดเจน และ *Sharpe ratio* อยู่ในระดับเกือบมีนัยสำคัญ ผลลัพธ์นี้จึงชี้ว่า หากเป้าหมายคือการเลือกชุดสูตรที่มีศักยภาพในการนำไปใช้จริง กลุ่มสูตรจากตารางที่ 1 ย่อมมีความได้เปรียบมากกว่าในภาพรวม

8 สรุปและอภิปราย

การศึกษาครั้งนี้มีจุดมุ่งหมายเพื่อประเมินว่าสูตรแอลฟาที่สร้างจากข้อมูล OHLCV สามารถใช้เป็นสัญญาณเชิงปริมาณสำหรับการคัดเลือกหลักทรัพย์และสร้างพอร์ตการลงทุนเชิงระบบได้มากน้อยเพียงใด โดยใช้ข้อมูลของหลักทรัพย์ในดัชนี S&P 500 เป็นตัวแทนโดยประมาณของจักรวาลสินทรัพย์ในแพลตฟอร์ม WorldQuant BRAIN และครอบคลุมช่วงเวลาระหว่างปี ค.ศ. 2019 ถึง 2024 ในการดำเนินงาน ผู้วิจัยได้พัฒนาสูตรแอลฟาหลายหมวดหมู่จากข้อมูลราคาและปริมาณซื้อขาย เช่น กลุ่มผลตอบแทนย้อนหลัง โมเมนตัม การกลับเข้าสู่ค่าเฉลี่ย ความผันผวนความสัมพันธ์ระหว่างราคาและปริมาณซื้อขาย และรูปแบบแท่งเทียน ก่อนนำไปตรวจสอบความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำ และทดสอบผลการดำเนินงานจริงบนแพลตฟอร์ม WorldQuant BRAIN

ผลการศึกษาชี้ให้เห็นว่า แนวคิดเรื่อง “ความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำ” มีความสำคัญในฐานะเกณฑ์คัดกรองเบื้องต้นของสูตรแอลฟา กล่าวคือ กลุ่มสูตรที่แสดงรูปแบบการกระจายของค่าสัญญาณแตกต่างกันอย่างชัดเจนระหว่างหุ้นผลตอบแทนสูงและหุ้นผลตอบแทนต่ำ มีแนวโน้มให้ผลการทดสอบบนแพลตฟอร์มดีกว่ากลุ่มสูตรที่ไม่สามารถแยกความแตกต่างดังกล่าวได้อย่างเด่นชัด เมื่อพิจารณาจากผลการทดสอบทางสถิติ พบว่ากลุ่มสูตรจากตารางที่ 1 ซึ่งเป็นกลุ่มสูตรที่มีลักษณะการแยกกลุ่มชัดกว่า มีค่าเฉลี่ยของ *Fitness*, *Returns* และ *Margin* สูงกว่ากลุ่มสูตรจากตารางที่ 2 อย่างมีนัยสำคัญทางสถิติ ขณะเดียวกันยังมีค่าเฉลี่ยของ *Turnover* ต่ำกว่าอย่างมีนัยสำคัญด้วย ซึ่งสะท้อนถึงความสามารถที่เปรียบทั้งในแง่ผลตอบแทนและความคุ้มค่าของการซื้อขาย

ในมุมมองของการตีความเชิงการเงิน ผลลัพธ์นี้มีนัยสำคัญอย่างมาก เพราะแสดงให้เห็นว่าการมีพลังในการจำแนก (*discriminative power*) ในระดับข้อมูลสามารถเชื่อมโยงไปสู่ผลการดำเนินงานในระดับพอร์ตได้จริง กล่าวคือ สูตรแอลฟาที่สามารถให้ค่าสัญญาณแตกต่างกันอย่างมีระบบระหว่างหุ้นที่ให้ผลตอบแทนดีและหุ้นที่ให้ผลตอบแทนต่ำ ย่อมมีแนวโน้มสูงกว่าที่จะนำไปสู่กระบวนการจัดอันดับหลักทรัพย์ที่มีคุณภาพ และท้ายที่สุดสะท้อนออกมาเป็นผลตอบแทนเฉลี่ยที่สูงกว่าเมื่อถูกนำไปใช้ในกรอบการจำลองแบบ long-short ของ WorldQuant BRAIN ผลลัพธ์ของ *Fitness* และ *Margin* ที่สูงกว่าในกลุ่มสูตรดังกล่าวยิ่งช่วยสนับสนุนแนวคิดนี้ เพราะไม่ได้หมายถึงเพียงการให้ผลตอบแทนสูงกว่าเท่านั้น แต่ยังบ่งชี้ว่าผลตอบแทนดังกล่าวมีคุณภาพเชิงโครงสร้างและมีประสิทธิภาพต่อหน่วยของการซื้อขายมากกว่าอีกด้วย

ประเด็นที่น่าสนใจอีกประการหนึ่งคือ *Turnover* ของกลุ่มสูตรที่ 1 ต่ำกว่ากลุ่มสูตรที่ 2 อย่างมีนัยสำคัญทางสถิติ ซึ่งในเชิงปฏิบัติถือเป็นผลลัพธ์ที่สำคัญมาก เนื่องจากสูตรที่มี turnover ต่ำกว่ามักมีต้นทุนธุรกรรมต่ำกว่า มีความเสี่ยงจาก *slippage* ต่ำกว่า และมีโอกาสนำไปใช้จริงได้มากกว่า สูตรแอลฟาที่สามารถสร้างผลตอบแทนได้ดีโดยไม่ต้องหมุนพอร์ตมากเกินไปย่อมมีคุณค่าเชิงการลงทุนสูงกว่าสูตรที่ต้องอาศัยการซื้อขายถี่เพื่อสร้างผลตอบแทน ผลลัพธ์ส่วนนี้จึงชี้ว่า สูตรในกลุ่มที่สามารถแยกหุ้นผลตอบแทนสูงและต่ำได้ชัดเจน ไม่เพียงแต่ให้ผลตอบแทนดีกว่าเท่านั้น แต่ยังมีประสิทธิภาพในการใช้งานจริงมากกว่าด้วย

อย่างไรก็ตาม ผลลัพธ์ไม่ได้ชี้ว่าทุกมิติของประสิทธิภาพแตกต่างกันอย่างชัดเจนเสมอไป โดยเฉพาะ *Draw-down* ซึ่งไม่พบความแตกต่างอย่างมีนัยสำคัญทางสถิติระหว่างสองกลุ่ม และ *Sharpe ratio* ซึ่งแม้กลุ่มสูตรที่ 1 จะมีค่าเฉลี่ยสูงกว่า แต่ผลยังอยู่ในระดับเพียงเกือบมีนัยสำคัญเท่านั้น ข้อค้นพบนี้มีความสำคัญในเชิงอภิปราย เพราะสะท้อนว่า ความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำอาจสัมพันธ์กับผลตอบแทนและประสิทธิภาพของการซื้อขายอย่างชัดเจน แต่ยังไม่เพียงพอที่จะรับประกันว่าโครงสร้างความเสี่ยงของพอร์ตจะดีขึ้นในทุกมิติอย่างมีนัยสำคัญ กล่าวอีกนัยหนึ่ง สูตรแอลฟาที่ดีในแง่การจำแนกข้อมูลอาจยังต้องผ่านการออกแบบเพิ่มเติมในระดับการจัดการพอร์ตและการควบคุมความเสี่ยง เพื่อให้ผลลัพธ์เด่นชัดในมิติของ *drawdown* และผลตอบแทนต่อความเสี่ยงโดยรวม

หากพิจารณาในมุมมองระเบียบวิธีวิจัย งานศึกษานี้สนับสนุนแนวคิดว่ากระบวนการวิเคราะห์คุณสมบัติของสัญญาณในระดับข้อมูลก่อนการทดสอบเต็มรูปแบบบนแพลตฟอร์มจริงเป็นแนวทางที่มีประโยชน์ การใช้ลักษณะการกระจายของค่าพีเจอร์หรือค่าของสูตรแอลฟาเป็นเครื่องมือคัดกรองเบื้องต้นช่วยลดภาระในการทดสอบสูตรจำนวนมากโดยไม่จำเป็น และช่วยให้กระบวนการสร้างแอลฟามีเหตุผลเชิงวิทยาศาสตร์มากขึ้น กล่าวคือ จากเดิมที่อาจอาศัยการทดลองสูตรจำนวนมากแล้วคัดจากผลลัพธ์ภายหลัง งานวิจัยนี้เสนอกรอบคิดที่เริ่มจากการประเมินว่าแอลฟานั้นมีอำนาจในการจำแนกเชิงข้อมูลเพียงพอหรือไม่ ก่อนจะนำเข้าสู่การทดสอบบนแพลตฟอร์มที่มีต้นทุนเวลาและทรัพยากรมากกว่า

แม้งานวิจัยนี้จะให้ผลลัพธ์ที่สนับสนุนสมมติฐานหลัก แต่ก็ยังมีข้อจำกัดที่ควรกล่าวถึง ประการแรก การใช้ข้อมูลของหลักทรัพย์ในดัชนี S&P 500 เป็นเพียงตัวแทนโดยประมาณของจักรวาลสินทรัพย์ใน WorldQuant BRAIN ซึ่งในทางปฏิบัติอาจมีองค์ประกอบด้านสภาพคล่อง ภูมิภาค หรือเงื่อนไขการคัดกรองสินทรัพย์ที่แตกต่างกัน ประการที่สอง งานวิจัยนี้มุ่งเน้นเฉพาะข้อมูล OHLCV และไม่ได้รวมข้อมูลปัจจัยพื้นฐาน ข้อมูลมหภาค หรือข้อมูลทางเลือกประเภทอื่น ซึ่งอาจช่วยเสริมอำนาจในการจำแนกของสูตรแอลฟาได้ ประการที่สาม การเปรียบเทียบค่าเฉลี่ยของสองกลุ่มสูตรแม้จะให้ภาพรวมที่ชัดเจน แต่ยังเป็นการวิเคราะห์ในระดับกลุ่ม มิได้ลงลึกถึงพลวัตของสูตรรายตัวหรือปฏิสัมพันธ์ระหว่างหลายสัญญาณที่อาจเกิดขึ้นเมื่อสร้างพอร์ตแบบผสม

สำหรับการศึกษาต่อไปในอนาคต อาจขยายผลในหลายทิศทาง เช่น การเพิ่มจำนวนสูตรแอลฟาให้ครอบคลุม

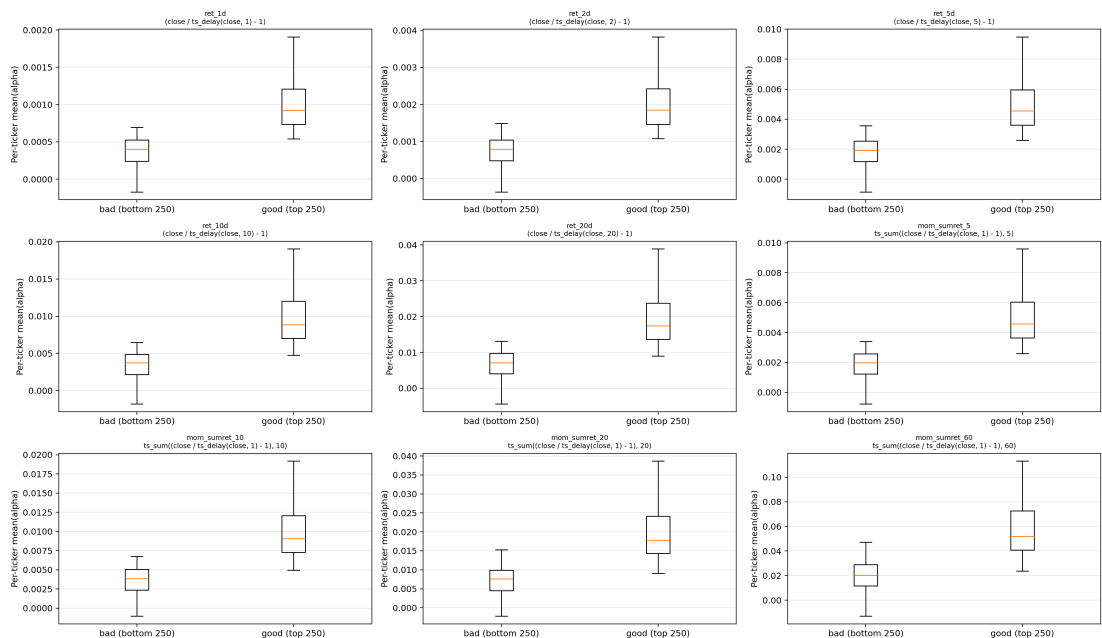
แนวคิดเชิงเทคนิคและเชิงปัจจัยที่หลากหลายมากขึ้น การพิจารณาวิธีการวัดอำนาจในการจำแนกของสัญญาณในเชิงปริมาณที่เข้มข้นขึ้น เช่น การทดสอบความแตกต่างของการกระจาย การใช้โมเดลจำแนกประเภท หรือการวัดความสัมพันธ์ระหว่างอันดับของสัญญาณกับผลตอบแทนในอนาคต ตลอดจนการนำหลายสูตรมารวมกันเป็นชุดสัญญาณ (*ensemble of alphas*) เพื่อศึกษาว่าการผสมสัญญาณที่มีคุณสมบัติแตกต่างกันสามารถเพิ่มคุณภาพของพอร์ตได้หรือไม่ นอกจากนี้ ยังสามารถขยายการศึกษาไปยังจักรวาลสินทรัพย์อื่นหรือช่วงเวลาตลาดที่ต่างต่างกันไปเพื่อประเมินความคงทนของข้อค้นพบ

โดยสรุป งานวิจัยนี้แสดงให้เห็นว่า สูตรแอลฟาที่สร้างจากข้อมูล OHLCV ไม่ได้มีคุณค่าเท่ากันทั้งหมด และความสามารถในการแยกหุ้นผลตอบแทนสูงและต่ำในระดับข้อมูลเป็นคุณสมบัติที่มีความสัมพันธ์กับผลการดำเนินงานจริงของสูตรบนแพลตฟอร์ม WorldQuant BRAIN อย่างมีนัยสำคัญในหลายมิติ โดยเฉพาะผลตอบแทน คุณภาพของสูตร และประสิทธิภาพในการซื้อขาย แม้ผลลัพธ์ในบางมิติของความเสี่ยงยังไม่แตกต่างกันอย่างชัดเจน แต่ข้อค้นพบโดยรวมสนับสนุนแนวคิดที่ว่า การคัดกรองแอลฟาจากคุณสมบัติของสัญญาณก่อนเข้าสู่การทดสอบเชิงพอร์ตเต็มรูปแบบ เป็นแนวทางที่มีประโยชน์ทั้งในเชิงวิชาการและเชิงปฏิบัติสำหรับการพัฒนากลยุทธ์การลงทุนเชิงปริมาณ

ตารางที่ 2: สูตรแอลฟาในกลุ่ม Returns and Momentum

Alpha formula	Mathematical formula	Financial hypothesis
ret_1d	$\alpha_{i,t} = \frac{C_{i,t}}{C_{i,t-1}} - 1$	หุ้นที่มีผลตอบแทนล่าสุดสูงกว่ามีแนวโน้มสะท้อนแรงซื้อระยะสั้นและความแข็งแกร่งเชิงภาคตัดขวาง
ret_2d	$\alpha_{i,t} = \frac{C_{i,t}}{C_{i,t-2}} - 1$	โมเมนตัมระยะสั้นในช่วงสองวันอาจสะท้อนการตอบสนองต่อข่าวหรือแรงซื้อที่ยังต่อเนื่อง
ret_5d	$\alpha_{i,t} = \frac{C_{i,t}}{C_{i,t-5}} - 1$	ผลตอบแทนย้อนหลังหนึ่งสัปดาห์ซื้อขายอาจช่วยจำแนกหุ้นที่กำลังอยู่ในช่วงแนวโน้มแข็งแกร่ง
ret_10d	$\alpha_{i,t} = \frac{C_{i,t}}{C_{i,t-10}} - 1$	ผลตอบแทนย้อนหลังระยะสั้นถึงกลางสะท้อนแรงส่งของราคาและความต่อเนื่องของแนวโน้ม
ret_20d	$\alpha_{i,t} = \frac{C_{i,t}}{C_{i,t-20}} - 1$	ผลตอบแทนย้อนหลังประมาณหนึ่งเดือนอาจใช้แยกหุ้นชนะและหุ้นแพ้ในเชิงโมเมนตัม
mom_sumret_5	$\alpha_{i,t} = \sum_{k=0}^4 \left(\frac{C_{i,t-k}}{C_{i,t-k-1}} - 1 \right)$	ผลตอบแทนสะสม 5 วันช่วยวัดความต่อเนื่องของแรงซื้อขายมากกว่าการดูเพียงวันเดียว
mom_sumret_10	$\alpha_{i,t} = \sum_{k=0}^9 \left(\frac{C_{i,t-k}}{C_{i,t-k-1}} - 1 \right)$	ผลตอบแทนสะสม 10 วันสะท้อน persistence ของแนวโน้มในช่วงเวลายาวขึ้น
mom_sumret_20	$\alpha_{i,t} = \sum_{k=0}^{19} \left(\frac{C_{i,t-k}}{C_{i,t-k-1}} - 1 \right)$	การสะสมผลตอบแทนระยะ 20 วันอาจบ่งชี้หุ้นที่มีความแข็งแกร่งของแนวโน้มอย่างต่อเนื่อง
mom_sumret_60	$\alpha_{i,t} = \sum_{k=0}^{59} \left(\frac{C_{i,t-k}}{C_{i,t-k-1}} - 1 \right)$	ผลตอบแทนสะสมระยะยาวช่วยวัดโมเมนตัมเชิงอนุกรมเวลาที่ฝังอยู่ในราคาหุ้น

Alpha Group 1/5 — 9 WorldQuant BRAIN formulas (boxplot of per-ticker mean since 2019-01-01)



รูปที่ 9: แผนภาพกล่องของกลุ่มสูตรแอลฟา Group 1: Returns / Momentum

ตารางที่ 3: สูตรแอลฟาในกลุ่ม Mean Reversion / Moving Average / Gaps / Intraday

Alpha formula	Mathematical formula	Financial hypothesis
rev_z_5	$\alpha_{i,t} = -\frac{C_{i,t} - \text{MA}_5(C_{i,t})}{\text{STD}_5(C_{i,t})}$	ราคาที่เบี่ยงเบนจากค่าเฉลี่ยระยะสั้นมากอาจมีแนวโน้มกลับเข้าสู่ค่าเฉลี่ย
rev_z_10	$\alpha_{i,t} = -\frac{C_{i,t} - \text{MA}_{10}(C_{i,t})}{\text{STD}_{10}(C_{i,t})}$	ส่วนเบี่ยงเบนจากค่าเฉลี่ยเคลื่อนที่ 10 วันอาจสะท้อนภาวะ overreaction ระยะสั้น
rev_z_20	$\alpha_{i,t} = -\frac{C_{i,t} - \text{MA}_{20}(C_{i,t})}{\text{STD}_{20}(C_{i,t})}$	ราคาที่สูงหรือต่ำเกินค่าเฉลี่ย 20 วันอาจมีแรงดึงดูดกลับเชิงสถิติ
rev_z_60	$\alpha_{i,t} = -\frac{C_{i,t} - \text{MA}_{60}(C_{i,t})}{\text{STD}_{60}(C_{i,t})}$	การเบี่ยงเบนจากค่าเฉลี่ยระยะยาวอาจบ่งชี้โอกาสการกลับตัวเมื่อราคาออกจากสมดุลมากเกินไป
rev_ma_ratio_5	$\alpha_{i,t} = -\left(\frac{C_{i,t}}{\text{MA}_5(C_{i,t})} - 1\right)$	ราคาที่สูงกว่าค่าเฉลี่ย 5 วันมากอาจมีโอกาสน้อยที่ราคาจะต่ำกว่าค่าเฉลี่ยมากอาจมีโอกาสนิ่งตัว
rev_ma_ratio_20	$\alpha_{i,t} = -\left(\frac{C_{i,t}}{\text{MA}_{20}(C_{i,t})} - 1\right)$	อัตราส่วนระหว่างราคาปัจจุบันกับค่าเฉลี่ย 20 วันสะท้อนระดับการยึดตัวของราคาออกจากแนวโน้มหลัก
rev_ma_ratio_60	$\alpha_{i,t} = -\left(\frac{C_{i,t}}{\text{MA}_{60}(C_{i,t})} - 1\right)$	ราคาที่ยังออกจากแนวโน้มระยะยาวมากอาจมีแรงกลับเข้าสู่ค่าเฉลี่ยในเวลาต่อมา
gap_1d	$\alpha_{i,t} = \frac{O_{i,t}}{C_{i,t-1}} - 1$	ช่องว่างระหว่างราคาเปิดกับราคาปิดวันก่อนอาจสะท้อนข้อมูลใหม่ที่เกิดขึ้นข้ามคืน
intraday_ret	$\alpha_{i,t} = \frac{C_{i,t} - O_{i,t}}{O_{i,t}}$	ผลตอบแทนภายในวันสะท้อนแรงซื้อขายสุทธิระหว่างช่วงเปิดถึงปิดตลาด

ตารางที่ 4: สูตรแอลฟาในกลุ่ม Trend / Slope / Price Position

Alpha formula	Mathematical formula	Financial hypothesis
trend_ma_gap_5_20	$\alpha_{i,t} = \frac{\text{MA}_5(C_{i,t})}{\text{MA}_{20}(C_{i,t})} - 1$	ค่าเฉลี่ยระยะสั้นที่สูงกว่าค่าเฉลี่ยระยะกลางบ่งชี้แนวโน้มขาขึ้นและ relative strength
trend_ma_gap_10_60	$\alpha_{i,t} = \frac{\text{MA}_{10}(C_{i,t})}{\text{MA}_{60}(C_{i,t})} - 1$	ความชันของแนวโน้มระยะสั้นเทียบระยะยาวช่วยจำแนกหุ้นที่กำลังอยู่ในเทรนด์ที่แข็งแกร่ง
trend_slope_5	$\alpha_{i,t} = \frac{\text{MA}_5(C_{i,t})}{\text{MA}_5(C_{i,t-5})} - 1$	หากค่าเฉลี่ยระยะสั้นเพิ่มขึ้นเมื่อเทียบกับอดีตแสดงถึงแนวโน้มบวกระยะใกล้
trend_slope_20	$\alpha_{i,t} = \frac{\text{MA}_{20}(C_{i,t})}{\text{MA}_{20}(C_{i,t-20})} - 1$	การเปลี่ยนแปลงของค่าเฉลี่ย 20 วันสะท้อนแนวโน้มระดับกลางที่ต่อเนื่อง
pos_in_range_20	$\alpha_{i,t} = \frac{C_{i,t} - \min(L_{i,t-19}, \dots, L_{i,t})}{\max(H_{i,t-19}, \dots, H_{i,t}) - \min(L_{i,t-19}, \dots, L_{i,t})}$	ราคาปิดที่อยู่ใกล้ด้านบนของกรอบ 20 วันมักสะท้อนแรงซื้อและความแข็งแกร่งเชิงโมเมนตัม
pos_in_range_60	$\alpha_{i,t} = \frac{C_{i,t} - \min(L_{i,t-59}, \dots, L_{i,t})}{\max(H_{i,t-59}, \dots, H_{i,t}) - \min(L_{i,t-59}, \dots, L_{i,t})}$	ตำแหน่งของราคาภายในกรอบ 60 วันช่วยบ่งชี้ว่าหุ้นอยู่ในโซนแข็งแกร่งหรืออ่อนแอเมื่อเทียบกับอดีต
dist_from_low_20	$\alpha_{i,t} = \frac{C_{i,t}}{\min(L_{i,t-19}, \dots, L_{i,t})} - 1$	ราคาที่อยู่สูงกว่าจุดต่ำสุด 20 วันมากอาจสะท้อน downside support และแนวโน้มเชิงบวก
dist_from_high_20	$\alpha_{i,t} = \frac{C_{i,t}}{\max(H_{i,t-19}, \dots, H_{i,t})} - 1$	ราคาที่อยู่ใกล้จุดสูงสุด 20 วันอาจสะท้อน breakout potential และแรงซื้อสะสม
ts_rank_close_20	$\alpha_{i,t} = \text{rank}_{20}(C_{i,t})$	อันดับของราคาปิดภายในหน้าต่าง 20 วันช่วยวัดความแข็งแกร่งของราคาปัจจุบันเมื่อเทียบกับอดีตของตัวเอง

ตารางที่ 5: สูตรแอลฟาในกลุ่ม Volatility / Range / True Range / Ranks

Alpha formula	Mathematical formula	Financial hypothesis
vol_ret_5	$\alpha_{i,t} = \text{STD}_5 \left(\frac{C_{i,t}}{C_{i,t-1}} - 1 \right)$	ความผันผวนของผลตอบแทนระยะสั้นอาจสะท้อนสถานะการค้นหาราคาที่เข้มข้นและความสนใจของตลาดต่อหุ้นรายตัว
vol_ret_10	$\alpha_{i,t} = \text{STD}_{10} \left(\frac{C_{i,t}}{C_{i,t-1}} - 1 \right)$	ความผันผวนย้อนหลัง 10 วันช่วยวัดการแกว่งตัวของราคาในระยะสั้นถึงกลาง และอาจสัมพันธ์กับโอกาสของผลตอบแทนในอนาคต
vol_ret_20	$\alpha_{i,t} = \text{STD}_{20} \left(\frac{C_{i,t}}{C_{i,t-1}} - 1 \right)$	หุ้นที่มีความผันผวนของผลตอบแทนเด่นในระดับรายเดือนอาจอยู่ในช่วงปรับโครงสร้างแนวโน้มหรือการเปลี่ยนระบอบราคา
vol_ret_60	$\alpha_{i,t} = \text{STD}_{60} \left(\frac{C_{i,t}}{C_{i,t-1}} - 1 \right)$	ความผันผวนระยะยาวช่วยบ่งชี้ลักษณะความเสี่ยงเชิงโครงสร้างของหุ้นและความไม่แน่นอนสะสมของราคา
true_range_pct	$\alpha_{i,t} = \frac{\max(H_{i,t}, C_{i,t-1}) - \min(L_{i,t}, C_{i,t-1})}{C_{i,t}}$	ค่าช่วงราคาจริงสะท้อนการเคลื่อนไหวของราคาโดยรวมผลจากช่องว่างข้ามคืนและการแกว่งตัวภายในวันไว้พร้อมกัน
atr_14	$\alpha_{i,t} = \text{MA}_{14} \left(\frac{\max(H_{i,t}, C_{i,t-1}) - \min(L_{i,t}, C_{i,t-1})}{C_{i,t}} \right)$	ค่าเฉลี่ยของ true range ระยะ 14 วันช่วยวัดระดับความผันผวนที่คงอยู่ต่อเนื่องในช่วงเวลาใกล้เคียง
range_pct	$\alpha_{i,t} = \frac{H_{i,t} - L_{i,t}}{C_{i,t}}$	สัดส่วนของช่วงราคาสูงสุด-ต่ำสุดต่อราคาปิดช่วยสะท้อนความรุนแรงของการแกว่งตัวภายในวัน
ts_rank_range_20	$\alpha_{i,t} = \text{rank}_{20} \left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}} \right)$	อันดับของช่วงราคาเมื่อเทียบกับอดีต 20 วันช่วยบ่งชี้ว่าการแกว่งตัวในวันปัจจุบันผิดปกติหรือโดดเด่นเพียงใด
vol_logvol_20	$\alpha_{i,t} = \text{STD}_{20}(\log V_{i,t})$	ความผันผวนของปริมาณซื้อขายในเชิงลอการิทึมอาจสะท้อนการเปลี่ยนแปลงของความสนใจจากตลาดและความไม่เสถียรของสภาพคล่อง

ตารางที่ 6: สูตรแอลฟาในกลุ่ม Breakouts / Stochastic / Candlestick Patterns

Alpha formula	Mathematical formula	Financial hypothesis
stoch_k_10	$\alpha_{i,t} = \frac{C_{i,t} - \min(L_{i,t-9}, \dots, L_{i,t})}{\max(H_{i,t-9}, \dots, H_{i,t}) - \min(L_{i,t-9}, \dots, L_{i,t})}$	ราคาปิดที่อยู่ใกล้ด้านบนของกรอบ 10 วันอาจสะท้อนแรงซื้อและความแข็งแกร่งของแนวโน้มระยะสั้น
stoch_k_20	$\alpha_{i,t} = \frac{C_{i,t} - \min(L_{i,t-19}, \dots, L_{i,t})}{\max(H_{i,t-19}, \dots, H_{i,t}) - \min(L_{i,t-19}, \dots, L_{i,t})}$	ตำแหน่งของราคาปิดภายในกรอบ 20 วันช่วยวัดความแข็งแกร่งสัมพัทธ์ของราคาปัจจุบันเทียบกับอดีต
breakout_high_10	$\alpha_{i,t} = \frac{C_{i,t}}{\max(H_{i,t-9}, \dots, H_{i,t})} - 1$	ราคาปิดที่เข้าใกล้หรือทะลุจุดสูงสุด 10 วันอาจสะท้อนโอกาสของการเกิด breakout ระยะสั้น
breakout_high_20	$\alpha_{i,t} = \frac{C_{i,t}}{\max(H_{i,t-19}, \dots, H_{i,t})} - 1$	การปิดใกล้จุดสูงสุดย้อนหลัง 20 วันอาจบ่งชี้แรงซื้อสะสมและความเป็นไปได้ของการต่อเนื่องของแนวโน้ม
bullish_engulfing	$\alpha_{i,t} = 1(C_{i,t-1} < O_{i,t-1}) 1(C_{i,t} > O_{i,t}) 1(O_{i,t} \leq C_{i,t-1}) 1(O_{i,t} \geq C_{i,t-1})$	รูปแบบ bullish engulfing อาจสะท้อนการเปลี่ยนผ่านจากแรงขายไปสู่แรงซื้อและแนวโน้มการกลับตัวขึ้น
bearish_engulfing	$\alpha_{i,t} = 1(C_{i,t-1} > O_{i,t-1}) 1(C_{i,t} < O_{i,t}) 1(O_{i,t} \geq C_{i,t-1}) 1(O_{i,t} \leq C_{i,t-1})$	รูปแบบ bearish engulfing อาจสะท้อนการเปลี่ยนผ่านจากแรงซื้อไปสู่แรงขายและโอกาสของการกลับตัวลง
engulfing_score	$\alpha_{i,t} = \alpha_{i,t}^{\text{bull}} - \alpha_{i,t}^{\text{bear}}$	ความต่างระหว่างสัญญาณ bullish และ bearish engulfing ใช้สรุปแนวโน้มเชิงสุทธิของรูปแบบแท่งเทียนกลับตัว
hammer	$\alpha_{i,t} = 1(\min(O_{i,t}, C_{i,t}) - L_{i,t} \geq 2 C_{i,t} - O_{i,t}) 1(H_{i,t} - \max(O_{i,t}, C_{i,t}) \leq 2 C_{i,t} - O_{i,t})$	แท่งเทียนแบบ hammer อาจสะท้อนการปฏิเสธราคาต่ำและแรงซื้อกลับที่อาจเกิดขึ้น (หรืออาจบ่งชี้การอ่อนตัวหากขึ้น)
shooting_star	$\alpha_{i,t} = 1(H_{i,t} - \max(O_{i,t}, C_{i,t}) \geq 2 C_{i,t} - O_{i,t}) 1(\min(O_{i,t}, C_{i,t}) - L_{i,t} \leq 2 C_{i,t} - O_{i,t})$	แท่งเทียนแบบ shooting star อาจสะท้อนแรงขายเมื่อราคาปรับตัวขึ้นและแรงซื้อกลับไม่มากเกิดขึ้น (หรืออาจบ่งชี้การอ่อนตัวหากขึ้น)

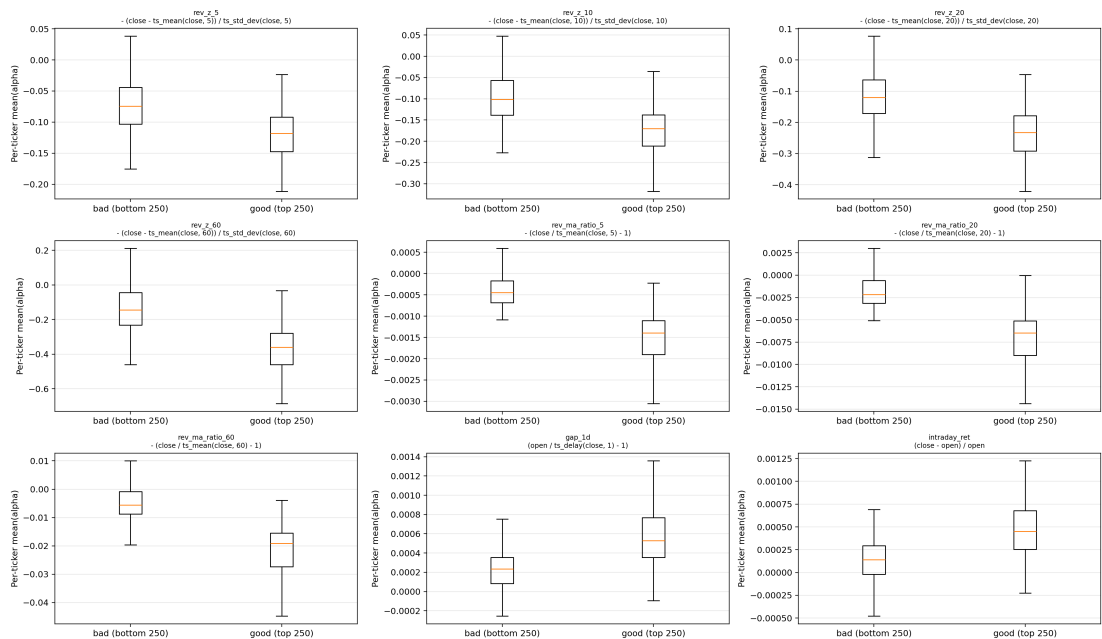
ตารางที่ 7: สูตรแอลฟาในกลุ่ม Dollar Volume / Flow Pressure

Alpha formula	Mathematical formula	Financial hypothesis
dollar_vol_ratio_20	$\alpha_{i,t} = \frac{C_{i,t}V_{i,t}}{MA_{20}(C_{i,t}V_{i,t})} - 1$	มูลค่าการซื้อขายที่สูงกว่าค่าเฉลี่ยระยะ 20 วันอาจสะท้อนความสนใจของตลาดและแรงไหลเข้าของเงินทุน
dollar_vol_ratio_60	$\alpha_{i,t} = \frac{C_{i,t}V_{i,t}}{MA_{60}(C_{i,t}V_{i,t})} - 1$	การเพิ่มขึ้นของ dollar volume เมื่อเทียบกับระยะยาวอาจบ่งชี้การเปลี่ยนแปลงเชิงโครงสร้างของสภาพคล่องและกระแสเงิน
corr_close_volume_20	$\alpha_{i,t} = \text{Corr}_{20}(C_{i,t}, V_{i,t})$	ความสัมพันธ์ระหว่างราคาปิดกับปริมาณซื้อขายในระยะ 20 วันช่วยสะท้อนว่าการขึ้นลงของราคามาพร้อมกับการยืนยันจาก volume หรือไม่
corr_close_volume_60	$\alpha_{i,t} = \text{Corr}_{60}(C_{i,t}, V_{i,t})$	ความสัมพันธ์ระยะยาวระหว่างราคาและปริมาณซื้อขายช่วยวัด persistence ของ flow-driven trend
signed_dollar_flow_10	$\alpha_{i,t} = \frac{\sum_{k=0}^9 \text{sign}\left(\frac{C_{i,t-k}}{C_{i,t-k-1}} - 1\right) C_{i,t-k} V_{i,t-k}}{MA_{10}(C_{i,t}V_{i,t})}$	กระแสเงินแบบมีทิศทางในระยะ 10 วันช่วยวัดว่าเงินทุนสุทธิเข้าออกหุ้นในฝั่งบวกหรือลบมากเพียงใด
signed_dollar_flow_20	$\alpha_{i,t} = \frac{\sum_{k=0}^{19} \text{sign}\left(\frac{C_{i,t-k}}{C_{i,t-k-1}} - 1\right) C_{i,t-k} V_{i,t-k}}{MA_{20}(C_{i,t}V_{i,t})}$	signed dollar flow ระยะ 20 วันสะท้อนแรงซื้อขายสะสมเชิงมูลค่าที่มีทิศทางอย่างต่อเนื่อง
volume_shock_z_20	$\alpha_{i,t} = \frac{V_{i,t} - MA_{20}(V_{i,t})}{STD_{20}(V_{i,t})}$	ปริมาณซื้อขายที่ผิดปกติเมื่อเทียบกับอดีตอาจสะท้อนข่าวใหม่หรือความสนใจจากตลาดที่เพิ่มขึ้นอย่างฉับพลัน
ret_x_volratio20	$\alpha_{i,t} = \left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right) \left(\frac{V_{i,t}}{MA_{20}(V_{i,t})} - 1\right)$	ผลตอบแทนที่เกิดร่วมกับปริมาณซื้อขายสูงกว่าปกติอาจสะท้อน price move ที่มี conviction มากกว่าการเคลื่อนไหวทั่วไป
absret_x_volratio20	$\alpha_{i,t} = \left \frac{C_{i,t}}{C_{i,t-1}} - 1\right \left(\frac{V_{i,t}}{MA_{20}(V_{i,t})} - 1\right)$	ขนาดการเคลื่อนไหวของราคาที่เกิดพร้อม volume สูงช่วยวัดความรุนแรงของแรงกดดันจาก order flow โดยไม่สนใจทิศทาง

ตารางที่ 8: สูตรแอลฟาในกลุ่ม Overnight / Gap Dynamics

Alpha formula	Mathematical formula	Financial hypothesis
overnight_gap	$\alpha_{i,t} = \frac{O_{i,t}}{C_{i,t-1}} - 1$	ช่องว่างระหว่างราคาเปิดกับราคาปิดวันก่อน สะท้อนข่าวใหม่หรือแรงซื้อขายที่เกิดขึ้นข้ามคืน
open_to_prev_high	$\alpha_{i,t} = \frac{O_{i,t}}{H_{i,t-1}} - 1$	ราคาเปิดที่เข้าใกล้หรือสูงกว่าจุดสูงสุดของวันก่อน อาจสะท้อนแรง breakout ตั้งแต่ก่อนเปิดตลาด
open_to_prev_low	$\alpha_{i,t} = \frac{O_{i,t}}{L_{i,t-1}} - 1$	ราคาเปิดที่อยู่สูงกว่าจุดต่ำสุดของวันก่อนมากอาจ บ่งชี้ว่าแรงอ่อนตัวข้ามคืนมีน้อย
close_to_prev_high	$\alpha_{i,t} = \frac{C_{i,t}}{H_{i,t-1}} - 1$	ราคาปิดที่เข้าใกล้หรือทะลุจุดสูงสุดของวันก่อนอาจ สะท้อน continuation ของแรงซื้อ
close_to_prev_low	$\alpha_{i,t} = \frac{C_{i,t}}{L_{i,t-1}} - 1$	ราคาปิดที่อยู่ห่างจากจุดต่ำสุดของวันก่อนมากอาจ สะท้อน closing strength และคุณภาพของแรงซื้อ
gap_reversal_score	$\alpha_{i,t} = - \left(\frac{O_{i,t}}{C_{i,t-1}} - 1 \right) \left(\frac{C_{i,t} - O_{i,t}}{O_{i,t}} \right)$	ช่องว่างข้ามคืนที่ถูกสวนกลับระหว่างวันอาจ สะท้อนแรง reversal หลังการตอบสนองต่อข่าว ช่วงเปิดตลาด
mean_gap_5	$\alpha_{i,t} = \frac{1}{5} \sum_{k=0}^4 \left(\frac{O_{i,t-k}}{C_{i,t-k-1}} - 1 \right)$	ค่าเฉลี่ยของ overnight gap ระยะสั้นช่วยวัด sentiment ข้ามคืนที่สะสมต่อเนื่อง
std_gap_20	$\alpha_{i,t} = \text{STD}_{20} \left(\frac{O_{i,t}}{C_{i,t-1}} - 1 \right)$	ความผันผวนของช่องว่างข้ามคืนอาจสะท้อนการ ไหลเข้าของข้อมูลใหม่และระดับ uncertainty ของ หุ้น
gap_rank_20	$\alpha_{i,t} = \text{rank}_{20} \left(\frac{O_{i,t}}{C_{i,t-1}} - 1 \right)$	อันดับของ overnight gap เทียบกับอดีต 20 วัน ช่วยบ่งชี้ว่าช่องว่างวันนี้ผิดปกติหรือโดดเด่นเพียงใด

Category 2/10 — Mean Reversion / MA / Gaps / Intraday (9 alphas, boxplot of per-ticker mean since 2019-01-01)



รูปที่ 10: แผนภาพกล่องของกลุ่มสูตร Category 2: Mean Reversion / Moving Average / Gaps / Intraday

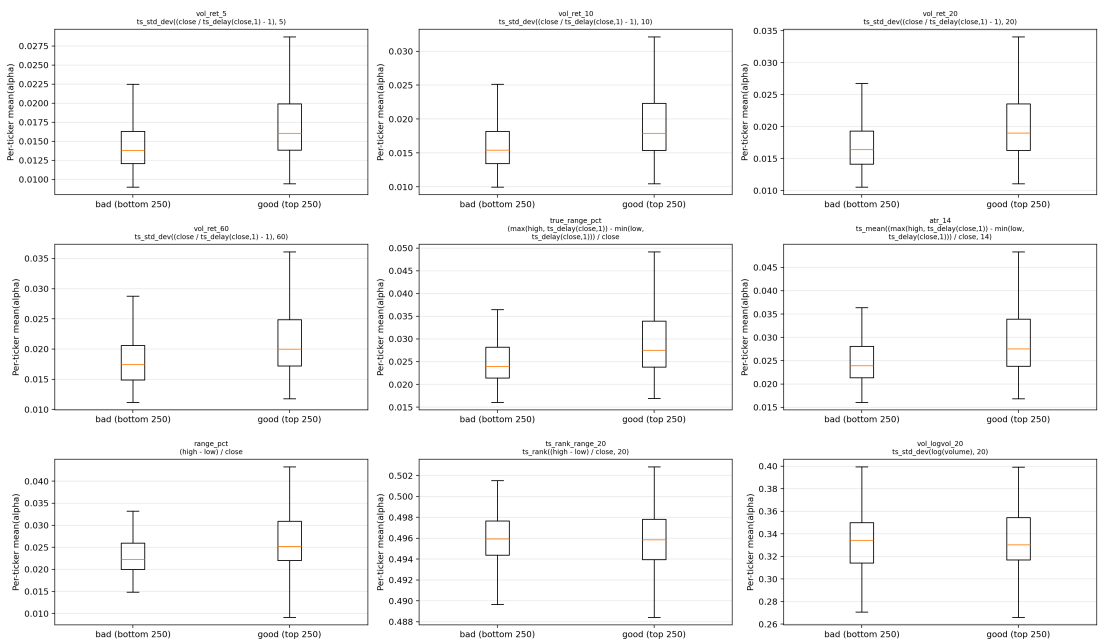
ตารางที่ 9: สูตรแอลฟาในกลุ่ม Volatility Regime / Compression

Alpha formula	Mathematical formula	Financial hypothesis
range_z_20	$\alpha_{i,t} = \frac{\frac{H_{i,t} - L_{i,t}}{C_{i,t}} - \text{MA}_{20}\left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}}\right)}{\text{STD}_{20}\left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}}\right)}$	ช่วงราคาภายในวันที่สูงหรือต่ำผิดปกติเมื่อเทียบกับอดีต 20 วันอาจสะท้อนการเปลี่ยนระบอบความผันผวนของหุ้น
vol_ratio_20_60	$\alpha_{i,t} = \frac{\text{STD}_{20}\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right)}{\text{STD}_{60}\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right)}$	อัตราส่วนความผันผวนระยะกลางต่อระยะยาวช่วยบ่งชี้ว่าหุ้นกำลังเข้าสู่ช่วง volatility expansion หรือ compression
vol_ratio_5_20	$\alpha_{i,t} = \frac{\text{STD}_5\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right)}{\text{STD}_{20}\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right)}$	ความผันผวนระยะสั้นที่สูงขึ้นเมื่อเทียบกับระยะกลางอาจสะท้อนแรงเร่งของการเคลื่อนไหวราคา
vol_spread_5_20	$\alpha_{i,t} = \text{STD}_5\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right) - \text{STD}_{20}\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right)$	ส่วนต่างระหว่างความผันผวนระยะสั้นและระยะกลางช่วยวัดการเปลี่ยนผ่านของสภาวะตลาดได้โดยตรง
range_ratio_5_20	$\alpha_{i,t} = \frac{\text{MA}_5\left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}}\right)}{\text{MA}_{20}\left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}}\right)}$	อัตราส่วนของช่วงราคาเฉลี่ยระยะสั้นต่อระยะกลางใช้ตรวจจับการบีบตัวหรือขยายตัวของ trading range
atr_ratio_14_60	$\alpha_{i,t} = \frac{\text{MA}_{14}\left(\frac{\max(H_{i,t}, C_{i,t-1}) - \min(L_{i,t}, C_{i,t-1})}{C_{i,t}}\right)}{\text{MA}_{60}\left(\frac{\max(H_{i,t}, C_{i,t-1}) - \min(L_{i,t}, C_{i,t-1})}{C_{i,t}}\right)}$	อัตราส่วนของ average true range ระยะสั้นต่อระยะยาวสะท้อนการเปลี่ยนแปลงของ realized volatility regime
ret_to_range_20	$\alpha_{i,t} = \frac{\left \frac{C_{i,t}}{C_{i,t-1}} - 1\right }{\text{MA}_{20}\left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}}\right)}$	ผลตอบแทนรายวันที่สูงเมื่อเทียบกับช่วงราคาเฉลี่ยอาจสะท้อน directional conviction ที่แข็งแกร่งผิดปกติ
range_breakout_pressure_20	$\alpha_{i,t} = \left(\frac{H_{i,t} - L_{i,t}}{C_{i,t}}\right) \left(\frac{C_{i,t} - \min(L_{i,t-19}, \dots, L_{i,t})}{\max(H_{i,t-19}, \dots, H_{i,t}) - \min(L_{i,t-19}, \dots, L_{i,t})}\right)$	ช่วงราคากว้างร่วมกับการปิดใกล้ด้านบนของกรอบ breakout pressure เชิงบวก
ts_rank_neg_vol5_60	$\alpha_{i,t} = \text{rank}_{60}\left(-\text{STD}_5\left(\frac{C_{i,t}}{C_{i,t-1}} - 1\right)\right)$	อันดับตามเวลาของค่าความผันผวนระยะสั้นที่ต่ำเครื่องหมายลบใช้ตรวจจับช่วง volatility compression ซึ่งอาจนำไปสู่การเคลื่อนไหวรุนแรงในอนาคต

ตารางที่ 10: สูตรแอลฟาในกลุ่ม Candlestick Body / Wick Structure

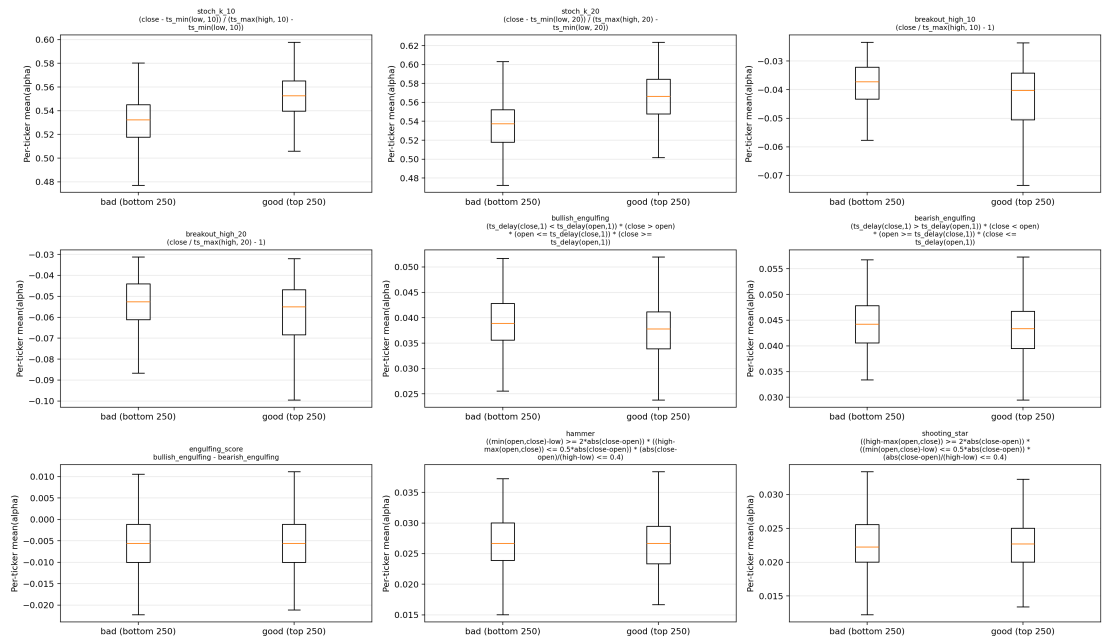
Alpha formula	Mathematical formula	Financial hypothesis
body_pct	$\alpha_{i,t} = \frac{ C_{i,t} - O_{i,t} }{C_{i,t}}$	ขนาดของตัวแท่งเทียนเมื่อเทียบกับราคาปิดสะท้อนความแรงของการเคลื่อนไหวสุทธิภายในวัน
upper_wick_pct	$\alpha_{i,t} = \frac{H_{i,t} - \max(O_{i,t}, C_{i,t})}{C_{i,t}}$	ใส่เทียนบนที่ยาวอาจสะท้อนแรงขายหรือแรงทำกำไรเมื่อราคาขึ้นไปแตะระดับสูงสุดระหว่างวัน
lower_wick_pct	$\alpha_{i,t} = \frac{\min(O_{i,t}, C_{i,t}) - L_{i,t}}{C_{i,t}}$	ใส่เทียนล่างที่ยาวอาจสะท้อนแรงซื้อกลับหลังจากราคาปรับลงลิ้นระหว่างวัน
body_to_range	$\alpha_{i,t} = \frac{ C_{i,t} - O_{i,t} }{H_{i,t} - L_{i,t}}$	สัดส่วนของตัวแท่งต่อช่วงราคาทั้งวันช่วยวัดว่าการเคลื่อนไหวภายในวันมีทิศทางชัดเจนเพียงใด
close_location	$\alpha_{i,t} = \frac{C_{i,t} - L_{i,t}}{H_{i,t} - L_{i,t}}$	ราคาปิดที่อยู่ใกล้จุดสูงของวันอาจสะท้อน closing strength และแรงซื้อที่ยังคงอยู่จนสิ้นวัน
open_location	$\alpha_{i,t} = \frac{O_{i,t} - L_{i,t}}{H_{i,t} - L_{i,t}}$	ตำแหน่งของราคาเปิดภายในกรอบวันช่วยสะท้อน sentiment เริ่มต้นของตลาดในวันนั้น
wick_imbalance	$\alpha_{i,t} = \frac{(H_{i,t} - \max(O_{i,t}, C_{i,t})) - (\min(O_{i,t}, C_{i,t}) - L_{i,t})}{C_{i,t}}$	ความไม่สมดุลระหว่างใส่เทียนบนและล่างช่วยวัดว่าระหว่างวันแรงขายหรือแรงซื้อเป็นฝ่ายกดดันมากกว่า
marubozu	$\alpha_{i,t} = 1 \left(\frac{ C_{i,t} - O_{i,t} }{H_{i,t} - L_{i,t}} \geq 0.8 \right)$	แท่งเทียนแบบ marubozu บ่งชี้วันที่ราคามีทิศทางชัดเจนและมีใส่เทียนสั้นมาก อาจสะท้อน conviction สูงของตลาด
doji	$\alpha_{i,t} = 1 \left(\frac{ C_{i,t} - O_{i,t} }{H_{i,t} - L_{i,t}} \leq 0.1 \right)$	แท่งเทียนแบบ doji สะท้อนความลังเลของตลาดและภาวะสมดุลชั่วคราวระหว่างแรงซื้อและแรงขาย

Category 3/10 — Volatility / Range / True Range / Ranks (9 alphas, boxplot of per-ticker mean since 2019-01-01)



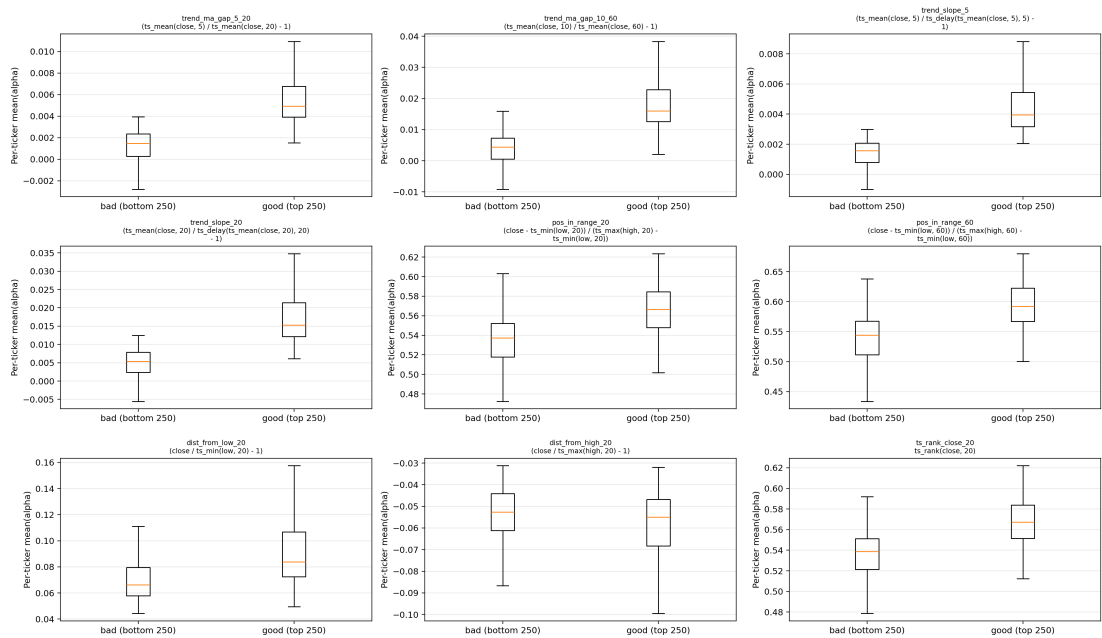
รูปที่ 11: แผนภาพกล่องของกลุ่มสูตร Category 3: Volatility / Range / True Range / Ranks

Category 5/10 — Breakouts / Stochastic / Candlestick Patterns (9 alphas, boxplot of per-ticker mean since 2019-01-01)



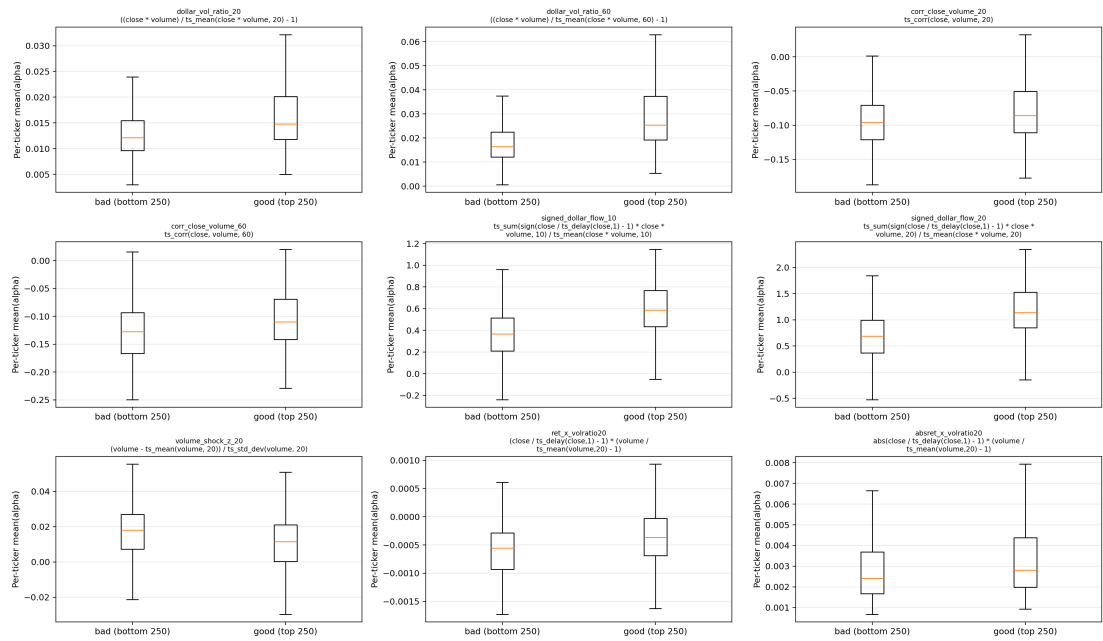
รูปที่ 12: แผนภาพกล่องของกลุ่มสูตร Category 5: Breakouts / Stochastic / Candlestick Patterns

Category 6/10 — Trend / Slope / Price Position (9 alphas, boxplot of per-ticker mean since 2019-01-01)



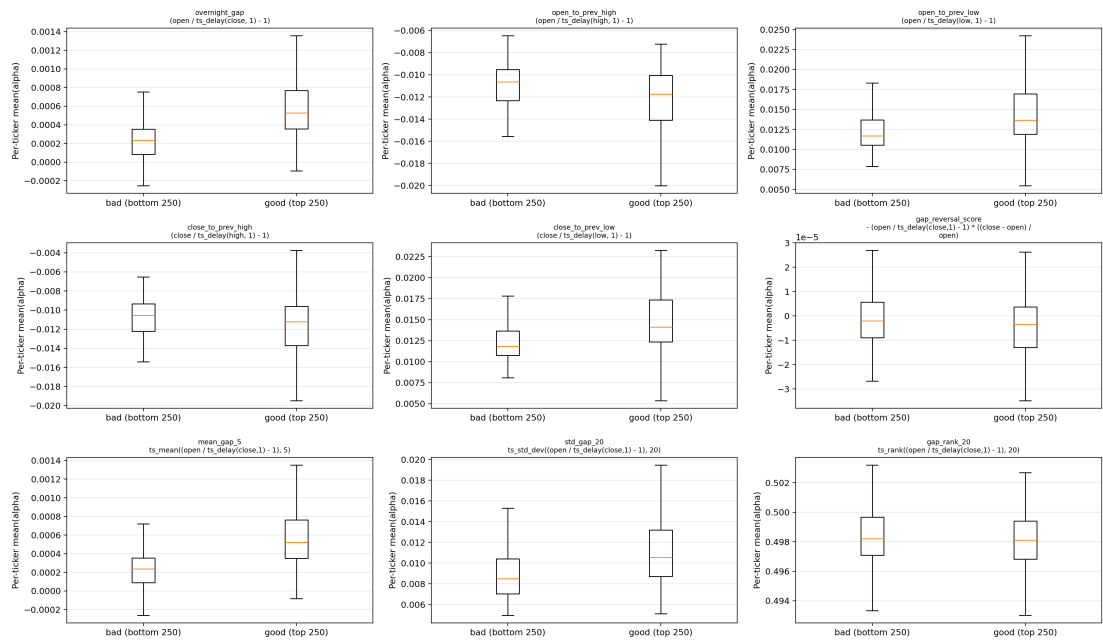
รูปที่ 13: แผนภาพกล่องของกลุ่มสูตร Category 6: Trend / Slope / Price Position

Category 7/10 — Dollar Volume / Flow Pressure (9 alphas, boxplot of per-ticker mean since 2019-01-01)



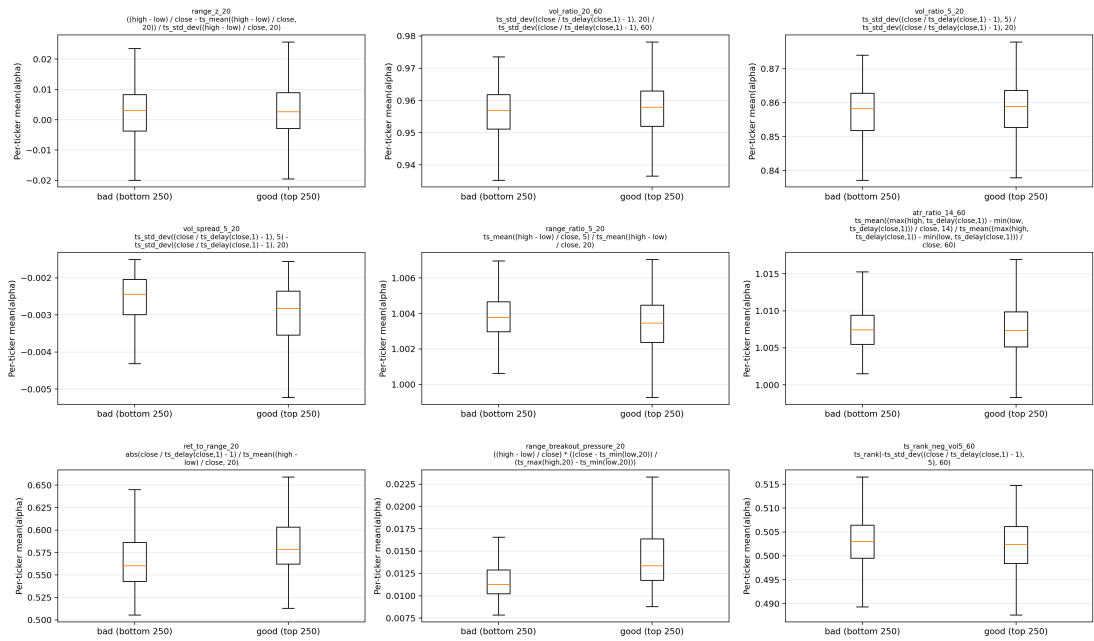
รูปที่ 14: แผนภาพกล่องของกลุ่มสูตร Category 7: Dollar Volume / Flow Pressure

Category 8/10 — Overnight / Gap Dynamics (9 alphas, boxplot of per-ticker mean since 2019-01-01)



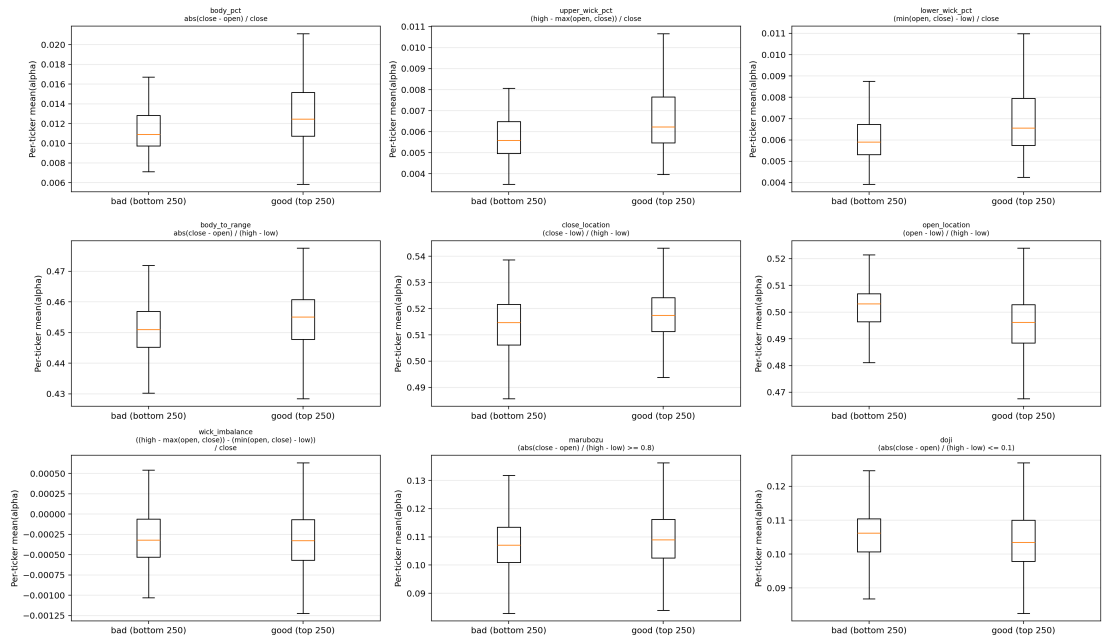
รูปที่ 15: แผนภาพกล่องของกลุ่มสูตร Category 8: Overnight / Gap Dynamics

Category 9/10 — Volatility Regime / Compression (9 alphas, boxplot of per-ticker mean since 2019-01-01)



รูปที่ 16: แผนภาพกล่องของกลุ่มสูตร Category 9: Volatility Regime / Compression

Category 10/10 — Candlestick Body / Wick Structure (9 alphas, boxplot of per-ticker mean since 2019-01-01)



รูปที่ 17: แผนภาพกล่องของกลุ่มสูตร Category 10: Candlestick Body / Wick Structure

ตารางที่ 11: ผลการเปรียบเทียบสูตร Alpha บนแพลตฟอร์ม WorldQuant BRAIN สำหรับกรณีที่การแผนภาพ
กล่องที่มีระดับของช่วงตัวต่างกัน * หมายถึงผ่านเกณฑ์ของ WorldQuant Brain

สูตร	Sharpe	Turnover	Fitness	Returns	Drawdown	Margin
$-(\text{close} / \text{ts_delay}(\text{close}, 1) - 1)$	0.91	81.51%	0.41	16.82%*	25.26%*	4.13%*
$-(\text{close} / \text{ts_delay}(\text{close}, 2) - 1)$	0.97	62.76%*	0.52	18.21%*	24.97%*	5.80%*
$-(\text{close} / \text{ts_delay}(\text{close}, 5) - 1)$	1.12	40.40%*	0.83	22.01%*	19.04%*	10.89%*
$-(\text{close} / \text{ts_delay}(\text{close}, 10) - 1)$	0.80	27.04%*	0.64	17.29%*	23.83%*	12.79%*
$-(\text{close} / \text{ts_delay}(\text{close}, 20) - 1)$	0.49	18.83%*	0.33	8.70%*	23.92%*	9.24%*
$-(\text{ts_sum}((\text{close} / \text{ts_delay}(\text{close}, 1) - 1), 5))$	1.08	40.41%*	0.76	20.24%*	18.15%*	10.02%*
$-(\text{ts_sum}((\text{close} / \text{ts_delay}(\text{close}, 1) - 1), 10))$	0.64	26.99%*	0.43	12.15%*	23.99%*	9.00%*
$-(\text{ts_sum}((\text{close} / \text{ts_delay}(\text{close}, 1) - 1), 20))$	0.84	18.76%*	0.76	15.21%*	22.96%*	16.22%*
$-(\text{ts_sum}((\text{close} / \text{ts_delay}(\text{close}, 1) - 1), 60))$	0.90	10.60%*	1.07*	17.56%*	38.71%*	33.12%*
$-(\text{ts_mean}(\text{close}, 5) / \text{ts_mean}(\text{close}, 20) - 1)$	0.35	18.71%*	0.20	6.35%*	26.25%*	6.79%*
$-(\text{ts_mean}(\text{close}, 10) / \text{ts_mean}(\text{close}, 60) - 1)$	0.34	8.23%*	0.24	6.33%*	23.01%*	15.38%*
$-(\text{ts_mean}(\text{close}, 5) / \text{ts_delay}(\text{ts_mean}(\text{close}, 5), 5) - 1)$	0.63	31.11%*	0.40	12.44%*	26.71%*	8.00%*
$-(\text{ts_mean}(\text{close}, 20) / \text{ts_delay}(\text{ts_mean}(\text{close}, 20), 20) - 1)$	0.31	9.68%*	0.21	5.59%*	23.44%*	11.55%*
$-(\text{close} / \text{ts_mean}(\text{close}, 5) - 1)$	0.97	50.70%*	0.58	18.14%*	21.97%*	7.15%*
$-(\text{close} / \text{ts_mean}(\text{close}, 20) - 1)$	0.58	22.59%*	0.40	10.98%*	22.07%*	9.72%*
$-(\text{close} / \text{ts_mean}(\text{close}, 60) - 1)$	0.57	12.56%*	0.53	11.03%*	21.61%*	17.56%*
$-(\text{ts_mean}((\text{open} / \text{ts_delay}(\text{close}, 1) - 1), 5))$	0.49	39.69%*	0.22	7.81%*	25.72%*	3.93%*

ตารางที่ 12: ผลการเปรียบเทียบสูตร Alpha บนแพลตฟอร์ม WorldQuant BRAIN สำหรับกรณีที่การแผนภาพ
กล่องที่มีระดับของช่วงตัวไม่ต่างกัน * หมายถึงผ่านเกณฑ์ของ WorldQuant Brain

Name	Formula	Sharpe	Turnover	Fitness	Returns	Drawdown	Margin
range_z_20	((high-low)/close - ts_mean((high-low)/close, 20)) / ts_std_dev((high-low)/close, 20)	0.56	72.36%	0.13	3.86%	16.60%	1.07%
vol_ratio_20_60	-ts_std_dev((close / ts_delay(close, 1) - 1), 20) / ts_std_dev((close / ts_delay(close, 1) - 1), 60)	0.05	16.46%	0.01	0.32%	11.17%	0.38%
vol_ratio_5_20	ts_std_dev((close / ts_delay(close, 1) - 1), 5) / ts_std_dev((close / ts_delay(close, 1) - 1), 20)	0.04	41.33%	0.00	0.26%	13.46%	0.12%
are_ratio_14_60	ts_mean((max(high, its_deley(close, 1)) - min(low, ts_delay(close, 1))) / close, 14) / ts_mean((max(high, its_delay(close, 1)) - min(low, ts_delay(close, 1))) / close, 60)	0.05	18.55%	0.01	0.47%	17.46%	0.50%
doji	-(abs(close-open)/(high-low) <= 0.1)	0.06	73.33%	0.00	0.22%	5.15%	0.06%
gap_rank_20	-ts_rank((open / its_delay(close, 1) - 1), 20)	1.14	80.79%	0.35	7.76%	40.80%	1.92%
shooting_star	-((high-max(open, close)) >= 2*abs(close-open)) * ((min(open, close)-low) <= 0.5*abs(close-open)) * (abs(close-open)/(high-low) <= 0.4)	0.61	73.49%	0.17	5.84%	33.96%	1.59%
hammer	((min(open, close)-low) >= *abs(close-open)) * ((high-max(open, close)) <= 0.5*abs(close-open)) * (abs(close-open)/(high-low) <= 0.4)	0.61	72.70%	0.14	4.04%	9.72%	1.11%
ts_rank_range_20	ts_rank((high-low)/close, 20)	0.72	73.56%	0.18	4.43%	14.55%	1.21%
vol_ratio_20_60	ts_std_dev(log(volume), 20)	0.06	16.92%	0.02	1.21%	39.61%	1.43%
bearish_engulfing	ts_jelly(close, 1) > ts_delay(open, 1)) * (close < open) * (open >= ts_delay(close, 1)) * (close <= ts_delay(open, 1))	0.02	72.30%	0.00	0.10%	17.34%	0.03%
range_ratio_20	-ts_mean((high-low)/close, 5) / ts_mean((hig-low)/close, 20)	0.13	37.61%	0.02	1.08%	16.20%	0.57%
ret_to_range_20	abs(close / ts_delay(close, 1) - 1) / ts_mean((high-low)/close, 20)	0.01	76.87%	0.00	0.10%	20.30%	0.03%
marubozu	(abs(close-open)/(high-low) >= 0.8)	0.44	73.59%	0.08	2.22%	8.49%	0.60%
close_location	-(close-low)/(high-low)	1.56	80.41%	0.59	11.39%	58.30%	2.83%
upper_wick_pct	-(high-max(open, close))/close	0.37	60.26%	0.14	8.71%	100.20%	2.89%
corr_close_volume_20	corr(close, volume, 20)	1.04	20.00%	0.55	5.63%	9.15%	5.63%

เอกสารอ้างอิง

- [1] Zura Kakushadze. 101 formulaic alphas. *Wilmott*, 2016(84):72–81, 2016.
- [2] Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.
- [3] William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.
- [4] John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–37, 1965.
- [5] Jan Mossin. Equilibrium in a capital asset market. *Econometrica*, 34(4):768–783, 1966.
- [6] Fischer Black. Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3):444–455, 1972.
- [7] André F. Perold. The capital asset pricing model. *Journal of Economic Perspectives*, 18(3):3–24, 2004.
- [8] Michael C. Jensen. The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, 23(2):389–416, 1968.
- [9] Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.
- [10] Mark M. Carhart. On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82, 1997.
- [11] William Brock, Josef Lakonishok, and Blake LeBaron. Simple technical trading rules and the stochastic properties of stock returns. *The Journal of Finance*, 47(5):1731–1764, 1992.
- [12] Andrew W. Lo, Harry Mamaysky, and Jiang Wang. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*, 55(4):1705–1765, 2000.
- [13] Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91, 1993.
- [14] Andrew W. Lo and A. Craig MacKinlay. When are contrarian profits due to stock market overreaction? *The Review of Financial Studies*, 3(2):175–205, 1990.
- [15] Clifford S. Asness, Tobias J. Moskowitz, and Lasse Heje Pedersen. Value and momentum everywhere. *The Journal of Finance*, 68(3):929–985, 2013.

- [16] Tobias J. Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. Time series momentum. *Journal of Financial Economics*, 104(2):228–250, 2012.
- [17] Frank J. Fabozzi, Sergio M. Focardi, and Petter N. Kolm. *Quantitative Equity Investing: Techniques and Strategies*. John Wiley & Sons, 2010.
- [18] Zvi Bodie, Alex Kane, and Alan J. Marcus. *Investments*. McGraw-Hill Education, 11 edition, 2018.
- [19] Richard C. Grinold. The fundamental law of active management. *The Journal of Portfolio Management*, 15(3):30–37, 1989.
- [20] Jack L. Treynor and Fischer Black. How to use security analysis to improve portfolio selection. *The Journal of Business*, 46(1):66–86, 1973.
- [21] Evan Gatev, William N. Goetzmann, and K. Geert Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, 19(3):797–827, 2006.
- [22] Andrea Frazzini and Lasse Heje Pedersen. Betting against beta. *Journal of Financial Economics*, 111(1):1–25, 2014.