การพยากรณ์ราคาทองคำโดยใช้เทคนิคการเรียนรู้ของเครื่องแบบกลุ่ม

FORECASTING GOLD PRICE USING ENSEMBLE MACHINE LEARNING TECHNIQUES

ธนกฤต พลอยเกตุ

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2567

การพยากรณ์ราคาทองคำโดยใช้เทคนิคการเรียนรู้ของเครื่องแบบกลุ่ม

ธนกฤต พลอยเกตุ

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2567
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

FORECASTING GOLD PRICE USING ENSEMBLE MACHINE LEARNING TECHNIQUES

THANAKRIT PLOYKATE

An Independent Study Submitted in Partial Fulfillment of the Requirements

for the Degree of MASTER OF SCIENCE

(Data Science)

Faculty of Science, Srinakharinwirot University

2024

Copyright of Srinakharinwirot University

สารนิพนธ์

เรื่อง

การพยากรณ์ราคาทองคำโดยใช้เทคนิคการเรียนรู้ของเครื่องแบบกลุ่ม

ของ

ธนกฤต พลอยเกตุ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

คณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบปากเปล่าสารนิพนธ์

..... ที่ปรึกษาหลัก

(อาจารย์ ดร.บรรพตรี คมขำ)

..... ประธาน

(ผู้ช่วยศาสตราจารย์ ดร.จันตรี ผลประเสริฐ)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ศิริสรรพ เหล่าหะเกียรติ)

ชื่อเรื่อง	การพยากรณ์ราคาทองคำโดยใช้เทคนิคการเรียนรู้ของเครื่องแบบกลุ่ม
ผู้วิจัย	ธนกฤต พลอยเกตุ
ปริญญา	วิทยาศาสตรมหาบัณฑิต
ปีการศึกษา	2567
อาจารย์ที่ปรึกษา	อาจารย์ ดร. บรรพตรี คมขำ

การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อพยากรณ์ราคาทองคำโดยใช้เทคนิคการเรียนรู้ของเครื่องแบบกลุ่ม เพื่อเพิ่มความแม่นยำในการทำนาย โดยใช้ข้อมูลทางการเงินที่มีความสัมพันธ์อย่างใกล้ชิดกับราคาทองคำ ได้แก่ ดัชนีตลาดหุ้น อัตราแลกเปลี่ยน และอัตราผลตอบแทนของพันธบัตร ในการสร้างแบบจำลองการพยากรณ์ งานวิจัยนี้ประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องที่หลากหลาย ได้แก่ ซัพพอร์ตเวกเตอร์รีเกรสชัน, การถดถอยแบบต้นไม้ตัดสินใจ, การถดถอยของป่าแบบสุ่ม, การถดถอยลาสโซ, การถดถอยแบบริดจ์, การถดถอยของริดจ์แบบเบส์, การถดถอยแบบกราเดียนส์บูสติง และสโคแคสติกกราเดียนเดสเซนท์ พร้อมทั้งใช้เทคนิคการคัดเลือกคุณลักษณะ ได้แก่ การถดถอยลาสโซ การถดถอยของป่าแบบสุ่ม และการค้นหาขั้นตอนเชิงพันธุกรรม เพื่อเลือกตัวแปรที่มีความสำคัญสูงสุด จากนั้นนำแบบจำลองที่มีประสิทธิภาพสูงมารวมกันโดยใช้เทคนิคการรวมโมเดลแบบ Voting และ Stacking ผลการทดลองแสดงให้เห็นว่าแบบจำลองที่รวมโมเดลสามารถให้ผลการพยากรณ์ที่แม่นยำกว่าแบบจำลองเดี่ยวในทุกช่วงเวลา โดยพิจารณาจากค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ ซึ่งแสดงให้เห็นว่าเทคนิคการรวมโมเดลช่วยเพิ่มความน่าเชื่อถือในการพยากรณ์ราคาทองคำได้อย่างมีนัยสำคัญ แบบจำลองที่พัฒนาขึ้นสามารถนำไปประยุกต์ใช้ในการตัดสินใจลงทุน การประเมินความเสี่ยงทางการเงิน และการวางกลยุทธ์ทางเศรษฐกิจในระดับมหภาคได้อย่างมีประสิทธิภาพ

คำสำคัญ : การพยากรณ์ราคาทองคำ, การเรียนรู้ของเครื่อง, การรวมโมเดล, การคัดเลือกคุณลักษณะ, แบบจำลองการถดถอย

Title	FORECASTING GOLD PRICE USING ENSEMBLE MACHINE LEARNING TECHNIQUES
Author	THANAKRIT PLOYKATE
Degree	MASTER OF SCIENCE
Academic Year	2024
Thesis Advisor	Dr. Banphatree Khomkham

This study aims to forecast gold prices using ensemble machine learning techniques to enhance prediction accuracy. Financial data that are closely related to gold prices, such as stock indices, exchange rates, and bond yields, were utilized to build the forecasting models. Various machine learning regression algorithms were applied, including Support Vector Regression, Decision Tree Regression, Random Forest, Lasso Regression, Ridge Regression, Bayesian Ridge Regression, Gradient Boosting, and Stochastic Gradient Descent. To improve model efficiency, feature selection techniques such as Lasso Regression, Random Forest, and Genetic Algorithms were employed to identify the most significant variables. The best-performing models were then combined using ensemble techniques, specifically Voting and Stacking. Experimental results demonstrate that the ensemble models provide more accurate forecasts than individual models across all time horizons, as measured by Root Mean Square Error (RMSE) and the R-squared metrics. These findings indicate that ensemble learning significantly enhances the reliability of gold price forecasting. The developed models can be effectively applied in investment decision-making, financial risk assessment, and macroeconomic strategy planning

Keyword : Gold Price Forecasting, Machine Learning, Ensemble Learning, Feature Selection, Regression Models

กิตติกรรมประกาศ

ข้าพเจ้าขอแสดงความขอบคุณอาจารย์ ดร.บรรพตรี คมขำ อาจารย์ที่ปรึกษาหลัก ที่ได้ให้คำแนะนำ ความรู้ ความเอาใจใส่ และการสนับสนุนตลอดกระบวนการทำการค้นคว้าอิสระฉบับนี้ด้วยความเมตตาและกรุณาอย่างสูง

ข้าพเจ้าขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.จันตรี ผลประเสริฐ ประธานกรรมการสอบ และ ผู้ช่วยศาสตราจารย์ ดร.ศิริสรรพ เหล่าหะเกียรติ กรรมการสอบ สำหรับข้อเสนอแนะอันมีคุณค่าและการตรวจสอบผลงานอย่างรอบคอบ ซึ่งช่วยให้ผลงานมีความสมบูรณ์ยิ่งขึ้น

ขอขอบคุณคณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ และอาจารย์ทุกท่านในหลักสูตรวิทยาการข้อมูล สำหรับความรู้และประสบการณ์อันมีคุณค่าที่ได้ถ่ายทอดตลอดระยะเวลาการศึกษา

สุดท้ายนี้ ข้าพเจ้าขอขอบคุณครอบครัวและเพื่อน ๆ ที่ให้กำลังใจ สนับสนุน และเป็นแรงผลักดันที่สำคัญตลอดระยะเวลาการศึกษาและการทำวิจัยนี้

ธนกฤต พลอยเกตุ

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ฐ

สารบัญรูปภาพ ..... ต

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของงานวิจัย 2

1.3 ขอบเขตงานวิจัย 3

1.4 สมมติฐานการวิจัย 3

1.5 ประโยชน์ที่คาดว่าจะได้รับ 4

บทที่ 2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง 5

2.1. ทฤษฎีที่เกี่ยวข้อง 6

2.1.1. Moving Average Convergence-Divergence (MACD) 6

2.1.2. ดัชนีความแข็งแกร่งสัมพัทธ์ (Relative Strength Index, RSI) 7

2.1.3. ค่าเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average) 8

2.1.4. เทคนิคการเลือกคุณลักษณะ (Feature Selection) 9

2.1.4.1. การค้นหาแบบขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithms) 9

2.1.5. แบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) 10

2.1.5.1. ซัพพอร์ตเวกเตอร์รีเกสชัน (Support Vector Regression) 10

2.1.5.2. การถดถอยแบบต้นไม่ตัดสินใจ (Decision Tree Regression) 11
2.1.5.3. การถดถอยของป่าแบบสุ่ม (Random Forest Regression) 11
2.1.5.4. การถดถอนลาสโซ (Lasso Regression) 12
2.1.5.5. การถดถอยแบบริดจ์ (Ridge Regression) 13
2.1.5.6. การถดถอยแบบริดจ์ของเบส์ (Bayesian Ridge Regression) 13
2.1.5.7. การถดถอยแบบกราเดียนส์บูสต์ติง (Gradient Boosting Regressor) 14
2.1.5.8. สโตแคสติกแกรเดียนต์เดสเซนท์ (Stochastic Gradient Decent) 15
2.1.6. เทคนิคการรวมโมเดล (Ensemble Method) 16
2.1.6.1. Bagging 16
2.1.6.2. Boosting 17
2.1.6.3. Stacking 18
2.1.7. เครื่องมือการแปรผล 18
2.1.7.1. ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error, RMSE) 18
2.1.7.2. ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple Determination, R-Squared) 19
2.1.7.3. ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉี่ย (Mean Absolute Percent Error, MAPE) 20
2.2. งานวิจัยที่เกี่ยวข้อง 20
บทที่ 3 การดำเนินการวิจัย 24
3.1. กระบวนการทำงานของแบบจำลอง (Model Overview Workflow) 25
3.2. กระบวนการสำรวจข้อมูล (Exploratory Data Analysis) 26
3.2.1. การรวบรวมข้อมูล (Data Collection) 26
3.2.1.1. กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ (Gold ETF, GLD) 27

3.2.1.2. ดัชนีตลาดหุ้นสหรัฐอเมริกา (Standard&Poor's 500 Index, S&P500) .. 28
3.2.1.3. ดัชนีดาวโจนส์ (Dow Jones ETF, DJI) 30
3.2.1.4. ดัชนีหุ้นเอลโดราโด โกลด์ (Eldorado Gold Corporation, EGO) 31
3.2.1.5. กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ (VanEck Gold Miners ETF, GDX)..... 32
3.2.1.6. ดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ (W&T Offshore Stock, WTI)..... 33
3.2.1.7. กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา (United States Brent Oil Fund, BNO) 35
3.2.1.8. กองทุนรวมราคาน้ำมันสหรัฐอเมริกา (United States Oil Fund, USO) .. 36
3.2.1.9. อัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว (United States 10-Year Bond Yield) 37
3.2.1.10. อัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ (Euro US Dollar, EURUSD) 39
3.2.1.11. ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา (US Dollar Index)..... 41
3.2.1.12. สัญญาซื้อขายแร่เงินล่วงหน้า (Silver Futures) 43
3.2.1.13. สัญญาซื้อขายแร่แพลทินัมล่วงหน้า (Platinum Futures) 44
3.2.1.14. สัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า (Palladium Futures) 46
3.2.2. การจัดเก็บข้อมูล (Data Storage) 48
3.2.3. การสำรวจข้อมูล (Data Exploratory)..... 50
3.2.3.1. การพิจารณาคุณลักษณะเดี่ยว (Univariant Data)..... 50
3.2.3.2. การพิจารณาความสัมพันธ์ (Correlation)..... 63
3.2.3. กระบวนการสร้างและปรับปรุงคุณลักษณะ (Feature Engineering) 72
3.2.4. กระบวนการแบ่งชุดข้อมูล (Splitting Dataset)..... 76
3.2.5. กระบวนการคัดเลือกคุณลักษณะ (Feature Selection) 78

3.2.5. การตั้งค่าและการฝึกฝนแบบจำลอง (Model Setup and Training) 83
3.2.5.1. ซัพพอร์ตเวกเตอร์รีเกสชัน (Support Vector Regression) 83
3.2.5.2. การถดถอยแบบต้นไม้ตัดสินใจ (Decision Tree Regression) 85
3.2.5.3. การถดถอยของป่าแบบสุ่ม (Random Forest Regression) 86
3.2.5.4. การถดถอยลาสโซ (Lasso Regression) 88
3.2.5.5. การถดถอยแบบริดจ์ (Ridge Regression) 89
3.2.5.6. การถดถอยของริดจ์แบบเบส์ (Bayesian Ridge Regression) 90
3.2.5.7. การถดถอยแบบกราเดียนส์บูสติง (Gradient Boosting Regression) 91
3.2.5.8. สโคแคสติกกราเดียนเดสเซนท์ (Stochastic Gradient Decent) 93
3.2.6. การประเมินผลแบบจำลอง (Model Evaluation) 94
3.2.7. การบวนการใช้เทคนิคการรวมโมเดล (Ensemble Algorithms) 95
บทที่ 4 ผลการดำเนินการวิจัย 96
4.1. การพยากรณ์แบบล่วงหน้า 1 วัน 96
4.1.1. Support Vector Machine 96
4.1.2. Decision Tree Regression 97
4.1.3. Random Forest Regression 98
4.1.4. Lasso Regression 98
4.1.5. Ridge Regression 99
4.1.6. Baysian Ridge Regression 100
4.1.7. Gradient Boosting Regression 100
4.1.8. Stochastic Gradient Descent 101
4.2. การพยากรณ์แบบล่วงหน้า 3 วัน 102
4.2.1. Support Vector Machine 102

4.2.2. Decision Tree Regression 102
4.2.3. Random Forest Regression 103
4.2.4. Lasso Regression 104
4.2.5. Ridge Regression 104
4.2.6. Baysian Ridge Regression 105
4.2.7. Gradient Boosting Regression 106
4.2.8. Stochastic Gradient Descent 106
4.3. การพยากรณ์แบบล่วงหน้า 5 วัน 107
4.3.1. Support Vector Machine 107
4.3.2. Decision Tree Regression 108
4.3.3. Random Forest Regression 108
4.3.4. Lasso Regression 109
4.3.5. Ridge Regression 110
4.3.6. Baysian Ridge Regression 110
4.3.7. Gradient Boosting Regression 111
4.3.8. Stochastic Gradient Descent 112
4.4. การพยากรณ์แบบล่วงหน้า 7 วัน 112
4.4.1. Support Vector Machine 112
4.4.2. Decision Tree Regression 113
4.4.3. Random Forest Regression 114
4.4.4. Lasso Regression 114
4.4.5. Ridge Regression 115
4.4.6. Baysian Ridge Regression 116

4.4.7. Gradient Boosting Regression 116
4.4.8. Stochastic Gradient Descent 117
4.5. การรวมโมเดล 121
บทที่ 5 สรุปผลการวิจัย 125
5.1. สรุปผลการวิจัย 125
5.2. อภิปรายผล 126
5.3. ข้อจำกัดและอุปสรรคของงานวิจัย 127
5.4. ข้อเสนอแนะ 127
บรรณานุกรม 129
ประวัติผู้เขียน 132

สารบัญตาราง

หน้า

ตาราง 1 ตารางแสดงโครงสร้างข้อมูลของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ..... 28
ตาราง 2 ตารางแสดงโครงสร้างข้อมูลของดัชนีตลาดหุ้นสหรัฐอเมริกา 29
ตาราง 3 ตารางแสดงโครงสร้างข้อมูลของดัชนีดาวโจนส์..... 30
ตาราง 4 ตารางแสดงโครงสร้างข้อมูลของดัชนีหุ้นเอลโดราโด โกลด์..... 32
ตาราง 5 ตารางแสดงโครงสร้างข้อมูลของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ 33
ตาราง 6 ตารางแสดงโครงสร้างข้อมูลของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ 34
ตาราง 7 ตารางแสดงโครงสร้างข้อมูลของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา..... 36
ตาราง 8 ตารางแสดงโครงสร้างข้อมูลของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา..... 37
ตาราง 9 ตารางแสดงโครงสร้างข้อมูลของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว 38
ตาราง 10 ตารางแสดงโครงสร้างข้อมูลของอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ 40
ตาราง 11 ตารางแสดงโครงสร้างข้อมูลของดัชนีค่าเงินดอลลาร์สหรัฐ 42
ตาราง 12 ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่เงินล่วงหน้า..... 44
ตาราง 13 ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า..... 45
ตาราง 14 ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า..... 48
ตาราง 15 ตารางแสดงการใช้ชื่อแทนของแต่ละชุดข้อมูล..... 49
ตาราง 16 ตารางสรุปการคัดเลือกคุณลักษณะแต่ละเทคนิค..... 80
ตาราง 17 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับซัพพอร์ตเวกเตอร์รีเกรสชัน . 84
ตาราง 18 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบต้นไม้ตัดสินใจ 86
ตาราง 19 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของป่าแบบสุ่ม.. 87
ตาราง 20 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยลาสโซ..... 88

ตาราง 21 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบริดจ์ 89

ตาราง 22 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของริดจ์แบบเบส์ 90

ตาราง 23 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบกราเดียนส์บูสติง 92

ตาราง 24 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับสโตแคสติกกราเดียนเดสเซนท์ 94

ตาราง 25 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 1 วัน 97

ตาราง 26 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 1 วัน 97

ตาราง 27 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 1 วัน 98

ตาราง 28 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 1 วัน 99

ตาราง 29 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 1 วัน 99

ตาราง 30 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 1 วัน 100

ตาราง 31 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 1 วัน 101

ตาราง 32 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 1 วัน 101

ตาราง 33 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 3 วัน 102

ตาราง 34 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 3 วัน 103

ตาราง 35 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 3 วัน 103

ตาราง 36 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 3 วัน 104

ตาราง 37 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 3 วัน 105

ตาราง 38 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 3 วัน 105

ตาราง 39 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 3 วัน 106

ตาราง 40 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 3 วัน 107

ตาราง 41 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 5 วัน 107

ตาราง 42 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 5 วัน 108

ตาราง 43 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 5 วัน 109

ตาราง 44 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 5 วัน 109

ตาราง 45 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 5 วัน 110

ตาราง 46 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 5 วัน 111

ตาราง 47 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 5 วัน 111

ตาราง 48 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 5 วัน 112

ตาราง 49 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 7 วัน 113

ตาราง 50 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 7 วัน 113

ตาราง 51 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 7 วัน 114

ตาราง 52 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 7 วัน 115

ตาราง 53 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 7 วัน 115

ตาราง 54 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 7 วัน 116

ตาราง 55 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 7 วัน 117

ตาราง 56 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 7 วัน 117

ตาราง 57 ตารางแสดงค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R^2) ของแต่ละแบบจำลอง 118

ตาราง 58 ตารางแสดงค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ของแต่ละแบบจำลอง 119

ตาราง 59 ตารางแสดงค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ของแต่ละแบบจำลอง 120

ตาราง 60 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 1 วัน 121

ตาราง 61 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 3 วัน 122
ตาราง 62 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 5 วัน 123
ตาราง 63 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 7 วัน 124

สารบัญรูปภาพ

หน้า

ภาพประกอบ 1 Moving Average Convergence Divergence (MACD) 6

ภาพประกอบ 2 ภาพแสดงโครงสร้างของ Support Vector Regression 10

ภาพประกอบ 3 ภาพแสดงโครงสร้างของ Decision Tree 11

ภาพประกอบ 4 ภาพแสดงโครงสร้างของ Random Forest 12

ภาพประกอบ 5 ภาพแสดงการรวมโมเดลแบบ Bagging 17

ภาพประกอบ 6 ภาพแสดงการรวมโมเดลแบบ Boosting 17

ภาพประกอบ 7 ภาพแสดงการรวมโมเดลแบบ Stacking 18

ภาพประกอบ 8 กระบวนการทำงานของแบบจำลอง (Medel Overview Workflow) 25

ภาพประกอบ 9 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ 50

ภาพประกอบ 10 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีตลาดหุ้นสหรัฐอเมริกา 51

ภาพประกอบ 11 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีดาวโจนส์ 52

ภาพประกอบ 12 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีหุ้นเอลโดราโด โกลด์ 53

ภาพประกอบ 13 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ 54

ภาพประกอบ 14 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ 55

ภาพประกอบ 15 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา 56

ภาพประกอบ 16 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา 57

ภาพประกอบ 17 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว 58

ภาพประกอบ 18 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ 59

ภาพประกอบ 19 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา 60

ภาพประกอบ 20 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่เงินล่วงหน้า 61

ภาพประกอบ 21 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า 62

ภาพประกอบ 22 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า 63

ภาพประกอบ 23 แสดงความสัมพันธ์ (Correlation) ระหว่างคุณลักษณะ 64

ภาพประกอบ 24 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันและการกระจายตัวของราคาปิดของกองทุนทองคำและดัชนีตลาดหุ้นสหรัฐอเมริกา 66

ภาพประกอบ 25 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีดาวโจนส์ 66

ภาพประกอบ 26 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีหุ้นเอลโดราโด โกลด์ 67

ภาพประกอบ 27 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ 67

ภาพประกอบ 28 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ 68

ภาพประกอบ 29 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา 68

ภาพประกอบ 30 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและกองทุนรวมราคาน้ำมันสหรัฐอเมริกา 69

ภาพประกอบ 31 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว 69

ภาพประกอบ 32 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ 70

ภาพประกอบ 33 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา 71

ภาพประกอบ 34 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและสัญญาซื้อขายแร่เงินล่วงหน้า 71

ภาพประกอบ 35 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและสัญญาซื้อขายแร่แพลทินัมล่วงหน้า 72

ภาพประกอบ 36 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า 72

ภาพประกอบ 37 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณ MACD ของราคาปิดกองทุนทองคำ 73

ภาพประกอบ 38 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณค่าเฉลี่ยเคลื่อนที่แบบง่าย 74

ภาพประกอบ 39 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณดัชนีความแข็งแกร่งสัมพัทธ์ 75

ภาพประกอบ 40 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่ 75

ภาพประกอบ 41 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณผลต่างของราคาปิดและราคาเปิด, ราคาสูงสุดและราคาต่ำสุด 76

ภาพประกอบ 42 แสดงข้อมูลตัวอย่างหลังจากผ่านการปรับขนาดของข้อมูล 77

ภาพประกอบ 43 แสดงข้อมูลตัวอย่างหลังจากผ่านการปรับขนาดของข้อมูล (ต่อ) 77

ภาพประกอบ 44 แสดงรูปร่างของข้อมูลหลังจากทำการแบ่งชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ 77

ภาพประกอบ 45 แสดงคุณลักษณะที่ถูกคัดเลือกจากการถดถอยลาสโซของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ..... 78

ภาพประกอบ 46 แสดงคุณลักษณะที่ถูกคัดเลือกจากการถดถอยของป่าแบบสุ่มของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ..... 79

ภาพประกอบ 47 แสดงคุณลักษณะที่ถูกคัดเลือกจากการค้นหาแบบขั้นตอนเชิงพันธุกรรมของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ..... 80

ภาพประกอบ 48 แสดงพารามิเตอร์ของ GridSearchCV สำหรับซัพพอร์ตเวกเตอร์รีเกรสชัน 84

ภาพประกอบ 49 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบต้นไม้ตัดสินใ . 85

ภาพประกอบ 50 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยของป่าแบบสุ่ม 87

ภาพประกอบ 51 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยลาสโซ..... 88

ภาพประกอบ 52 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบริดจ์ 89

ภาพประกอบ 53 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยของริดจ์แบบเบส... 90

ภาพประกอบ 54 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบกราเดียนส์บูสติง 92

ภาพประกอบ 55 แสดงพารามิเตอร์ของ GridSearchCV สำหรับสโตแคสติกกราเดียนเดสเซนท์ 93

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เศรษฐกิจโลกในปัจจุบันมีการเปลี่ยนแปลงไปอย่างรวดเร็ว ส่งผลให้การลงทุนกลายเป็นส่วนสำคัญในการจัดการทางการเงินของประชากรทั่วโลก อาทิ การลงทุนในหุ้น ตราสารหนี้ พันธบัตร อสังหาริมทรัพย์ สกุลเงินดิจิทัล ไปจนกระทั่งการลงทุนในรูปแบบของทองคำที่เป็นทรัพย์สินที่มีความสำคัญต่อระบบเศรษฐกิจและการเงินมาอย่างยาวนาน ซึ่งทั่วโลกยอมรับทองคำในฐานะที่เรียกว่าสินทรัพย์ปลอดภัย (Safe Haven) จากการวิจัยการป้องกันความเสี่ยงจากการถล่มของตลาดหุ้นและความไม่แน่นอนทองเศรษฐกิจ โดยใช้ช่วงการระบาดของโรคโควิด-19 เป็นช่วงเวลาที่ตรวจสอบบทบาทของราคาทองคำ (Is gold a hedge or a safe-haven asset in the COVID–19 crisis?) และเปรียบเทียบผลตอบแทนที่ได้รับจากทองคำและดัชนีหุ้น ซึ่งประกอบไปด้วย S&P 500, Euro Stoxx 50, Nikkei 225 และดัชนี China FTSE A50 นั้น พบว่าดัชนีหุ้นมีค่าเป็นลบ กล่าวคือ ดัชนีหุ้นถูกลดมูลค่าลงในช่วงการระบาดของโรคโควิด-19 ซึ่งในช่วงเวลาดังกล่าวราคาทองคำก็มีค่าเป็นลบเช่นเดียวกัน แต่เมื่อเปรียบเทียบมูลค่าการสูญเสียแล้วพบว่าดัชนีหุ้นมีการสูญเสียมากกว่าราคาทองคำ ซึ่งเป็นตัวบ่งชี้ที่ทำให้เห็นว่าทองคำเป็นสินทรัพย์ปลอดภัย (Akhtaruzzaman et al., 2021)

อย่างไรก็ตามทองคำที่กล่าวว่าเป็นสินทรัพย์ปลอดภัย ก็ยังคงมีความผันผวนของราคาเช่นเดียวกับสินทรัพย์รูปแบบอื่น จากการวิจัยการใช้ทองคำเพื่อเป็นเครื่องมือป้องกันความเสี่ยงหรือสินทรัพย์ปลอดภัยที่มีประสิทธิภาพเมื่อเทียบกับหุ้น ตราแลกเปลี่ยนรูปี-ดอลล่า และน้ำมันดิบของประเทศอินเดีย (Is gold an effective hedge and/or safe haven instrument against stocks, rupee–dollar rate and crude: an empirical investigation from India) พบว่าในช่วงเวลาที่เกิดวิกฤตการณ์การเงินในแต่ละช่วง ทองคำเป็นสินทรัพย์ปลอดภัยได้ที่ดีเมื่อเปรียบเทียบกับดัชนีหุ้น Nifty และอัตราแลกเลี่ยนรูปี-ดอลล่า แต่ไม่สามารถป้องกันความเสี่ยงจากราคาน้ำมันดิบได้ (Shahani & Bansal, 2021) ซึ่งชี้ให้เห็นว่าทองคำไม่ใช่เครื่องมือป้องกันความเสี่ยงหรือสินทรัพย์ปลอดภัยได้ในทุกกรณี

เมื่อทองคำที่เป็นที่เชื่อถือกันว่าเป็นสินทรัพย์ปลอดภัยหรือสินทรัพย์ป้องกันความเสี่ยงไม่สามารถเป็นสินทรัพย์ป้องกันความเสี่ยงได้ในทุกกรณี จึงก่อให้เกิดการศึกษาเกี่ยวกับการวิเคราะห์ทางเศรษฐมิติ (Econometrics) เพื่อใช้ในการทำนายราคาทองคำ โดยใช้ความสัมพันธ์ระหว่างทองคำและสินทรัพย์รูปแบบอื่น และนำมาสู่การพัฒนาการเรียนรู้ของเครื่อง (Machine

Learning) เพื่อพัฒนาการทำนายราคาทองคำให้แม่นยำขึ้น และวิเคราะห์ราคาทองคำจากราคาของสินทรัพย์อื่นที่มีนัยสำคัญกับทองคำได้มากยิ่งขึ้น ทั้งจากราคาน้ำมันดิบ ดัชนีตลาดหุ้น กองทุน และพันธบัตร ซึ่งล้วนแล้วแต่มีนัยสำคัญกับราคาทองทั้งสิ้น

จากการวิจัยก่อนหน้ามุ่งเน้นการทำนายราคาทองคำในลักษณะเชิงเส้น กล่าวคือเป็นการทำนายที่ใช้ราคาทองวันก่อนหน้าในการทำนายราคาทองเพียงเท่านั้น หรือในลักษณะการทำนายราคาทองไม่เชิงเส้น คือการใช้ราคาสินทรัพย์อื่นในการทำนายราคาทอง แต่ยังคงเลือกเฉพาะราคาปิดปรับปรุงของสินทรัพย์อื่นมาทำนายเพียงเท่านั้น ซึ่งในด้านเศรษฐมิติยังมีตัวช่วยทางสถิติอีกมากมายที่สามารถช่วยในการทำนายราคาทอง เช่น Moving Average Convergence-Divergence, ดัชนีความแข็งแกร่งสัมพัทธ์ (Relative Strength Index) และค่าเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average) เป็นต้น ซึ่งสามารถนำมาเป็นคุณลักษณะเพิ่มเติมเพื่อให้การเรียนรู้ของเครื่อง (Machine Learning) สามารถเรียนรู้ได้อย่างมีประสิทธิภาพมายิ่งขึ้น แต่การเรียนรู้คุณลักษณะที่มากเดินไปของการเรียนรู้ของเครื่องก็ไม่ได้เป็นผลดีมากนัก ซึ่งทำให้ประสิทธิภาพการเรียนรู้ลดลง มีการทำนายหรือพยากรณ์ที่ไม่แม่นยำ จึงใช้เทคนิคการคัดเลือกคุณลักษณะ (Feature Selection) มาปรับใช้เพื่อคัดเลือกคุณลักษณะที่มีความจำเป็นในการเรียนรู้ของเครื่องเพียงเท่านั้น และเมื่อได้โมเดล และการเรียนรู้ของเครื่องที่มีประสิทธิภาพที่สุดแล้วจะนำเข้าสู้เทคนิคการเรียนรู้ของเครื่องแบบกลุ่ม (Ensemble Machine Learning) เพื่อให้การทำนายและการพยากรณ์มีความแม่นยำมากยิ่งขึ้น

1.2 วัตถุประสงค์ของงานวิจัย

การวิจัยครั้งนี้ผู้วิจัยได้ตั้งจุดมุ่งหมายไว้ดังนี้

1. เพื่อศึกษาเทคนิคการเรียนรู้ของเครื่องและเปรียบเทียบประสิทธิภาพในแต่ละอัลกอริทึม (Algorithms) โดยแบบจำลอง (Model) ที่ใช้คือ ซัพพอร์ตเวกเตอร์รีเกสชัน (Support Vector Regression), การถดถอยแบบต้นไม้ตัดสินใจ (Decision Tree Regression), การถดถอยของป่าแบบสุ่ม (Random Forest Regression), การถดถอยลาสโซ (Lasso Regression), การถดถอยแบบริดจ์ (Ridge Regression), การถดถอยของริดจ์แบบเบส์ (Bayesian Ridge Regression), การถดถอยแบบกราเดียนส์บูสติง (Gradient Boosting Regression) และสโคแคสติกกราเดียนเดสเซนท์ (Stochastic Gradient Decent)

2. เพื่อศึกษาเทคนิคการออกแบบคุณลักษณะ (Feature Engineering) และเปรียบเทียบประสิทธิภาพในแต่ละอัลกอริทึม โดยเทคนิคที่ใช้คือ การถดถอยลาสโซ, การถดถอยของป่าแบบสุ่ม และการค้นหาแบบขั้นตอนเชิงพันธุกรรม (Genetic Algorithms)

3. เพื่อศึกษาเทคนิคการรวมโมเดล และเปรียบเทียบประสิทธิภาพในแต่ละอัลกอริทึม โดยใช้เทคนิค Voting และ Stacking

4. เพื่อศึกษาการเปรียบเทียบการทำนายราคาปิดตลาดของกองทุนทองคำกับราคาปิดตลาดของกองทุนทองคำจริงที่ผ่านมาแล้วในอดีตโดยแบ่งการเปรียบเทียบแบบการพยากรณ์ 1, 3, 5 และ 7 วัน ในแบบจำลองที่มีประสิทธิภาพ

1.3 ขอบเขตงานวิจัย

1. การวิจัยใช้ข้อมูลกองทุนและตลาดหุ้นของประเทศสหรัฐอเมริกา จาก National Association of Securities Dealers Automated Quotations และ investing.com

2. นำข้อมูลมาวิเคราะห์ เพื่อทำนายราคากองทุนทองคำ จากคุณลักษณะที่มีความสัมพันธ์ใกล้ชิดกับราคาปิดตลาดของกองทุนทองคำ

3. ใช้เทคนิคการเลือกคุณลักษณะ ที่มีความใกล้ชิดกับราคากองทุนทองคำ โดยใช้ 3 เทคนิค ได้แก่ การถดถอยลาสโซ่, การถดถอยของป่าแบบสุ่ม และการค้นหาขั้นตอนเชิงพันธุกรรม

4. ใช้เทคนิคการเรียนรู้ของเครื่อง ในการทำนายราคากองทุนทองคำ โดยใช้ 8 เทคนิค ได้แก่ ซัพพอร์ตเวกเตอร์รีเกสชัน, การถดถอยแบบต้นไม้ตัดสินใจ, การถดถอยของป่าแบบสุ่ม, การถดถอยลาสโซ, การถดถอยแบบริดจ์, การถดถอยของริดจ์แบบเบส์, การถดถอยแบบกราเดียนส์บูสติง และสโคแคสติกกราเดียนเดสเซนท์

5. ใช้เทคนิคการรวมโมเดล โดยเลือกจากการเรียนรู้ของเครื่องที่มีความแม่นยำที่สุด 1, 2 และ 3 อันดับ

6. สร้างแบบจำลองเพื่อทำนายราคากองทุนทองคำล่วงหน้า 1, 3, 5 และ 7 วัน

7.วัดประสิทธิภาพด้วยค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error), ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R-Square) และค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (Mean Absolute Percentage Error)

1.4 สมมติฐานการวิจัย

1. คุณลักษณะ ที่เหมาะสมสำหรับการทำนายราคาปิดของกองทุนทองคำมีบางส่วนที่เหมาะสม จึงใช้เทคนิคการเลือกคุณลักษณะ เพื่อเลือกคุณลักษณะที่เหมาะสม โดยการนำคุณลักษณะที่ถูกเลือกแล้ว ผ่านการเรียนรู้ของเครื่อง และเปรียบเทียบประสิทธิภาพของแต่ละเทคนิคที่ผ่านการเรียนรู้ เพื่อค้นหาเทคนิคที่ดีที่สุด 2, 3 และ 4 อันดับ เพื่อใช้เทคนิคการรวมโมเดล

2. จากแบบจำลองการเรียนรู้ของเครื่องที่นำมาทดลองมีเพียงบางแบบจำลองที่มีความเหมาะสมกับการทำนายราคาปิดตลาดของกองทุนทองคำยิ่งกว่าแบบจำลองอื่น โดยการทำการเปรียบเทียบประสิทธิภาพของแต่ละแบบจำลอง เพื่อค้นหาแบบจำลองที่ดีที่สุด 2, 3 และ 4 อันดับ เพื่อใช้เทคนิคการรวมโมเดล

3. เทคนิคการรวมโมเดล ใช้เทคนิคการรวมการเรียนรู้ของเครื่องเพื่อทำนายราคาปิดตลาดของกองทุนทองคำ เพื่อเปรียบเทียบการผสมการเรียนรู้ของเครื่องที่ดีที่สุดใน 2, 3 และ 4 อันดับ โดยทั้งผ่านเทคนิคการเลือกคุณลักษณะ และไม่ผ่านเทคนิคการเลือกคุณลักษณะ โดยการเปรียบเทียบเทคนิคการรวมโมเดลที่ดีที่สุดของการทำนายราคาปิดของกองทุนทองคำ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำแบบจำลองที่ได้ไปช่วยทำนายราคาปิดตลาดของกองทุนทองคำ เพื่อใช้ในการประกอบการตัดสินใจการลงทุนในทองคำ

2. สามารถนำแบบจำลองที่ได้ไปช่วยทำนายราคาปิดของกองทุนทองคำ เพื่อใช้ในการประเมินความเสี่ยงเพื่อปรับอัตราส่วนการลงทุน

3. สามารถนำแบบจำลองที่ได้ไปช่วยทำนายราคาปิดของกองทุนทองคำ เพื่อใช้ในการควบคุมการจัดการเงินเฟ้อ

4. สามารถนำแบบจำลองที่ได้ไปช่วยทำนายราคาปิดของกองทุนทองคำ เพื่อใช้ในการปรับกลยุทธ์ในตลาดการเงิน

บทที่ 2
ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ผู้วิจัยได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้องกับการเปรียบเทียบแบบจำลองในการทำนายราคาปิดของกองทุนทองคำและได้มีการนำเสนอตามหัวข้อต่อไปนี้

2.1. ทฤษฎีที่เกี่ยวข้อง

2.1.1. Moving Average Convergence-Divergence

2.1.2. ดัชนีความแข็งแกร่งสัมพัทธ์ (Relative Strength Index)

2.1.3. ค่าเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average)

2.1.4. เทคนิคการเลือกคุณลักษณะ (Feature Selection)

2.1.4.1. การค้นหาแบบขั้นตอนเชิงพันธุกรรม (Genetic Algorithms)

2.1.5. แบบจำลองการเรียนรู้ของเครื่อง (Machine Learning)

2.1.5.1. ซัพพอร์ตเวกเตอร์รีเกสชัน (Support Vector Regression)

2.1.5.2. การถดถอยแบบต้นไม้ตัดสินใจ (Decision Tree Regressor)

2.1.5.3. การถดถอยของป่าแบบสุ่ม (Random Forest Regression)

2.1.5.4. การถดถอยลาสโซ (Lasso Regressor)

2.1.5.5. การถดถอยแบบริดจ์ (Ridge Regression)

2.1.5.6. การถดถอยของริดจ์แบบเบส์ (Bayesian Ridge Regression)

2.1.5.7. การถดถอยแบบกราเดียนท์บูสติง (Gradient Boosting Regressor)

2.1.5.8. สโตแคสติกแกรเดียนท์เดสเซนท์ (Stochastic Gradient Decent)

2.1.6. เทคนิคการรวมโมเดล (Ensemble Method)

2.1.7. เครื่องมือการแปรผล

2.1.7.1. ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error, RMSE)

2.1.7.2. ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (Coefficient of Multiple Determination, R-Squared)

2.1.7.3. ค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (Mean Absolute Percentage Error, MAPE)

2.2. งานวิจัยที่เกี่ยวข้อง

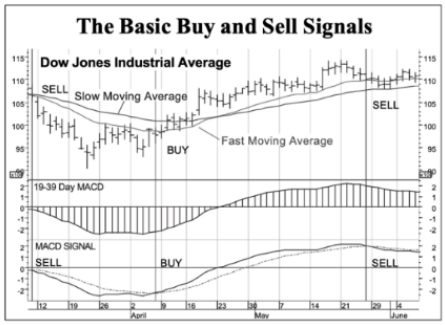
2.1. ทฤษฎีที่เกี่ยวข้อง

2.1.1. Moving Average Convergence-Divergence (MACD)

Moving Average Convergence Divergence (MACD) เป็นเครื่องมือที่ถูกพัฒนาโดย Gerald Appel ในช่วงปี 1970 ซึ่งใช้กันอย่างแพร่หลายในการวิเคราะห์หุ้น ฟอเร็กซ์ และตลาดการเงินอื่น ๆ โดยใช้เทคนิคที่ใช้ตัวชี้วัดที่รวมค่าเฉลี่ยเคลื่อนที่สองตัวเพื่อแสดง Momentum ของการซื้อขาย ได้แก่ MACD Line และ Signal Line โดย MACD Line เกิดจากค่าของผลต่างระหว่างค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โพเนนเชียล (Exponential Moving Average, EMA) แบบระยะสั้น และแบบระยะยาว โดยทั่วไปจะกำหนดให้แบบระยะสั้นมีจำนวน 12 วัน และแบบระยะยาวมีจำนวน 26 วัน และ Signal Line เกิดจากค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โพเนนเชียลของ MACD โดยทั่วไปจะกำหนดค่าเป็น 9 วัน ซึ่ง Signal Line ใช้เพื่อยืนยันสัญญาณซื้อขายที่เกิดขึ้นจาก MACD Line

การทำงานของ MACD จะแบ่งออกเป็น 2 สัญญาณ คือ สัญญาณซื้อ (Bullish Signal) เกิดขึ้นเมื่อ MACD Line ตัดขึ้นผ่าน Signal Line ซึ่งแสดงถึงกำลังซื้อที่เพิ่มขึ้น เป็นสัญญาณที่บ่งบอกถึงโอกาสในการซื้อ และสัญญาณขาย (Bearish Signal) เกิดขึ้นเมื่อ MACD Line ตัดลงผ่าน Signal Line ซึ่งแสดงถึงกำลังขายที่เพิ่มขึ้น เป็นสัญญาณที่บ่งบอกถึงการขาย

นอกจากนี้ เนื่องจาก MACD แสดงความแตกต่างระหว่างค่าเฉลี่ยเคลื่อนที่แบบเร็วและช้า ผลลัพธ์จะผันผวนรอบค่า 0 ทั้งด้านบวกและด้านลบ พื้นที่ที่อยู่เหนือค่า 0 แสดงถึงพื้นที่ที่มีการซื้อมากเกินไป (Overbought Area) และพื้นที่ที่อยู่ใต้ค่า 0 แสดงถึงพื้นที่ที่มีการขายมากเกินไป ดังนั้น สัญญาณการเข้าสู่ตำแหน่งซื้อที่แข็งแกร่งคือการเกิดขึ้นในพื้นที่ที่มีการขายมากเกินไป ในขณะที่สัญญาณการเข้าสู่ตำแหน่งขายที่แข็งแกร่งคือการเกิดขึ้นในพื้นที่ที่มีการซื้อมากเกินไป (Cardoso, 2019)



ภาพประกอบ 1 Moving Average Convergence Divergence (MACD)

ที่มา: (Appel, 2005)

$$MACD = EMA_f - EMA_s \quad (2.1)$$

โดย EMA_f หมายถึง ค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โปแนนเชียลแบบระยะสั้น

EMA_s หมายถึง ค่าเฉลี่ยเคลื่อนที่แบบอ็กซ์โปแนนเชียลแบบระยะยาว

โดยค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โพเนนเชียลในแต่ละช่วงเวลา สามารถคำนวณได้โดยการใช้ Smoothing Constant (α) เพื่อหาความแตกต่างระหว่างราคาปิดในช่วงเวลานั้น

$$\alpha = \frac{2}{t+1} \quad (2.2)$$

โดย t หมายถึง จำนวนข้อมูลที่ใช้ในการคำนวณ

และเมื่อได้ Smoothing Constant (α) จะสามารถหาเคลื่อนที่แบบเอ็กซ์โพเนนเชียลในแต่ละช่วงเวลาได้โดยการหาผลต่างระหว่างราคาปิดในช่วงเวลาที่ต้องการ กับค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โพเนนเชียลในช่วงเวลาก่อนหน้า และนำ Smoothing Constant (α) รวมกับค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โพเนนเชียลช่วงเวลาก่อนหน้า

$$EMA_n = \left(CP_n - EMA_{n-1}\right) \times \alpha + EMA_{n-1} \quad (2.3)$$

โดย CP_n หมายถึง ผลต่างระหว่างราคาปิดในช่วงเวลาที่ต้องการ

EMA_{n-1} หมายถึง ค่าเฉลี่ยเคลื่อนที่แบบเอ็กซ์โปแนนเชียลในช่วงเวลาก่อนหน้า

α หมายถึง Smoothing Constant

2.1.2. **ดัชนีความแข็งแกร่งสัมพัทธ์** (Relative Strength Index, RSI)

ดัชนีความแข็งแกร่งสัมพัทธ์ (Relative Strength Index : RSI) เป็นหนึ่งในเครื่องมือการวิเคราะห์ทางเทคนิคที่ใช้ในการวัดความแรง (Strength) และความเร็ว (Momentum) ของการเปลี่ยนแปลงราคา RSI ช่วยให้ผู้ใช้งานทราบว่าตลาดอยู่ในสภาวะซื้อเกิน (Overbought) หรือ ขายเกิน (Oversold) ซึ่ง RSI ถูกพัฒนาขึ้นโดย J. Welles Wilder ในปี 1978

การคำนวณ RSI จะทำการคำนวณโดยการเปรียบเทียบขนาดของการเปลี่ยนแปลงราคาที่เป็นบวก (Gain) กับขนาดของการเปลี่ยนแปลงราคาที่เป็นลบ (Loss) ในช่วงเวลาที่กำหนด เช่น 14 วัน ซึ่งเป็นช่วงเวลาที่นิยมใช้มากที่สุด

$$RSI = 100\frac{100}{1+RS} \tag{2.4}$$

โดย RS หมายถึง Relative Strength

$$RS = \frac{\text{Average Gain}}{\text{Average Loss}} \tag{2.5}$$

การคำนวณค่าเฉลี่ยของกำไร (Average Gain) คำนวณการเปลี่ยนแปลงเฉลี่ยของราคาที่เป็นบวกในช่วงเวลาที่กำหนด และค่าเฉลี่ยของการขาดทุน (Average Loss) คำนวณจากการเปลี่ยนแปลงเฉลี่ยของราคาที่เป็นลบในช่วงเวลาที่กำหนด

การวิเคราะห์ค่า RSI แบ่งออกเป็น 2 กรณี คือ กรณี RSI > 70 หมายความว่า ตลาดอาจอยู่ในภาวะ ซื้อเกิน (Overbought) ซึ่งมีโอกาสที่ราคาอาจเริ่มปรับตัวลง เป็นสัญญาณเตือนว่าราคาอาจมีแนวโน้มกลับตัว และRSI < 30 หมายความว่า ตลาดอาจอยู่ในภาวะขายเกิน (Oversold) ซึ่งมีโอกาสที่ราคาอาจเริ่มปรับตัวขึ้น เป็นสัญญาณเตือนว่าราคาอาจมีแนวโน้มกลับตัวขึ้น

2.1.3. **ค่าเฉลี่ยเคลื่อนที่แบบง่าย** (Simple Moving Average)

ค่าเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average, SMA) คือค่าเฉลี่ยเคลื่อนที่แบบธรรมดา ซึ่งเป็นหนึ่งในเครื่องมือการวิเคราะห์ทางเทคนิคที่ใช้กันอย่างแพร่หลายเพื่อหาค่าเฉลี่ยของราคาหุ้นหรือสินทรัพย์ในช่วงเวลาหนึ่ง SMA เป็นวิธีการวัดแนวโน้มของราคาโดยการคำนวณค่าเฉลี่ยของราคาในช่วงที่กำหนด เช่น 10 วัน, 50 วัน หรือ 200 วัน

การคำนวณ SMA จะคำนวณโดยการนำราคาปิดของช่วงเวลาที่กำหนด (เช่น n วัน) มารวมกัน แล้วหารด้วยจำนวนวัน

$$SMA = \frac{\sum_{i=1}^{n} P_i}{n} \tag{2.6}$$

โดย P_i หมายถึง ราคาปิดในแต่ละวัน

n หมายถึง จำนวนวันหรือเวลาที่ใช้ในการคำนวณค่าเฉลี่ย

2.1.4. **เทคนิคการเลือกคุณลักษณะ** (Feature Selection)

2.1.4.1. **การค้นหาแบบขั้นตอนวิธีเชิงพันธุกรรม** (Genetic Algorithms)

การเลือกคุณลักษณะแบบขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithms) เป็นเทคนิคการคัดเลือกคุณลักษณะที่ได้รับแรงบันดาลใจจากกระบวนการคัดเลือกตามธรรมชาติ (Natural Selection) ในชีววิทยา โดยมีแนวคิดนี้ในการค้นหาชุดคุณลักษณะที่ดีที่สุดเพื่อนำไปสร้างโมเดลที่แม่นยำ

การทำงานของ Genetic Algorithm จะทำงานโดยการจำลองกระบวนการวิวัฒนาการของสิ่งมีชีวิต ซึ่งจะคัดเลือกเฉพาะโครโมโซม (Chromosomes) หรือชุดของคุณลักษณะ (Feature) ที่มีคุณภาพสูงไปสืบพันธุ์ สร้างรุ่นใหม่ โดยปรับปรุงประสิทธิภาพของการเลือกคุณลักษณะไปเรื่อย ๆ จนกว่าจะได้ชุดคุณลักษณะที่เหมาะสมที่สุด

กระบวนการของ Genetic Algorithm ประกอบไปด้วยขั้นตอนต่อไปนี้

1. Initialization เป็นขั้นตอนการสร้างประชากรเริ่มต้นของคุณลักษณะเป็นไปได้แบบสุ่ม

2. Selection เป็นขั้นตอนการเลือกชุดคุณลักษณะที่มีคะแนน (Fitness) สูงที่สุดเพื่อไปเป็นต้นแบบในขั้นต่อไป

3. Crossover เป็นขั้นตอนการผสมชุดคุณลักษณะสองชุดเข้าด้วยกันเพื่อสร้างลูกหลาน (Offspring) ซึ่งเป็นการสร้างชุดคุณลักษณะใหม่

4. Mutation เป็นขั้นตอนการเปลี่ยนคุณลักษณะในบางจุดแบบสุ่มเพื่อหลีกเลี่ยงการติดอยู่ใน Local Optima

5. Evaluation เป็นขั้นตอนการประเมินประสิทธิภาพของแต่ละชุดคุณลักษณะและคัดเลือกชุดที่มีคะแนนสูงที่สุด

การใช้งาน Genetic Algorithms จะใช้งานผ่านฟังก์ชัน GAFeatureSelectionCV ซึ่งเป็นฟังก์ชันใน Scikit-learn เป็นการใช้ Genetic Algorithm เพื่อเลือกคุณลักษณะ ในกระบวนการ Cross-Validation โดยจะประเมินความแม่นยำของโมเดลที่ใช้ชุดคุณลักษณะแตกต่างกันหลายชุด โดยการทดสอบแต่ละชุดผ่านขั้นตอนการฝึกและทดสอบโมเดลซ้ำ ๆ ในหลาย ๆ ชุดรูปแบบข้อมูลเพื่อหาค่าเฉลี่ยความแม่นยำ ซึ่งช่วยให้ GAFeatureSelectionCV ค้นหาชุดคุณลักษณะที่เหมาะสมที่สุดกับโมเดล

2.1.5. แบบจำลองการเรียนรู้ของเครื่อง (Machine Learning)

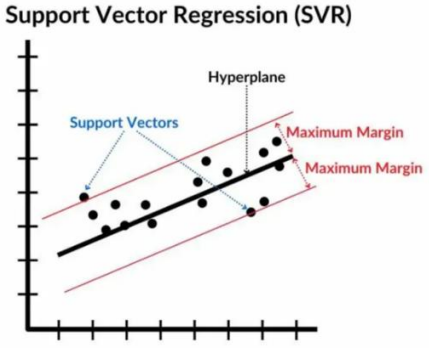
2.1.5.1. ซัพพอร์ตเวกเตอร์รีเกสชัน (Support Vector Regression)

Support Vector Regression (SVR) เป็นหนึ่งในรูปแบบของ Support Vector Machine (SVM) ที่นำมาใช้กับปัญหาการถดถอย (Regression) เพื่อทำนายค่าตัวแปรเป้าหมายที่เป็นตัวเลขต่อเนื่อง โดย SVR ใช้แนวคิดในการวาดเส้นโค้งหรือเส้นตรงที่ดีที่สุด (Hyperplane) โดยเน้นการรักษาความแม่นยำสูงสุดและควบคุมความผิดพลาดให้อยู่ในช่วงที่ยอมรับได้ เพื่อสร้างเส้นทำนายที่สามารถคาดการณ์ได้แม่นยำที่สุด

การทำงานของ SVR เริ่มจากการสร้างเส้น (Hyperplane) ที่อยู่ในรูปของเส้นตรงหรือโค้ง ซึ่งพยายามที่จะสร้างการทำนายใกล้เคียงกับค่าจริงของข้อมูลโดยมีค่าความผิดพลาดให้น้อยที่สุด ซึ่งเป้าหมายของ SVR คือการทำให้เส้นนี้มีค่า Margin กว้างที่สุดเท่าที่จะเป็นไปได้ โดย Margin คือตำแหน่งของจุดที่อยู่ห่างจาก Hyperplane มากที่สุดในช่วงความคลาดเคลื่อนที่ยอมรับได้ ต่อมา SVR จะใช้ค่าพารามิเตอร์ Epsilon (ε) กำหนดความคลาดเคลื่อนที่ยอมรับได้หรือช่วงความผิดพลาดที่สามารถยอมรับได้ เช่น การยอมให้จุดข้อมูลอยู่ห่างจาก Hyperplane ไม่เกินค่า Epsilon

โดยจุดข้อมูลที่อยู่นอกขอบเขตนี้จะถูกพิจารณาเป็นค่าคลาดเคลื่อนและจะถูกลงโทษ (Penalty) และจุดข้อมูลที่อยู่ใกล้หรืออยู่นอกขอบเขต Epsilon จะถูกเรียกว่า Support Vectors ซึ่งจะมีบทบาทสำคัญในการสร้างและกำหนดตำแหน่งของ Hyperplane ซึ่ง Support Vectors เป็นจุดข้อมูลที่สำคัญ เพราะการเปลี่ยนแปลงของมันจะส่งผลโดยตรงต่อเส้นทำนาย

เป้าหมายของ SVR คือการหาตำแหน่งของ Hyperplane ที่ทำให้การคาดการณ์แม่นยำที่สุด โดยคำนึงถึงความคลาดเคลื่อนที่อยู่นอกช่วง Epsilon ซึ่งสามารถควบคุมได้ผ่านค่าพารามิเตอร์ C โดย Loss Function ของ SVR ใช้ค่า Epsilon-Insensitive Loss โดยยอมรับค่าผิดพลาดที่อยู่ในขอบเขต Epsilon และจะคำนวณค่า Loss เฉพาะค่าผิดพลาดที่อยู่นอกช่วงนั้น

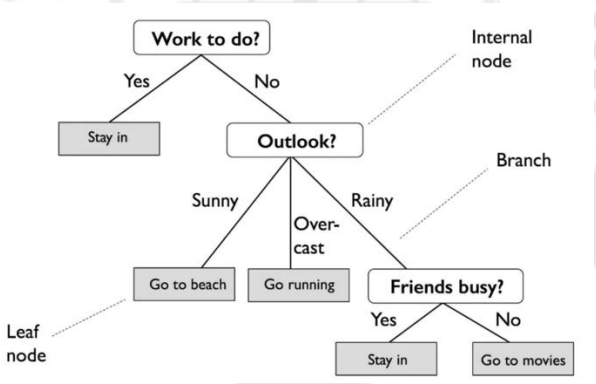


ภาพประกอบ 2 ภาพแสดงโครงสร้างของ Support Vector Regression

2.1.5.2. **การถดถอยแบบต้นไม่ตัดสินใจ** (Decision Tree Regression)

การถดถอยแบบต้นไม้ตัดสินใจ (Decision Tree Regression) เป็นหนึ่งในอัลกอริทึมการเรียนรู้แบบมีผู้สอน (Supervised Learning) ที่ใช้สำหรับปัญหาการถดถอย (Regression) โดยมีแนวคิดการสร้างโครงสร้างต้นไม้ (Tree Structure) เพื่อตัดสินใจแบ่งข้อมูลเป็นกลุ่มย่อยและสร้างการคาดการณ์ตามค่าเฉลี่ยของข้อมูลในแต่ละกลุ่มย่อยเหล่านั้น โครงสร้างการตัดสินใจนี้ถูกสร้างขึ้นมาโดยการเลือกคุณลักษณะ ที่ช่วยแบ่งข้อมูลให้มีความแม่นยำมากที่สุด

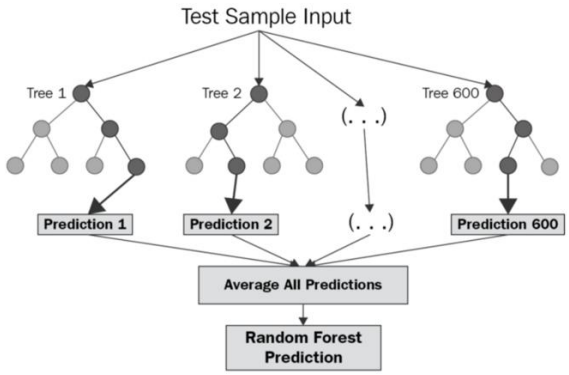
การทำงานของ Decision Tree จะแบ่งข้อมูลออกเป็นสองกลุ่มซ้ำ ๆ ตามค่าเกณฑ์ของคุณลักษณะที่เลือก ซึ่งค่าเกณฑ์นี้เป็นค่าที่ทำให้ความแปรปรวน (Variance) ของข้อมูลในกลุ่มย่อยลดลงมากที่สุด โดยกระบวนการนี้เรียกว่า Recursive Binary Splitting โดยกระบวนการแบ่งนี้จะดำเนินไปจนกว่าจะถึงเกณฑ์ที่กำหนด เช่น ความลึกของต้นไม้ หรือจำนวนข้อมูลในแต่ละกลุ่มย่อยเล็กสุด เมื่อต้นไม้ถูกสร้างขึ้นแล้ว การคาดการณ์สำหรับข้อมูลใหม่จะถูกกำหนดโดยการเดินทางตามโหนดตามค่าของคุณลักษณะไปจนถึงโหนดใบ (Leaf Node) ที่อยู่ล่างสุดของต้นไม้ โดยค่าผลลัพธ์ของข้อมูลจะถูกคาดการณ์เป็นค่าเฉลี่ยของข้อมูลทั้งหมดที่อยู่ในโหนดใบนี้



ภาพประกอบ 3 ภาพแสดงโครงสร้างของ Decision Tree

2.1.5.3. **การถดถอยของป่าแบบสุ่ม** (Random Forest Regression)

การถดถอยของป่าแบบสุ่ม (Random Forest Regression) เป็นเทคนิคการเรียนรู้ของเครื่องที่ใช้สำหรับการทำนายค่าตัวแปรเชิงต่อเนื่อง (Continuous variable) โดยอิงจากแนวคิดของการรวมกันที่ใช้หลาย ๆ ต้นไม้ตัดสินใจ (Decision Trees) ในการสร้างการคาดการณ์ซึ่งจะช่วยเพิ่มความแม่นยำและลดความแปรปรวนของโมเดลได้ดีกว่าใช้ต้นไม้เพียงต้นเดียว



ภาพประกอบ 4 ภาพแสดงโครงสร้างของ Random Forest

2.1.5.4. **การถดถอนลาสโซ (Lasso Regression)**

Lasso Regression (Least Absolute Shrinkage and Selection Operator) เป็นเทคนิคการถดถอย ที่ใช้ในการวิเคราะห์ข้อมูลเพื่อทำนายค่าตัวแปรเชิงต่อเนื่อง โดยใช้วิธีการควบคุมการซับซ้อนของโมเดลเพื่อป้องกันการ Overfitting ผ่านการเพิ่มพูน (Regularization) ที่มีผลกระทบต่อการเลือกคุณลักษณะ (Feature Selection) และค่าพารามิเตอร์ในโมเดล Lasso ใช้เทคนิคที่เรียกว่า L1 Regularization ซึ่งหมายถึงการเพิ่มค่าความผิดพลาด (Loss Function) ที่เกิดจากค่าพารามิเตอร์ของโมเดลด้วยผลรวมของค่าแอลเบลโท (Absolute Values) ของพารามิเตอร์

Lasso Regression มีวัตถุประสงค์เพื่อลดค่าความผิดพลาดโดยใช้การรวมค่าความผิดพลาดของโมเดล (Residual Sum of Squares. RSS) กับค่าพารามิเตอร์ที่ถูกควบคุมด้วยค่าพารามิเตอร์ λ (Lambda) หรือ α (Alpha)

$$Objective\ Function\ =\ RSS + \lambda \sum_{j=1}^{p} \left|\beta_j\right| \tag{2.7}$$

โดย β_j หมายถึง พารามิเตอร์ที่ถูกฝึกฝน

λ ค่าพารามิเตอร์ที่ใช้กำหนดขนาดของการเพิ่มพูน ถ้า λ เป็น 0 จะเท่ากับการทำ Linear Regression ธรรมดา

$$\mathrm{RSS} = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 \tag{2.8}$$

การเพิ่มค่าของ λ จะทำให้ค่าแอลเบลโท (Absolute Value) ของพารามิเตอร์บางตัวถูกบีบให้มีค่าลงเป็นศูนย์ ซึ่งหมายความว่าพารามิเตอร์เหล่านั้นจะถูกตัดออกจากโมเดลโดยอัตโนมัติ ด้วยเหตุนี้ Lasso จึงเป็นเครื่องมือที่ดีในการจัดการกับข้อมูลที่มีคุณลักษณะจำนวนมากและสามารถลดความซับซ้อนของโมเดลได้

2.1.5.5. **การถดถอยแบบริดจ์** (Ridge Regression)

Ridge Regression เป็นเทคนิคการถดถอยที่ใช้ในการวิเคราะห์ข้อมูลเพื่อลดปัญหาการ Overfitting โดยการเพิ่มพูน ในการวิเคราะห์โมเดลทางสถิติ ผ่านการเพิ่มค่าความผิดพลาด ด้วยผลรวมของพารามิเตอร์ที่มีการยกกำลังสอง ซึ่งเรียกว่า L2 Regularization

Ridge Regression มีวัตถุประสงค์เพื่อลดค่าความผิดพลาดโดยการรวมค่าความผิดพลาดของโมเดล กับการเพิ่มพูนที่เกิดจากค่าพารามิเตอร์ในโมเดล

$$Objective\ Function\ = RSS + \lambda \sum_{j=1}^{p} \left|\beta_j\right| \tag{2.9}$$

โดย β_j หมายถึง พารามิเตอร์ที่ถูกฝึกฝน

λ ค่าพารามิเตอร์ที่ใช้กำหนดขนาดของการเพิ่มพูน ถ้า λ เป็น 0 จะเท่ากับการทำ Linear Regression ธรรมดา

$$\mathrm{RSS} = \sum_{i=1}^{n} \left(y_i - \hat{y}_i\right)^2 \tag{2.10}$$

การเพิ่มค่าของ λ จะช่วยลดขนาดของค่าพารามิเตอร์ แต่จะไม่ทำให้พารามิเตอร์ใด ๆ ถูกบีบให้เป็นศูนย์ ซึ่งแตกต่างจาก Lasso Regression ที่อาจบีบให้บางพารามิเตอร์เป็นศูนย์ได้ ทำให้ Ridge Regression สามารถใช้กับข้อมูลที่มีหลายพารามิเตอร์ที่มีความสัมพันธ์กันได้ดีกว่า

2.1.5.6. **การถดถอยแบบริดจ์ของเบส์** (Bayesian Ridge Regression)

การถดถอยแบบริดจ์ของเบส์ (Bayesian Ridge Regression) เป็นวิธีการถดถอยที่ใช้แนวทาง Bayesian ในการปรับค่าพารามิเตอร์ของโมเดล โดยมุ่งหวังที่จะจัดการกับปัญหา Overfitting และความไม่แน่นอนของข้อมูล โดยการใช้ความเชื่อมั่น (Prior Beliefs)

เกี่ยวกับค่าพารามิเตอร์ของโมเดล ซึ่งสามารถช่วยเพิ่มความเสถียรในการทำนายและทำให้การตีความโมเดลง่ายขึ้น

การทำงานของ Bayesian Ridge Regression เริ่มจากการสร้างแบบจำลองโดยการใช้การกระจายของค่าพารามิเตอร์ (Prior Distribution) แทนที่จะใช้ค่าคงที่เหมือนในวิธีการถดถอยทั่วไป เมื่อมีข้อมูลใหม่เข้ามา ค่าพารามิเตอร์จะถูกปรับโดยใช้กฎของ Bayes เพื่อให้ได้การกระจายแบบ Posterior ซึ่งรวมถึงข้อมูลใหม่ที่ได้มา

การตั้งค่าการกระจายเบื้องต้น (Prior Distribution) ใน Bayesian Ridge Regression ค่าเบื้องต้นสำหรับพารามิเตอร์จะถูกตั้งอยู่ในลักษณะ Gaussian (Normal Distribution) โดยมีค่าเฉลี่ยและความแปรปรวนเป็นพารามิเตอร์ที่สามารถปรับได้ และค่าความเชื่อมั่นในค่าพารามิเตอร์เหล่านี้สามารถปรับได้ผ่าน Hyperparameters ซึ่งอาจรวมถึงค่าของ Variance สำหรับ Prior Distribution โดยการคำนวณการกระจาย Posterior จะถูกคำนวณโดยการรวม Prior Distribution เข้ากับ Likelihood ของข้อมูลซึ่งผลลัพธ์คือการกระจายที่ปรับค่าพารามิเตอร์ซึ่งสามารถนำไปใช้ในการทำการคาดการณ์ และ Bayesian Ridge Regression จะทำการคาดการณ์โดยใช้ค่าเฉลี่ยของการกระจาย Posterior ของพารามิเตอร์ ซึ่งให้ความเชื่อมั่นเกี่ยวกับความไม่แน่นอนในผลลัพธ์ที่ได้

2.1.5.7. **การถดถอยแบบกราเดียนส์บูสต์ติง** (Gradient Boosting Regressor)

การถดถอยแบบกราเดียนส์บูสต์ติง (Gradient Boosting Regressor) เป็นเทคนิคการถดถอยที่ใช้กระบวนการเพิ่มพูน (Boosting) โดยการรวมโมเดลพื้นฐานหลาย ๆ โมเดลเข้าด้วยกันเพื่อให้ได้ผลลัพธ์ที่แม่นยำขึ้น โมเดลพื้นฐานที่มักใช้ใน Gradient Boosting คือ Decision Trees ที่มีขนาดเล็ก เรียกว่า Weak Learners โดยการรวมหลาย ๆ weak learner เหล่านี้จะช่วยเพิ่มความแม่นยำในการทำนาย

การทำงานของ Gradient Boosting Regressor เริ่มจากการเพิ่มพูนแบบเป็นขั้นตอน (Sequential Boosting) โดย Gradient Boosting จะสร้างโมเดลใหม่เป็นขั้น ๆ (Iteratively) โดยแต่ละขั้นจะสร้างโมเดลใหม่เพื่อลดความผิดพลาดของโมเดลก่อนหน้า ซึ่งความผิดพลาดนี้จะใช้วิธีการคำนวณจาก Gradient ของค่าความผิดพลาด (Loss Function) ที่คำนวณจากผลลัพธ์ที่คาดการณ์กับค่าจริง โดยเริ่มต้นจากการสร้างโมเดลแรกที่คาดการณ์ค่าเฉลี่ยของข้อมูลเป้าหมาย (Target Variable) และสร้างโมเดลต่อมาเพื่อลดข้อผิดพลาดที่เกิดขึ้น

การคำนวณ Gradient ในแต่ละขั้นตอน Gradient Boosting จะคำนวณ Gradient ของค่าความผิดพลาดในแต่ละจุดข้อมูล เพื่อบ่งบอกถึงทิศทางการเปลี่ยนแปลงที่ควรปรับปรุงในโมเดล โดยโมเดลใหม่ที่สร้างขึ้นจะเรียนรู้จากค่า Gradient นี้เพื่อลดข้อผิดพลาด (เช่น ลด Mean Squared Error ในกรณีของการถดถอย) ให้มากที่สุด

การปรับปรุงค่าด้วย Learning Rate ซึ่งผลลัพธ์ของโมเดลใหม่แต่ละขั้นตอนจะถูกปรับตามค่า Learning Rate เพื่อควบคุมความเร็วในการปรับพารามิเตอร์ของโมเดลโดยค่า Learning Rate มีผลอย่างมากต่อความเสถียรของโมเดล ค่า Learning Rate ที่ต่ำจะทำให้การเรียนรู้ช้าลงแต่มีโอกาสลด Overfitting ในขณะที่ค่าที่สูงอาจทำให้การเรียนรู้เร็วขึ้นแต่มีความเสี่ยงในการ Overfitting และผลลัพธ์ที่ได้จากโมเดลในแต่ละขั้นตอนจะถูกรวมเข้าด้วยกันเพื่อสร้างผลลัพธ์สุดท้าย ซึ่งเป็นการทำนายที่แม่นยำมากขึ้น

2.1.5.8. **สโตแคสติกแกรเดียนต์เดสเซนท์** (Stochastic Gradient Decent)

สโตแคสติกส์แกรเดียนต์เดสเซนท์ (Stochastic Gradient Descent, SGD) เป็นอัลกอริทึมที่ใช้สำหรับการหาค่าที่เหมาะสมที่สุดของพารามิเตอร์ของโมเดลในการเรียนรู้เชิงลึกและการถดถอย โดยการปรับค่าพารามิเตอร์ในทิศทางที่ลดค่าความผิดพลาด (Loss Function) ของโมเดล โดยการใช้เทคนิค Gradient Descent ในการปรับปรุงค่า ซึ่ง SGD มีข้อแตกต่างจาก Gradient Descent ทั่วไปในแง่ของการเลือกข้อมูลในการปรับปรุงค่าพารามิเตอร์

การทำงานของ Stochastic Gradient Descent จะทำการสุ่มเลือกข้อมูล (Stochastic Sampling) แทนที่จะคำนวณ Gradient จากข้อมูลทั้งหมดในทุก ๆ ขั้นตอนการปรับปรุง (เช่นใน Batch Gradient Descent) SGD จะสุ่มเลือกข้อมูลเพียงหนึ่งตัวหรือจำนวนเล็ก ๆ ของข้อมูล (Mini-Batch) มาใช้ในการคำนวณ Gradient ในแต่ละครั้ง การสุ่มนี้ทำให้การปรับปรุงเกิดขึ้นเร็วกว่าและช่วยลดเวลาคำนวณได้มาก โดยเฉพาะกับชุดข้อมูลที่มีขนาดใหญ่

พารามิเตอร์ของโมเดลจะถูกปรับปรุงในทิศทางตรงข้ามกับ Gradient ของค่าความผิดพลาดโดยตรง ซึ่งทำให้ค่าพารามิเตอร์เคลื่อนที่ไปในทิศทางที่ลดค่าความผิดพลาดของโมเดล

$$\theta := \theta - \eta \cdot \nabla_{\theta} j\left(\theta; x_i, y_i\right) \tag{2.11}$$

โดย θ หมายถึง ค่าพารามิเตอร์ของโมเดล

η หมายถึง Learning Rate

$\nabla_{\theta} j\left(\theta; x_i, y_i\right)$ หมายถึง Gradient ของค่าความผิดพลาดของข้อมูลตัวอย่าง

ค่า Learning Rate มีผลสำคัญต่อการทำงานของ SGD ค่า Learning Rate ที่สูงเกินไปอาจทำให้การหาค่าที่เหมาะสมไม่นิ่ง (Oscillating) ส่วนค่าที่ต่ำเกินไปอาจทำให้การเรียนรู้ช้า ดังนั้นการตั้งค่า Learning Rate ที่เหมาะสมเป็นสิ่งสำคัญ และการทำ Mini-Batch Gradient Descent คือการใช้ข้อมูลจำนวนน้อยในแต่ละขั้นตอนของการปรับปรุงค่า (เช่น 32 หรือ 64 ตัวอย่าง) ซึ่งทำให้การปรับปรุงค่ามีความเสถียรมากกว่าแบบสุ่มอย่างเดียวและเร็วกว่าแบบ Batch Gradient Descent ที่ใช้ข้อมูลทั้งหมด

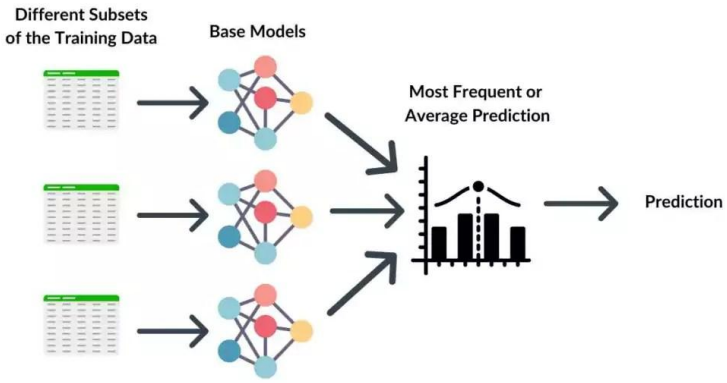
2.1.6. **เทคนิคการรวมโมเดล** (Ensemble Method)

เทคนิคการรวมโดมเดล (Ensemble Method) เป็นเทคนิคใน Machine Learning ที่รวมผลลัพธ์ของโมเดลหลาย ๆ โมเดลเข้าด้วยกันเพื่อเพิ่มความแม่นยำและความเสถียรในการทำนาย ซึ่งการใช้โมเดลหลายโมเดลช่วยลดปัญหา Overfitting และช่วยเพิ่มความน่าเชื่อถือในการพยากรณ์หรือจำแนกข้อมูล วิธี Ensemble นี้ได้รับความนิยมเนื่องจากสามารถนำเสนอผลลัพธ์ที่ดีขึ้นกว่าโมเดลเดี่ยว ๆ ที่ทำงานเพียงลำพัง

การทำงานของ Ensemble Method ทำงานโดยการผสมผสานผลลัพธ์ของโมเดลหลายตัวเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด วิธีการผสมผสานนี้มีหลายวิธีที่เป็นที่นิยม เช่น การลงคะแนนเสียง (Voting) การถ่วงน้ำหนัก หรือการใช้โมเดลหนึ่งมาเป็นพื้นฐานและสร้างโมเดลต่อไปให้เหมาะสมยิ่งขึ้นตามผลลัพธ์ของโมเดลก่อนหน้า

2.1.6.1. Bagging

Bagging (Bootstrap Aggregating) เป็นเทคนิคการสร้างตัวอย่างย่อย (Subsample) ของข้อมูลเพื่อสร้างโมเดลหลายตัวแยกจากกัน ตัวอย่างที่เป็นที่นิยมของ Bagging คือ Random Forest ใน Random Forest จะใช้การสุ่มเลือกข้อมูลและฟีเจอร์ในการสร้าง Decision trees หลาย ๆ ต้น จากนั้นจะนำผลลัพธ์จากต้นไม้แต่ละต้นมารวมกันเพื่อตัดสินใจ และ Bagging ช่วยลดความผันผวนของโมเดลโดยการลดการเปลี่ยนแปลงของโมเดลที่เกิดจากการเปลี่ยนแปลงข้อมูล (Variance) ซึ่งช่วยให้ได้ผลลัพธ์ที่เสถียรมากขึ้น

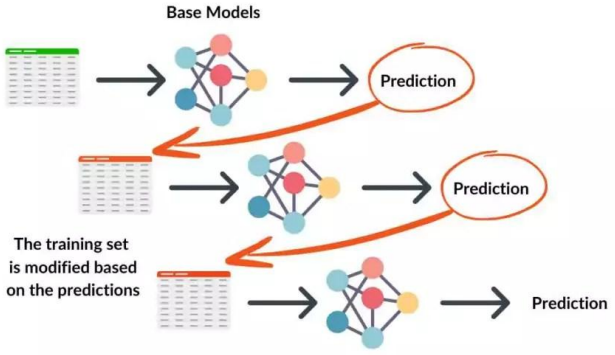


ภาพประกอบ 5 ภาพแสดงการรวมโมเดลแบบ Bagging

ที่มา : https://spotintelligence.com/2024/03/18/bagging-boosting-stacking/

2.1.6.2. Boosting

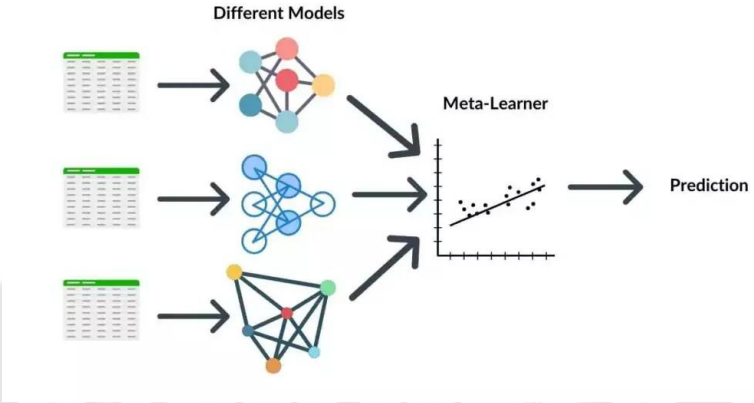
Boosting เป็นเทคนิคที่สร้างโมเดลอย่างต่อเนื่อง โดยโมเดลใหม่จะถูกสร้างขึ้นเพื่อปรับปรุงข้อผิดพลาดของโมเดลก่อนหน้า โดยโมเดลที่ได้รับความนิยมในกลุ่ม Boosting ได้แก่ AdaBoost และ Gradient Boosting ซึ่งแต่ละโมเดลจะให้ความสำคัญกับตัวอย่างที่ทำนายผิดพลาดในการรอบก่อน ๆ มากขึ้น โดย Boosting ช่วยให้โมเดลมีความแม่นยำสูงขึ้น เนื่องจากสามารถปรับปรุงตามจุดอ่อนของโมเดลที่สร้างขึ้นในรอบก่อนหน้า



ภาพประกอบ 6 ภาพแสดงการรวมโมเดลแบบ Boosting

2.1.6.3. Stacking

Stacking เป็นการนำโมเดลหลาย ๆ ตัวมาใช้ร่วมกันในลักษณะที่ซับซ้อนขึ้น โดยใช้โมเดลระดับแรก (Base Models) ทำนายค่า จากนั้นจะนำค่าทำนายเหล่านั้นมาใช้เป็นอินพุตให้กับโมเดลระดับที่สอง (Meta-Model) โดย Meta-model จะทำการทำนายอีกครั้งโดยอิงจากผลลัพธ์ของ Base Models ทำให้การคาดการณ์มีความแม่นยำมากยิ่งขึ้น โดย Stacking เหมาะสำหรับปัญหาที่มีความซับซ้อนมากและต้องการการผสมผสานจากโมเดลหลายรูปแบบเพื่อให้ได้ผลลัพธ์ที่ดียิ่งขึ้น



ภาพประกอบ 7 ภาพแสดงการรวมโมเดลแบบ Stacking

2.1.7. เครื่องมือการแปรผล

2.1.7.1. ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error, RMSE)

ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error, RMSE) เป็นตัวชี้วัดที่ใช้ประเมินความแม่นยำของโมเดลในการทำนาย โดย RMSE เป็นการคำนวณระยะห่างระหว่างค่าที่โมเดลทำนายกับค่าจริงในรูปแบบของรากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนยกกำลังสอง ซึ่งช่วยให้เราเข้าใจว่าผลการทำนายของโมเดลมีความแม่นยำมากเพียงใด และบอกถึงระดับของความผิดพลาดที่เกิดขึ้นจากการทำนาย

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{2.12}$$

โดย y_i หมายถึง ค่าจริงของข้อมูลในลำดับที่ i

$\hat{y}_i$ หมายถึง ค่าทำนายของข้อมูลในลำดับที่ i

n หมายถึง จำนวนตัวอย่างทั้งหมด

RMSE จะให้ค่าในหน่วยเดียวกับค่าที่เราทำนาย เช่น ถ้าค่าทำนายคืออุณหภูมิในองศาเซลเซียส RMSE จะให้ค่าเป็นองศาเซลเซียสเช่นกัน ซึ่ง RMSE สามารถใช้เปรียบเทียบโมเดลหลายโมเดลได้ โดยโมเดลที่มีค่า RMSE ต่ำกว่าจะเป็นโมเดลที่มีความแม่นยำมากกว่าแต่เนื่องจาก RMSE ใช้การยกกำลังสองในการคำนวณ ทำให้มีการลงโทษค่าที่ทำนายผิดพลาดมาก ๆ อย่างสูง เช่น การทำนายที่แตกต่างจากค่าจริงในระดับมาก

2.1.7.2. **ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ** (Coefficient of Multiple Determination, R-Squared)

ค่าสัมประสิทธิ์การตัดสอนใจพหุคูณ (Coefficient of Determination) หรือที่เรียกว่า R-Squared (R^2) เป็นตัวชี้วัดที่ใช้บอกว่าโมเดลการถดถอย (Regression Model) สามารถอธิบายความแปรปรวนของข้อมูลที่มีได้ดีเพียงใด ค่า R^2 แสดงเป็นเปอร์เซ็นต์ของความแปรปรวนของตัวแปรตาม (Dependent Variable) ที่สามารถอธิบายได้ด้วยตัวแปรอิสระ (Independent Variable) ของโมเดลนั้น โดยค่านี้จะมีค่าตั้งแต่ 0 ถึง 1 หรือคิดเป็น 0% ถึง 100%

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_1 - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_1 - \bar{y}\right)^2} \qquad (2.13)$$

โดย y_i หมายถึง ค่าจริงของข้อมูลในลำดับที่ i

$\hat{y}_i$ หมายถึง ค่าทำนายของข้อมูลในลำดับที่ i

$\bar{y}$ หมายถึง ค่าเฉลี่ยของตัวแปรตามทั้งหมด

n หมายถึง จำนวนตัวอย่างทั้งหมด

การตีความผลที่ได้จากการคำนวณ โดยจะกำหนดให้กรณี $R^2 = 1$ หมายถึงโมเดลสามารถอธิบายความแปรปรวนของข้อมูลได้ทั้งหมด หรือ 100% ซึ่งแสดงถึงการทำนายที่แม่นยำมาก กรณี $R^2 = 0$ หมายถึง โมเดลไม่สามารถอธิบายความแปรปรวนของข้อมูลได้เลย และ R^2 ใกล้เคียง 1 แสดงว่าโมเดลมีความแม่นยำสูงและสามารถอธิบายความแปรปรวนของข้อมูลได้ดี ในขณะที่ค่า R^2 ต่ำแสดงว่าโมเดลมีความแม่นยำน้อยกว่า โดยค่า R^2 บ่งบอกถึงสัดส่วนของความแปรปรวนที่โมเดลสามารถอธิบายได้ แต่ค่า R^2 สูงไม่จำเป็นต้องหมายถึงโมเดลที่ดีที่สุดเสมอไป โมเดลอาจเกิดปัญหา Overfitting หากค่า R^2 สูงเกินไป นอกจากนี้ยังมีค่า Adjusted R-Squared ที่พิจารณาความซับซ้อนของโมเดลเพื่อช่วยประเมินโมเดลที่มีตัวแปรหลายตัวได้แม่นยำมากขึ้น

2.1.7.3. ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percent Error, MAPE)

ค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย หรือ MAPE เป็นตัวชี้วัดที่สามารถอธิบายความแม่นยำของแบบจำลอง ที่เกิดจากการพยากรณ์อนุกรมเวลา เช่น การพยากรณ์ราคาทองคำ หรือยอดขาย โดยสามารถคำนวณได้จากค่าเฉลี่ยของอัตราร้อยละของค่าความคลาดเคลื่อนสัมบูรณ์ระหว่างค่าจริงและค่าที่พยากรณ์ได้ ซึ่งแสดงในรูปแบบร้อยละ หรือเปอร์เซ็นต์ ทำให้ผลลัพธ์ที่สามารถตีความได้ง่ายและสะดวกต่อการเปรียบเทียบระหว่างแบบจำลอง

$$\text{MAPE} = \frac{\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100}{n} \tag{2.14}$$

โดย y_i หมายถึง ค่าจริงของข้อมูลในลำดับที่ i

$\hat{y}_i$ หมายถึง ค่าทำนายของข้อมูลในลำดับที่ i

n หมายถึง จำนวนตัวอย่างทั้งหมด

การเปรียบเทียบค่า MAPE หากน้อยกว่า 10% หมายถึงการพยากรณ์อยู่ในระดับที่ค่อนข้างแม่นยำ หากอยู่ในช่วง 10% ถึง 20% หมายถึงการพยากรณ์ใช้ได้ดี หากอยู่ในช่วง 20% ถึง 50% หมายถึงการพยากรณ์อยู่ในระดับพอใช้ และหากมากกว่า 50% หมายถึงการพยากรณ์ไม่มีความแม่นยำ ซึ่งแปลความหมายได้ว่าค่า MAPE ที่ต่ำแสดงให้เห็นว่าแบบจำลองมีความแม่นยำในการพยากรณ์มากกว่าค่า MAPE ที่สูง

2.2. งานวิจัยที่เกี่ยวข้อง

2.2.1. บทความวิจัยเรื่อง Gold Price Prediction using Machine Learning **โดย** Rushikesh Ghule **และ** Abhijeet Gadhave (Ghule et al., 2022)

งานวิจัยนี้นำเสนอการคาดการณ์ราคาทองคำในกองทุนรวมที่ซื้อขายในตลาดหลักทรัพย์ (Gold Exchange-Traded Fund, ETF) โดยตัวแปลด้านการตลาด 22 ตัวและเทคนิคการเรียนรู้ของเครื่องโดยใช้การถดถอยของป่าแบบสุ่ม และการถดถอยเชิงเส้น ซึ่งเป็นอัลกอริทึมการเรียนรู้ของเครื่องแบบมีผู้สอน โดยใช้เครื่องมือการแปรผลแบบค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง หรือ R^2 เป็นเครื่องมือการแปรผลความผิดพลาดของการทำงานของเครื่อง

การทดลองของงานวิจัยนี้พบว่าการคาดการณ์ราคาทองคำโดยใช้การถดถอยของป่าแบบสุ่มสามารถคาดการราคาทองคำได้ดีกว่า โดยมีค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสองอยู่ที่ 0.975509 หรือ 97% ซึ่งดีกว่าการการทำนายแบบการถดถอยเชิงเส้น

2.2.2. **บทความวิจัยเรื่อง** Efficient Machine Learning Algorithm for Future Gold Price Prediction **โดย** MinalGhute **และ** Mridula Korde (Ghute & Korde, 2023)

งานวิจัยนี้นำเสนอการคาดการณ์ราคาทองคำจากการเทคนิคการเรียนรู้ของเครื่องแบบหลากหลายรูปแบบ ได้แก่ ต้นไม้ตัดสินใจ, การถดถอยเชิงเส้น, การถดถอยของป่าแบบสุ่ม, ซัพพอร์ตเวกเตอร์ และการถดถอยแบบริดจ์ โดยใช้เครื่องมือการแปรผลแบบค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง หรือ RMSE และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ หรือ R^2 เป็นเครื่องมือการแปรผลความผิดพลาดของการทำงานของเครื่อง ซึ่งข้อมูลจะถูกเก็บรวบรวมและทำการประมวลผลล่วงหน้า จากนั้นแบ่งข้อมูลเป็น 80% สำหรับการฝึกโมเดลและ 20% สำหรับการทดสอบ

การทดลองของงานวิจัยนี้พบว่าการคาดการณ์ราคาทองคำโดยใช้การถดถอยด้วยป่าแบบสุ่ม (Random Forest Regression) ให้ผลลัพธ์ที่แม่นยำที่สุด โดยมีค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสองอยู่ที่ 2.3869 และค่าสัมประสิทธิ์การตัดสินใจพหุคูณอยู่ที่ 0.9891 ซึ่งดีที่สุดเมื่อเปรียบเทียบกับอัลกอริทึมอื่น

2.2.3. **บทความวิจัยเรื่อง** Forecasting Gold Prices Using Multiple Linear Regression Method **โดย** Z. Ismail, A. Yahya **และ** A. Shabri (Ismail et al., 2009)

งานวิจัยนี้นำเสนอกระบวนการในการประมาณสถานการณ์ในอนาคตที่ยังไม่ทราบโดยศึกษาการพัฒนาโมเดลการคาดการณ์ราคาทองคำโดยอิงจากปัจจัยทางเศรษฐกิจ เช่น อัตราเงินเฟ้อ การเคลื่อนไหวของอัตราแลกเปลี่ยน และอื่น ๆ โดยใช้เทคนิคการเรียนรู้ของเครื่อง แบบการถดถอยเชิงเส้นหลายตัวแปร (Multiple Linear Regression, MLR) ซึ่งศึกษาเกี่ยวกับความสัมพันธ์ระหว่างตัวแปรตาม หรือราคาทองคำ และตัวแปรอิสระหลายตัว และใช้เครื่องมือการแปรผลแบบค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง หรือ RMSE เป็นเครื่องมือการแปรผลความผิดพลาดของการทำงานของเครื่อง

การทดลองของงานวิจัยนี้พบว่าการคาดการณ์ราคาทองคำโดยใช้การถดถอยเชิงเส้นหลายตัวแปร โดยใช้ตัวแปรอิสระทั้งหมดสามารถอธิบายความแปรผันของราคาทองคำรายเดือนได้ถึง 85.2% และเมื่อพิจารณาใช้ตัวแปรอิสระที่สำคัญ ได้แก่ Research Bureau future index

(CRB), อัตราแลกเปลี่ยน USD/Euro (EUROUSD), อัตราเงินเฟ้อ (INF) และ อุปทานเงิน (M1) ซึ่งแสดงผลการคาดการณ์ที่มีความแม่นยำสูง โดยอธิบายความแปรผันได้ประมาณ 70%

2.2.4. **บทความวิจัยเรื่อง** Forecasting Gold Prices: Leveraging Machine Learning Algorithms **โดย** Alok Kumar Pati, Samitinjay Mishra, Subham Shankar Sahoo **และ** Arabinda Kar (Pati et al., 2024)

งานวิจัยนี้มุ่งเนการคาดการณ์ราคาทองคำ โดยใช้ดัชนีทางการเงิน เช่น ดัชนีหุ้นสหรัฐอเมริกา (Standard & Poor's 500 Index, S&P500), กองทุนซื้อขายแลกเปลี่ยนที่ติดตามราคาน้ำมันดิบ (United States Oli Exchange Traded Fund), Silver Trust (SLV), และอัตราแลกเปลี่ยนยูโรต่อดอลลาร์สหรัฐ (EUR/USD) โดยใช้ข้อมูลเริ่มต้นตั้งแต่วันที่ 2 มกราคม 2008 โดยใช้เทคนิคการเรียนรู้ของเครื่อง แบบต้นไม้ตัดสินใจ, การถดถอยเชิงเส้น, การถดถอยลาสโซ, การถดถอยริดจ์, และการถดถอยของป่าแบบสุ่ม โดยใช้เครื่องมือการแปรผลแบบค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง หรือ RMSE เป็นเครื่องมือการแปรผลความผิดพลาดของการทำงานของเครื่อง

การทดลองของงานวิจัยนี้พบว่าการคาดการณ์ราคาทองคำโดยใช้การถดถอยของป่าแบบสุ่มให้ผลลัพธ์ที่แม้นยำที่สุด โดยมีค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสองอยู่ที่ 5.375175

2.2.5. **บทความวิจัยเรื่อง** Gold Price Prediction using Ensemble based Machine Learning Techniques **โดย** Manjula K A **และ** Karthikeyan P (Manjula & Karthikeyan, 2019)

งานวิจัยนี้มุ่งเน้นการศึกษาที่ดำเนินการเพื่อเข้าใจความสัมพันธ์ระหว่างราคาทองคำและปัจจัยที่มีอิทธิพล ซึ่งได้แก่ ตลาดหุ้น, ราคาน้ำมันดิบ, อัตราแลกเปลี่ยนรูปีต่อดอลลาร์, อัตราเงินเฟ้อ และอัตราดอกเบี้ย โดยใช้ข้อมูลระหว่าง เดือนมกราคม 2000 ถึงเดือนธันวาคม 2018 โดยแบ่งข้อมูลออกเป็น 2 ช่วง คือ ในช่วงที่ราคาทองมีแนวโน้มเพิ่มมากขึ้น และในช่วงที่ราคาทองมีแนวโน้มคงที่ โดยใช้เทคนิคการเรียนรู้ของเครื่อง ได้แก่ การถดถอยเชิงเส้น, การถดถอยของป่าแบบสุ่ม และการถดถอยแบบ Gradient Boosting

การเปรียบเทียบความแม่นยำในการคาดการณ์โดยใช้ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ย (Mean Squared Error, MSE), ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง หรือ R^2 และค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Error, MAE) ซึ่งการ

ถดถอยของป่าแบบสุ่มมีค่าต่ำที่สุดสำหรับการคาดการณ์ข้อมูลทั้งหมดทุกช่วงเวลา ในขณะที่โมเดลกการถดถอยแบบ Gradient Boosting มีค่าต่ำที่สุดสำหรับช่วงเวลาที่เหลืออีกสองช่วง แม้ว่าโมเดลการถดถอยของป่าแบบสุ่มจะมีความแม่นยำดีกว่าสำหรับช่วงเวลายาว แต่สำหรับช่วงเวลาสั้น โมเดลการถดถอยแบบ Gradient Boosting ดีกว่า

บทที่ 3
การดำเนินการวิจัย

งานวิจัยนี้มุ่งเน้นการศึกษาเพื่อการทำนายราคากองทุนทองคำ และหาค่าความแม่นยำระหว่างราคากองทุนทองคำจากการทำนายจากการเรียนรู้ของเครื่อง และราคาจริง เพื่อช่วยให้นักลงทุน และผู้ที่เกี่ยวข้องเข้าใจแนวโน้มของราคากองทุนทองคำในอนาคตได้ดีขึ้น ช่วยในการตัดสินใจลงทุน และช่วยในการตัดสอนใจเพื่อปรับกลยุทธ์การลงทุนได้ดีขึ้น ซึ่งงานวิจัยนี้จะศึกษาวิเคราะห์ เปรียบเทียบแบบจำลองการเรียนรู้ของเครื่อง ประกอบไปด้วย ซัพพอร์ตเวกเตอร์รีเกสชัน, การถดถอยแบบต้นไม้ตัดสินใจ, การถดถอยของป่าแบบสุ่ม, การถดถอยลาสโซ, การถดถอยแบบริดจ์, การถดถอยของริดจ์แบบเบส์, ถดถอยแบบกราเดียนท์บูสติง และสโตแคสติกแกรเดียนท์เดสเซนท์

ในการวิจัยครั้งนี้ ผู้วิจัยได้ดำเนินการวิจัยตามขั้นตอน ดังนี้

3.1. กระบวนการทำงานของแบบจำลอง (Model Overview Workflow)

3.2. กระบวนการสำรวจข้อมูล (Exploratory Data Analysis)

3.3. กระบวนการสร้างและปรับปรุงคุณลักษณะ (Feature Engineering)

3.4. กระบวนการคัดเลือกคุณลักษณะ (Feature Selection)

3.5. การตั้งค่าและการฝึกฝนแบบจำลอง (Model Setup and Training)

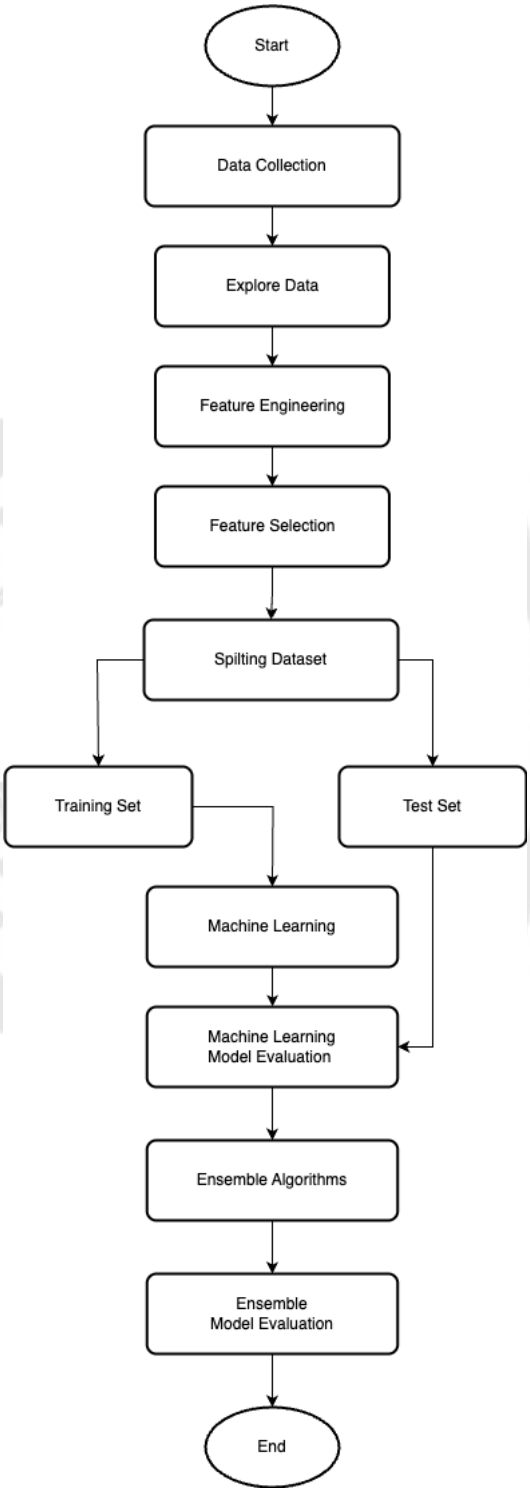
3.5.1. การแบ่งชุดข้อมูลฝึกฝนและทดสอบ (Splitting Dataset)

3.5.2. การตั้งค่าแบบจำลอง (Hyperparameter Optimization)

3.6. การประเมินผลแบบจำลอง (Model Evaluation)

3.7. การบวนการใช้เทคนิคการรวมโมเดล (Ensemble Algorithms)

3.1. กระบวนการทำงานของแบบจำลอง (Model Overview Workflow)



ภาพประกอบ 8 กระบวนการทำงานของแบบจำลอง (Medel Overview Workflow)

ภาพประกอบที่ 7 อธิบายกระบวนการทำงานของแบบจำลอง โดยเริ่มจากการนำเข้าข้อมูลเพื่อใช้ในการเรียนรู้และทดสอบ หลังจากนำเข้าข้อมูลสำเร็จแล้วจึงทำการสำรวจข้อมูล เพื่อทำความเข้าใจ และวิเคราะห์ความสัมพันธ์เบื้องต้นของข้อมูล และเพิ่มตัวชี้วัดที่สำคัญทางสถิติเพื่อเป็นตัวเลือกที่เพิ่มขึ้นให้กับการทำนาย และปรับเปลี่ยนช่วงของข้อมูลตัวเลขโดยวิธีการ Min Max Scalar เพื่อให้ง่ายต่อการทำงาน และเพิ่มเติมคุณลักษณะโดยวิธีการทางวิศวกรรมคุณลักษณะ จากนั้นทำการคัดเลือกคุณลักษณะที่สำคัญ โดยใช้เทคนิคการเลือกคุณลักษณะ แบบการถดถอยลาสโซ, การถดถอยของป่าแบบสุ่ม และการค้นหาขั้นตอนวิธีเชิงพันธุกรรม เมื่อได้รับการคัดเลือกคุณลักษณะที่เหมาะสม จากนั้นทำการแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลสำหรับการเรียนรู้ของแบบจำลอง (Training Set) และข้อมูลสำหรับทดสอบประสิทธิภาพของแบบจำลอง (Test Set) โดยงานวิจัยนี้เลือกใช้อัลกอริทึม 8 อัลกอริทึม ได้แก่ ซัพพอร์ตเวกเตอร์รีเกสชัน, การถดถอยแบบต้นไม้ตัดสินใจ, การถดถอยของป่าแบบสุ่ม, การถดถอยลาสโซ, การถดถอยแบบริดจ์, การถดถอยของริดจ์แบบเบส์, การถดถอยแบบกราเดียนส์บูสติง และสโคแคสติกกราเดียนเดสเซนท์ พร้อมทั้งการปรับจูนพารามิเตอร์ และทดสอบความแม่นยำของการทำนายโดยการนำข้อมูลที่ได้จากการทำนายมาเทียบกับข้อมูลสำหรับทดสอบ โดยใช้ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง, ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ และค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ซึ่งเมื่อได้ค่าความแม่นยำจากการเรียนรู้ของเครื่องแล้ว จากนั้นทำการเลือกการเลือกอัลกอริทึมที่ดีที่สุด 1, 2, 3 และ 4 อันดับ เพื่อทดสอบเทคนิคการรวมโมเดล และทำการทดสอบความแม่นยำของเทคนิคการรวมโมเดล โดยใช้ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง, ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ และค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย

3.2. **กระบวนการสำรวจข้อมูล** (Exploratory Data Analysis)

3.2.1. **การรวบรวมข้อมูล** (Data Collection)

การเก็บรวบรวมข้อมูลสำหรับการดำเนินการวิจัยนี้มุ่งเน้นการจัดหาข้อมูลจากแหล่งที่มาที่หลากหลาย เพื่อให้ได้ข้อมูลที่มีความทันสมัย ครอบคลุม และครบถ้วนสมบูรณ์ ซึ่งกระบวนการดังกล่าวถือเป็นขั้นตอนสำคัญที่ส่งผลต่อคุณภาพและความน่าเชื่อถือของผลการวิจัย โดยข้อมูลสำคัญที่ใช้ในการดำเนินการวิจัยในครั้งนี้ ได้แก่ ราคาทองคำ ซึ่งเป็นตัวแปรหลักที่มีบทบาทสำคัญในกระบวนการวิเคราะห์เพื่อพยากรณ์ราคาทองคำ ร่วมกับการคัดเลือกคุณลักษณะที่มีความสัมพันธ์ใกล้ชิดกับราคาทองคำ ได้แก่ ดัชนีตลาดหุ้นสหรัฐอเมริกา (Standard & Poor's 500 Index : S&P500), ดัชนีดาวโจนส์ (Down Jones Index), ดัชนีหุ้นเอลโดราโด โกลด์ คอร์ปอเรชัน(Eldorado Gold Corporation : EGO), อัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ (Euro

US Dollar), กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา (United States Brent Oil Fund), ดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ (W&T Offshore Stock), สัญญาซื้อขายแร่เงินล่วงหน้า (Silver Futures), อัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว (United States 10-Year Bond Yield), สัญญาซื้อขายแร่แพลทินัมล่วงหน้า (Platinum Futures), สัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า (Palladium Futures), ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา (US Dollar Index), กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ (VanEck Gold Miners ETF) และกองทุนรวมราคาน้ำมันสหรัฐเมริกา (United States Oil Fund)

3.2.1.1. กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ (Gold ETF, GLD)

งานวิจัยการพยากรณ์ราคาทองคำจะมุ่งเน้นไปที่การทำนายราคาทองคำโดยอาศัยข้อมูลจากกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ ซึ่งเป็นเครื่องมือการลงทุนที่ได้รับความนิยมในตลาดหลักทรัพย์และมีการติดตามราคาทองคำในตลาดโลก โดยแต่ละหน่วยของกองทุนดังกล่าวมีมูลค่าเท่ากับราคาทองคำ 1 กรัม

งานวิจัย Gold vs Gold Exchange Traded Funds: An Empirical Study in India (Kumar & Raj, 2019)ซึ่งเป็นงานวิจัยที่มีวัตถุประสงค์เพื่อศึกษาความแตกต่างระหว่างการลงทุนในทองคำจริง และกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ (Gold ETFs) ภายในบริบทของตลาดการเงินในประเทศอินเดีย โดยมุ่งเน้นไปที่การเปรียบเทียบผลตอบแทนและความเสี่ยงที่เกิดจากการลงทุนในสินทรัพย์ทั้งสองประเภท เพื่อให้ข้อมูลที่มีความสำคัญสำหรับนักลงทุนในการตัดสินใจเลือกเครื่องมือการลงทุนที่เหมาะสม พบว่าการลงทุนในกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำและทองคำจริงมีผลตอบแทนที่ใกล้เคียงกัน เนื่องจากราคาของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ จะสะท้อนถึงราคาทองคำในตลาดโลกอย่างใกล้ชิด ทั้งสองประเภทการลงทุนได้รับผลกระทบจากปัจจัยภายนอกเช่น การเปลี่ยนแปลงในเศรษฐกิจโลกและอัตราดอกเบี้ย อย่างไรก็ตาม การลงทุนในกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีข้อได้เปรียบในแง่ของความสะดวกในการซื้อขายและค่าใช้จ่ายที่ต่ำกว่าการลงทุนในทองคำจริง ซึ่งไม่ต้องกังวลเกี่ยวกับการเก็บรักษาหรือประกันภัยทองคำ ทั้งนี้ การลงทุนกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ อาจมีความเสี่ยงจากความผันผวนของราคาทองคำในระยะสั้นและความผันผวนในตลาดการเงิน ขณะที่การลงทุนในทองคำจริงอาจเผชิญความเสี่ยงในด้านการเก็บรักษาและการขโมย โดยสรุปแล้ว การเลือกเครื่องมือการลงทุนในทองคำทั้งสองประเภทขึ้นอยู่กับความเหมาะสมของนักลงทุนตามวัตถุประสงค์การลงทุนและระดับความเสี่ยงที่สามารถยอมรับได้

ข้อมูลกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations ซึ่งเป็นตลาดหลักทรัพย์ที่มีความสำคัญระดับโลก ซึ่งดำเนินการซื้อขายหลักทรัพย์ต่าง ๆ และยังเป็นแหล่งข้อมูลสำคัญสำหรับการติดตามราคาหุ้น, ดัชนีตลาด, ข่าวสารทางการเงิน, การวิเคราะห์ทางการเงิน และข้อมูลที่เกี่ยวข้องกับการลงทุนในตลาดทุน โดยข้อมูลกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 1 ตารางแสดงโครงสร้างข้อมูลของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ
2	Close/Last	ราคาปิดตลาดของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ
3	Volume	ปริมาณการซื้อขายภายในวันของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ
4	Open	ราคาเปิดตลาดของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ
5	High	ราคาสูงสุดภายในวันของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ
6	Low	ราคาต่ำสุดภายในวันของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ

3.2.1.2. **ดัชนีตลาดหุ้นสหรัฐอเมริกา** (Standard&Poor's 500 Index, S&P500)

ดัชนีตลาดหุ้นสหรัฐอเมริกาเป็นเครื่องมือหนึ่งที่นิยมใช้ในทำนายและพยากรณ์ราคาทองคำ โดยดัชนีตลาดหุ้นสหรัฐอเมริกาเป็นดัชนีที่ประกอบด้วยหุ้นของบริษัทใหญ่ 500 แห่งในสหรัฐอเมริกา โดยมักใช้เป็นดัชนีที่บ่งบอกถึงสภาพเศรษฐกิจของสหรัฐอเมริกา และการเคลื่อนไหวในตลาดหุ้น สัญญาณดัชนีตลาดหุ้นสหรัฐอเมริกามีความสำคัญในการแสดงทิศทางของตลาดการเงินและเศรษฐกิจโดยรวม ซึ่งอาจมีผลกระทบต่อตลาดทองคำ เนื่องจากนักลงทุนมักใช้ทองคำเป็นสินทรัพย์ที่ปลอดภัยในช่วงที่มีความไม่แน่นอนในตลาดหุ้น

งานวิจัย Gold and the S&P500: An analysis of the return and volatility relationship (Kuck, 2021) เป็นงานวิจัยที่มุ่งเน้นการวิเคราะห์ความสัมพันธ์ระหว่างผลตอบแทนและความผันผวนของราคาทองคำและดัชนีดัชนีตลาดหุ้นสหรัฐอเมริกา โดยมีวัตถุประสงค์เพื่อ

สำรวจว่าความเคลื่อนไหวของผลตอบแทนและความผันผวนของดัชนีตลาดหุ้นสหรัฐอเมริกาสามารถอธิบายหรือทำนายการเปลี่ยนแปลงของราคาทองคำได้หรือไม่ โดยผลการศึกษาเผยให้เห็นถึงความสัมพันธ์เชิงลบระหว่างผลตอบแทนของดัชนีตลาดหุ้นสหรัฐอเมริกาและราคาทองคำ โดยเฉพาะในช่วงที่ความผันผวนของตลาดหุ้นเพิ่มขึ้น ซึ่งนักลงทุนมักหันไปลงทุนในทองคำในฐานะสินทรัพย์ที่ปลอดภัยในช่วงที่มีความไม่แน่นอนทางเศรษฐกิจและการเงิน ผลการวิจัยยังระบุว่า เมื่อดัชนีดัชนีตลาดหุ้นสหรัฐอเมริกามีการปรับตัวลดลง ราคาทองคำมักมีการปรับตัวเพิ่มขึ้น และในทางกลับกัน เมื่อดัชนีดัชนีตลาดหุ้นสหรัฐอเมริกามีการปรับตัวเพิ่มขึ้น ราคาทองคำจะมีแนวโน้มลดลง และในส่วนของการวิเคราะห์ความผันผวน พบว่าความผันผวนของดัชนีตลาดหุ้นสหรัฐอเมริกาสามารถส่งผลกระทบต่อความผันผวนของราคาทองคำได้ โดยในช่วงที่ตลาดหุ้นมีความผันผวนสูง ราคาทองคำจะมีความผันผวนตามไปด้วย ซึ่งบ่งชี้ถึงความสัมพันธ์เชิงสาเหตุระหว่างทั้งสองตลาดในแง่ของความเสี่ยงและการลงทุน

ข้อมูลดัชนีตลาดหุ้นสหรัฐอเมริกาถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลดัชนีตลาดหุ้นสหรัฐอเมริกาถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 2 ตารางแสดงโครงสร้างข้อมูลของดัชนีตลาดหุ้นสหรัฐอเมริกา

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของดัชนีตลาดหุ้นสหรัฐอเมริกา
2	Close/Last	ราคาปิดตลาดของดัชนีตลาดหุ้นสหรัฐอเมริกา
3	Open	ราคาเปิดตลาดของดัชนีตลาดหุ้นสหรัฐอเมริกา
4	High	ราคาสูงสุดภายในวันของดัชนีตลาดหุ้นสหรัฐอเมริกา
5	Low	ราคาต่ำสุดภายในวันของดัชนีตลาดหุ้นสหรัฐอเมริกา

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลดัชนีตลาดหุ้นสหรัฐอเมริกาโดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง (Null Value)

3.2.1.3. **ดัชนีดาวโจนส์** (Dow Jones ETF, DJI)

ดัชนีดาวโจนส์เป็นหนึ่งในดัชนีตลาดหลักทรัพย์ที่มีชื่อเสียงและเก่าแก่ที่สุดในโลก มีบทบาทสำคัญในฐานะตัวชี้วัดความเคลื่อนไหวของตลาดหุ้นสหรัฐอเมริกา และสะท้อนถึงสภาพเศรษฐกิจโดยรวมอย่างมีนัยสำคัญ ดัชนีดาวโจนส์ประกอบด้วยหุ้นของบริษัทขนาดใหญ่จำนวน 30 บริษัทในสหรัฐอเมริกา ซึ่งครอบคลุมหลากหลายอุตสาหกรรม เช่น เทคโนโลยี การเงิน พลังงาน และสินค้าอุปโภคบริโภค หุ้นเหล่านี้ล้วนเป็นตัวแทนของอุตสาหกรรมสำคัญและมีบทบาทต่อเศรษฐกิจในระดับประเทศและโลก

งานวิจัย Gold and portfolio diversification: A stochastic dominance analysis of the Dow Jones Islamic indices (Alkhazali & Zoubi, 2020) เป็นงานวิจัยที่มุ่งเน้นการวิเคราะห์บทบาทของทองคำในการกระจายพอร์ตการลงทุน โดยใช้การวิเคราะห์ความโดดเด่นเชิงสถิติ (Stochastic Dominance Analysis) เพื่อศึกษาความสัมพันธ์ระหว่างดาวโจนส์กับการลงทุนในทองคำ ซึ่งพบว่าทองคำมีบทบาทสำคัญในการกระจายความเสี่ยงของพอร์ตการลงทุน โดยสามารถใช้เป็นเครื่องมือป้องกันความเสี่ยงในช่วงที่ตลาดมีความผันผวนสูง โดยเฉพาะเมื่อผสมผสานเข้ากับดัชนี Islamic ที่เน้นการลงทุนในสินทรัพย์ที่มีการปฏิบัติตามหลักการทางศาสนา อาทิ การหลีกเลี่ยงการลงทุนในอุตสาหกรรมที่เกี่ยวข้องกับแอลกอฮอล์ การพนัน หรือธุรกิจที่ผิดหลักศาสนา ผลการศึกษาพบว่า การเพิ่มทองคำในพอร์ตการลงทุนจะช่วยเพิ่มประสิทธิภาพในการกระจายความเสี่ยง และสามารถปรับปรุงผลตอบแทนรวมของพอร์ตได้ในระยะยาว โดยเฉพาะในสภาวะตลาดที่มีความไม่แน่นอนสูง ซึ่งแสดงให้เห็นว่าราคาทองคำและดัชนีดาวโจนส์มีความสัมพันธ์กันในด้านการกระจายความเสี่ยงของพอร์ตการลงทุน

ข้อมูลดัชนีดาวโจนส์ถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลดัชนีดาวโจนส์ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 3 ตารางแสดงโครงสร้างข้อมูลของดัชนีดาวโจนส์

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของดัชนีดาวโจนส์
2	Close/Last	ราคาปิดตลาดของดัชนีดาวโจนส์
3	Volume	ปริมาณการซื้อขายภายในวันของดัชนีดาวโจนส์

ตาราง 3 (ต่อ) ตารางแสดงโครงสร้างข้อมูลของดัชนีดาวโจนส์

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
4	Open	ราคาเปิดตลาดของดัชนีดาวโจนส์
5	High	ราคาสูงสุดภายในวันของดัชนีดาวโจนส์
6	Low	ราคาต่ำสุดภายในวันของดัชนีดาวโจนส์

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลดัชนีดาวโจนส์ โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.4. **ดัชนีหุ้นเอลโดราโด โกลด์** (Eldorado Gold Corporation, EGO)

เอลโดราโด โกลด์ คอร์ปอเรชัน (Eldorado Gold Corporation) เป็นบริษัทเหมืองแร่ทองคำระดับโลกที่ก่อตั้งขึ้นในปี 1992 โดยมีสำนักงานใหญ่ตั้งอยู่ในเมืองแวนคูเวอร์ ประเทศแคนาดา บริษัทมุ่งเน้นการสำรวจ พัฒนา และดำเนินกิจการเหมืองแร่ทองคำ รวมถึงแร่โลหะอื่น ๆ เช่น เงิน ทองแดง และเหล็ก ในหลายภูมิภาคทั่วโลก โดยเฉพาะในประเทศตุรกี กรีซ แคนาดา และโรมาเนีย เอลโดราโด โกลด์ มีจุดเด่นด้านการดำเนินงานด้วยต้นทุนที่ต่ำ และมุ่งมั่นในการบริหารจัดการที่ยั่งยืน ทั้งในด้านสิ่งแวดล้อม สังคม และธรรมาภิบาล (ESG) บริษัทถือเป็นหนึ่งในผู้นำของอุตสาหกรรมเหมืองแร่ทองคำ และได้รับความสนใจจากนักลงทุนทั่วโลก เนื่องจากราคาหุ้นของบริษัทมีความเชื่อมโยงโดยตรงกับราคาทองคำในตลาดโลก

งานวิจัยเรื่อง The Dynamics of Gold Prices, Gold Mining Stock Prices and Stock Market Prices Comovements (Gilmore et al., 2009) เป็นการศึกษาที่วิเคราะห์ความสัมพันธ์ระหว่างราคาทองคำ ราคาหุ้นของบริษัทเหมืองแร่ทองคำ และดัชนีตลาดหุ้นทั่วไป โดยมุ่งเน้นไปที่การเคลื่อนไหวร่วมกัน (Comovements) ของสินทรัพย์เหล่านี้ในบริบทของเศรษฐกิจมหภาคและตลาดการเงิน ผลการวิจัยพบว่าราคาทองคำมีความสัมพันธ์เชิงบวกอย่างชัดเจนกับราคาหุ้นของบริษัทเหมืองแร่ทองคำ เนื่องจากทองคำเป็นผลิตภัณฑ์หลักที่ส่งผลโดยตรงต่อรายได้ของบริษัทเหมืองแร่ อย่างไรก็ตาม ความสัมพันธ์ดังกล่าวอาจถูกปรับเปลี่ยนได้ด้วยปัจจัยอื่น เช่น ต้นทุนการผลิตและการบริหารจัดการของบริษัท นอกจากนี้ ความสัมพันธ์ระหว่างราคาทองคำและดัชนีตลาดหุ้นทั่วไปมีลักษณะที่แตกต่างออกไป โดยในช่วงที่ตลาดหุ้นอยู่ในภาวะผันผวนหรือเกิดวิกฤตเศรษฐกิจ ราคาทองคำมักจะปรับตัวสูงขึ้นในฐานะสินทรัพย์ปลอดภัย ซึ่งนักลงทุนใช้เพื่อป้องกันความเสี่ยงจากการลดลงของมูลค่าพอร์ตการลงทุน

ผลการศึกษาแสดงให้เห็นว่าราคาทองคำสามารถส่งผลกระทบโดยตรงต่อราคาหุ้นของบริษัทเหมืองแร่ทองคำ และมีบทบาทสำคัญในการกำหนดกลยุทธ์การลงทุนในตลาดหุ้นและสินค้าโภคภัณฑ์ งานวิจัยนี้จึงมีคุณค่าอย่างมากในการช่วยนักลงทุนและนักวิเคราะห์ในการทำความเข้าใจพลวัตของสินทรัพย์เหล่านี้ในบริบทของการบริหารความเสี่ยงและการสร้างพอร์ตการลงทุนที่หลากหลาย โดยเฉพาะในช่วงเวลาที่ตลาดการเงินมีความไม่แน่นอนสูง

ข้อมูลดัชนีหุ้นเอลโดราโด โกลด์ ถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลดัชนีหุ้นเอลโดราโด โกลด์ ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 4 ตารางแสดงโครงสร้างข้อมูลของดัชนีหุ้นเอลโดราโด โกลด์

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของดัชนีหุ้นเอลโดราโด โกลด์
2	Close/Last	ราคาปิดตลาดของดัชนีหุ้นเอลโดราโด โกลด์
3	Volume	ปริมาณการซื้อขายภายในวันของดัชนีหุ้นเอลโดราโด โกลด์
4	Open	ราคาเปิดตลาดของดัชนีหุ้นเอลโดราโด โกลด์
5	High	ราคาสูงสุดภายในวันของดัชนีหุ้นเอลโดราโด โกลด์
6	Low	ราคาต่ำสุดภายในวันของดัชนีหุ้นเอลโดราโด โกลด์

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลดัชนีหุ้นเอลโดราโด โกลด์ โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.5. **กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ** (VanEck Gold Miners ETF, GDX)

กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ เป็นกองทุนรวมดัชนีซื้อขายในตลาดหลักทรัพย์ (ETF) ที่มุ่งเน้นลงทุนในหุ้นของบริษัทเหมืองแร่ทองคำและบริษัทที่เกี่ยวข้องกับอุตสาหกรรมเหมืองทองคำทั่วโลก โดยมีเป้าหมายเพื่อให้ผลตอบแทนใกล้เคียงกับดัชนี NYSE Arca Gold Miners Index ซึ่งสะท้อนผลการดำเนินงานของบริษัทเหมืองทองคำขนาดใหญ่และขนาดกลาง กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ เหมาะสำหรับนักลงทุนที่ต้องการเก็งกำไร

จากราคาทองคำโดยไม่ต้องถือสินทรัพย์จริง และต้องการใช้เป็นเครื่องมือกระจายความเสี่ยงในพอร์ตการลงทุน ความผันผวนของกองทุนขึ้นอยู่กับราคาทองคำและผลประกอบการของบริษัทเหมืองทองคำ ทำให้กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำได้รับความนิยมในช่วงที่ราคาทองคำปรับตัวสูงขึ้น และยังสามารถซื้อขายในตลาดหลักทรัพย์ได้อย่างสะดวก เช่นเดียวกับหุ้นทั่วไป

ข้อมูลกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ ถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 5 ตารางแสดงโครงสร้างข้อมูลของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ
2	Close/Last	ราคาปิดตลาดของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ
3	Volume	ปริมาณการซื้อขายภายในวันของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ
4	Open	ราคาเปิดตลาดของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ
5	High	ราคาสูงสุดภายในวันของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ
6	Low	ราคาต่ำสุดภายในวันของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.6. **ดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์** (W&T Offshore Stock, WTI)

W&T Offshore, Inc. เป็นบริษัทพลังงานที่มุ่งเน้นการสำรวจ พัฒนา และผลิตน้ำมันดิบและก๊าซธรรมชาติ โดยมีพื้นที่ดำเนินงานหลักในบริเวณอ่าวเม็กซิโก ประเทศสหรัฐอเมริกา ซึ่งเป็นภูมิภาคสำคัญสำหรับอุตสาหกรรมน้ำมันและก๊าซธรรมชาติ บริษัทมีความ

เชี่ยวชาญด้านการจัดการทรัพยากรในพื้นที่ซับซ้อนและการใช้เทคโนโลยีขั้นสูงเพื่อเพิ่มประสิทธิภาพการผลิต หุ้นของ W&T Offshore มีชื่อย่อว่า WTI และจดทะเบียนในตลาดหลักทรัพย์นิวยอร์ก (NYSE) โดยมักถูกจับตามองในฐานะตัวแทนการลงทุนในภาคพลังงาน ความผันผวนของราคาสินค้าโภคภัณฑ์ เช่น น้ำมันและก๊าซธรรมชาติ ส่งผลโดยตรงต่อผลประกอบการของบริษัท แต่ก็เป็นโอกาสสำหรับนักลงทุนที่มองหาผลตอบแทนในช่วงราคาพลังงานขาขึ้น นอกจากนี้บริษัทยังเผชิญกับความท้าทายด้านสิ่งแวดล้อมและการเปลี่ยนแปลงของนโยบายพลังงานสะอาด แต่มีศักยภาพในการขยายแหล่งผลิตและสร้างมูลค่าเพิ่มผ่านการพัฒนาแหล่งทรัพยากรใหม่

งานวิจัย Oil and Gold Prices: Correlation or Causation? (Le & Chang, 2011) ศึกษาความสัมพันธ์ระหว่างราคาน้ำมันและราคาทองคำ โดยใช้วิธีการวิเคราะห์ทางสถิติและเชิงเศรษฐมิติ เพื่อประเมินว่า ความสัมพันธ์ดังกล่าวเป็นผลจากการเคลื่อนไหวที่สอดคล้องกันของตลาดหรือมีความเป็นเหตุและผลที่ชัดเจน ผลการวิจัยพบว่าราคาน้ำมันและราคาทองคำมีความสัมพันธ์เชิงบวกในระยะยาว ซึ่งสะท้อนถึงความเชื่อมโยงของตลาดสินค้าโภคภัณฑ์ในบริบทเศรษฐกิจโลก นอกจากนี้ ผลการวิเคราะห์เชิงสาเหตุด้วยวิธี Granger Causality Test พบว่าราคาน้ำมันมีแนวโน้มที่จะส่งผลต่อราคาทองคำในระยะสั้น ในขณะที่ในระยะยาว ความสัมพันธ์ระหว่างทั้งสองมีลักษณะของการตอบสนองร่วมกันต่อปัจจัยทางเศรษฐกิจมหภาค เช่น อัตราเงินเฟ้อ ค่าเงินดอลลาร์สหรัฐ และความไม่แน่นอนทางการเมือง

ผลการศึกษาชี้ให้เห็นว่าราคาน้ำมันสามารถใช้เป็นตัวชี้วัดเชิงนำ (Leading Indicator) สำหรับการเปลี่ยนแปลงของราคาทองคำในบางช่วงเวลา แต่ยังต้องคำนึงถึงปัจจัยอื่น ๆ ที่มีผลต่อทั้งสองตลาดอย่างซับซ้อน

ข้อมูลดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ ถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 6 ตารางแสดงโครงสร้างข้อมูลของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์
2	Close/Last	ราคาปิดตลาดของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์

ตาราง 6 (ต่อ) ตารางแสดงโครงสร้างข้อมูลของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
3	Volume	ปริมาณการซื้อขายภายในวันของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์
4	Open	ราคาเปิดตลาดของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์
5	High	ราคาสูงสุดภายในวันของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์
6	Low	ราคาต่ำสุดภายในวันของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.7. **กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา** (United States Brent Oil Fund, BNO)

กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา เป็นกองทุนซื้อขายแลกเปลี่ยน (Exchange-Traded Fund, ETF) ที่ลงทุนในสัญญาซื้อขายล่วงหน้าของน้ำมันดิบเบรนต์ ซึ่งเป็นประเภทน้ำมันที่ได้รับความนิยมในการอ้างอิงราคาน้ำมันในตลาดโลก โดยเฉพาะในตลาดยุโรป กองทุนนี้มีวัตถุประสงค์เพื่อให้ผู้ลงทุนสามารถเข้าถึงการเคลื่อนไหวของราคาน้ำมันดิบเบรนต์ได้ง่ายขึ้น โดยไม่ต้องซื้อขายสัญญาซื้อขายล่วงหน้าโดยตรงหรือจัดการกับการขนส่งน้ำมันดิบ กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา ลงทุนในสัญญาซื้อขายล่วงหน้า (Futures Contracts) ของน้ำมันดิบเบรนต์ที่จะหมดอายุในอนาคต ซึ่งมูลค่าของกองทุนจะสะท้อนถึงการเปลี่ยนแปลงของราคาน้ำมันดิบเบรนต์บนตลาดโลก นักลงทุนที่ซื้อหน่วยของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกาจะมีการเข้าถึงราคาน้ำมันดิบเบรนต์โดยผ่านการลงทุนในกองทุนนี้ ซึ่งทำให้สามารถกระจายการลงทุนได้โดยไม่ต้องมีการจัดการเองในส่วนของสินค้าฟิวเจอร์ส

ข้อมูลกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา ถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 7 ตารางแสดงโครงสร้างข้อมูลของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
1	Date	วันที่เก็บข้อมูลของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา
2	Close/Last	ราคาปิดตลาดของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา
3	Volume	ปริมาณการซื้อขายภายในวันของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา
4	Open	ราคาเปิดตลาดของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา
5	High	ราคาสูงสุดภายในวันของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา
6	Low	ราคาต่ำสุดภายในวันของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลกองทุนน้ำมัน-เบรนต์ของสหรัฐอเมริกา โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.8. **กองทุนรวมราคาน้ำมันสหรัฐอเมริกา** (United States Oil Fund, USO)

กองทุนรวมราคาน้ำมันสหรัฐอเมริกา เป็นกองทุนซื้อขายแลกเปลี่ยนที่ออกแบบมาเพื่อให้การลงทุนในน้ำมันดิบ West Texas Intermediate (WTI) ซึ่งเป็นเกณฑ์ราคาน้ำมันดิบที่ใช้ในตลาดสหรัฐอเมริกา ผ่านสัญญาซื้อขายล่วงหน้าที่เกี่ยวข้องกับน้ำมันดิบ West Texas Intermediate โดยผู้ลงทุนสามารถเข้าถึงการเปลี่ยนแปลงของราคาน้ำมันได้โดยการซื้อหุ้นของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา กองทุนนี้มักถูกใช้เป็นเครื่องมือสำหรับการเก็งกำไรเกี่ยวกับราคาน้ำมันหรือการป้องกันความเสี่ยงจากความผันผวนของราคาน้ำมันในตลาดโลก กองทุนรวมราคาน้ำมันสหรัฐอเมริกาจึงเหมาะสำหรับนักลงทุนที่ต้องการลงทุนในน้ำมันดิบโดยไม่ต้องซื้อขายสัญญาซื้อขายล่วงหน้าด้วยตัวเอง นอกจากนี้ กองทุนรวมราคาน้ำมันสหรัฐอเมริกา ยังได้รับความนิยมจากนักลงทุนที่ต้องการกระจายพอร์ตการลงทุนให้มีความหลากหลายผ่านการลงทุนในสินค้าโภคภัณฑ์ประเภทพลังงาน

ข้อมูลกองทุนรวมราคาน้ำมันสหรัฐอเมริกา ถูกรวบรวมโดย National Association of Securities Dealers Automated Quotations โดยข้อมูลข้อมูลกองทุนรวมราคาน้ำมันสหรัฐอเมริกา ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,516 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 8 ตารางแสดงโครงสร้างข้อมูลของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
1	Date	วันที่เก็บข้อมูลของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา
2	Close/Last	ราคาปิดตลาดของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา
3	Volume	ปริมาณการซื้อขายภายในวันของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา
4	Open	ราคาเปิดตลาดของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา
5	High	ราคาสูงสุดภายในวันของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา
6	Low	ราคาต่ำสุดภายในวันของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลกองทุนรวมราคาน้ำมันสหรัฐอเมริกา โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.9. **อัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว** (United States 10-Year Bond Yield)

อัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว เป็นอัตราผลตอบแทนจากพันธบัตรรัฐบาลสหรัฐ ที่มีระยะเวลาครบกำหนด 10 ปี ซึ่งเป็นตัวชี้วัดที่สำคัญในตลาดการเงินระหว่างประเทศ โดยพันธบัตรรัฐบาลสหรัฐ ถือเป็นสินทรัพย์ที่มีความเสี่ยงต่ำที่สุดในตลาดตราสารหนี้ และมักถูกใช้เป็นเกณฑ์ในการประเมินความเชื่อมั่นในเสถียรภาพทางเศรษฐกิจของสหรัฐเมริกา การเปลี่ยนแปลงในอัตราผลตอบแทนของพันธบัตรดังกล่าวสามารถสะท้อนถึงแนวโน้มทางเศรษฐกิจทั้งในระยะสั้นและระยะยาว รวมถึงการคาดการณ์อัตราดอกเบี้ยและนโยบายการเงินจากธนาคารกลางสหรัฐอเมริกา นอกจากนี้ยังมีบทบาทสำคัญในการกำหนดอัตราผลตอบแทนของสินทรัพย์อื่น ๆ ในตลาดการเงิน โดยเฉพาะในตราสารหนี้และตลาดทุน อีกทั้งยังเป็นตัวชี้วัดความเสี่ยงและความไม่แน่นอนทางเศรษฐกิจในช่วงเวลานั้น ๆ. เป็นตัวชี้วัดสำคัญในการประเมินคาดการณ์การเติบโตเศรษฐกิจและการดำเนินนโยบายการเงินของธนาคารกลางสหรัฐอเมริกา

งานวิจัย The Causal Relationship Between Stocks, Gold, Crude Oil, and Bond Returns in Poland (Mamcarz, 2024)มุ่งเน้นการวิเคราะห์ความสัมพันธ์ระหว่างผลตอบแทนจากหุ้น ทองคำ น้ำมันดิบ และพันธบัตรในตลาดการเงินของประเทศโปแลนด์ โดย

การศึกษานี้มีการพิจารณาความสัมพันธ์ทางสาเหตุระหว่างพันธบัตรกับราคาทองคำซึ่งเป็นหนึ่งในหัวข้อที่สำคัญในการทำความเข้าใจการเคลื่อนไหวของราคาสินทรัพย์ในตลาดการเงิน โดยจากการศึกษาพบว่ามีความสัมพันธ์ทางสาเหตุระหว่างผลตอบแทนจากพันธบัตรและราคาทองคำ โดยเมื่อผลตอบแทนจากพันธบัตรสูงขึ้น มักจะทำให้ราคาทองคำลดลง ความสัมพันธ์นี้สามารถอธิบายได้จากการที่นักลงทุนมีแนวโน้มที่จะหันไปลงทุนในพันธบัตรมากขึ้นเมื่ออัตราผลตอบแทนพันธบัตรสูงขึ้น ซึ่งอาจทำให้ความต้องการทองคำลดลง เนื่องจากการลงทุนในพันธบัตรให้ผลตอบแทนที่น่าสนใจในช่วงที่อัตราผลตอบแทนสูง ดังนั้นเมื่อพันธบัตรมีผลตอบแทนที่สูง การลงทุนในทองคำจึงอาจลดลงตามไปด้วย

ข้อมูลอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว ถูกรวบรวมโดย investing.com ซึ่งเป็นแพลตฟอร์มออนไลน์ที่ให้บริการข้อมูลและเครื่องมือด้านการเงินและการลงทุนอย่างครบวงจร โดยมีข้อมูลราคาสินทรัพย์ทางการเงินต่างๆ เช่น หุ้น, พันธบัตร, สกุลเงิน, สินค้าโภคภัณฑ์ และดัชนีตลาดหลักทรัพย์จากทั่วโลก เว็บไซต์นี้ให้ข้อมูลแบบเรียลไทม์พร้อมทั้งเครื่องมือการวิเคราะห์ทางเทคนิคที่สามารถช่วยนักลงทุนในการวิเคราะห์แนวโน้มของราคาสินทรัพย์ต่าง ๆ โดยใช้กราฟและตัวบ่งชี้ทางเทคนิคต่างๆ นอกจากนี้ยังมีข่าวสารและบทวิเคราะห์ทางเศรษฐกิจและการเงิน ซึ่งสามารถสนับสนุนการตัดสินใจลงทุนในตลาดการเงินได้อย่างมีประสิทธิภาพ เว็บไซต์นี้ได้รับการใช้งานอย่างแพร่หลายจากนักลงทุนทั้งในระดับมือใหม่และผู้ที่มีประสบการณ์ เพื่อเพิ่มประสิทธิภาพในการจัดการพอร์ตการลงทุนและการวางกลยุทธ์การลงทุนในสภาวะตลาดที่มีความผันผวน โดยข้อมูลอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,607 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 9 ตารางแสดงโครงสร้างข้อมูลของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว
2	Price	ราคาปิดตลาดของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว
3	Open	ราคาเปิดตลาดของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว

ตาราง 9 (ต่อ) ตารางแสดงโครงสร้างข้อมูลของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
4	High	ราคาสูงสุดภายในวันของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว
5	Low	ราคาต่ำสุดภายในวันของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว
6	Change%	อัตราการเปลี่ยนแปลงภายในวันของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับไม่มีค่าว่าง

3.2.1.10. **อัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ** (Euro US Dollar, EURUSD)

อัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐถือเป็นหนึ่งในอัตราแลกเปลี่ยนที่สำคัญที่สุดในระบบการเงินและการค้าระหว่างประเทศ เนื่องจากทั้งยูโรและดอลลาร์สหรัฐต่างเป็นสกุลเงินหลักที่ใช้ในระบบเศรษฐกิจโลก ยูโรเป็นสกุลเงินหลักของกลุ่มประเทศในสหภาพยุโรปที่ใช้ร่วมกันในโครงการสกุลเงินร่วมของยุโรป ซึ่งประกอบด้วยประเทศสมาชิกจำนวนมาก ส่วนดอลลาร์สหรัฐเป็นสกุลเงินหลักของสหรัฐอเมริกาและยังคงมีสถานะเป็นสกุลเงินสำรองหลักในระบบการเงินโลก ซึ่งการเปลี่ยนแปลงของอัตราแลกเปลี่ยนยูโรเป็นดอลลาร์สะท้อนถึงปัจจัยหลายประการที่มีอิทธิพลทั้งในระดับภายในและระดับโลก ปัจจัยสำคัญที่ส่งผลต่อการเคลื่อนไหวของอัตราแลกเปลี่ยนนี้ ได้แก่ นโยบายการเงินและการคลังของธนาคารกลางสหรัฐ และธนาคารกลางยุโรป (European Central Bank), สถานการณ์เศรษฐกิจภายในสหรัฐ และกลุ่มประเทศยูโรโซน, อัตราดอกเบี้ย, การคาดการณ์ทางเศรษฐกิจ รวมถึงเหตุการณ์ทางการเมืองและเศรษฐกิจที่มีผลกระทบต่อความเชื่อมั่นของนักลงทุน และในฐานะที่ราคาทองคำถูกกำหนดในสกุลเงินดอลลาร์สหรัฐ อัตราแลกเปลี่ยนระหว่างยูโรและดอลลาร์สหรัฐ จึงมีอิทธิพลโดยตรงต่อการเคลื่อนไหวของราคาทองคำ ความสัมพันธ์นี้สามารถอธิบายได้ผ่านมุมมองทางเศรษฐกิจที่เกี่ยวข้องกับการเปลี่ยนแปลงในมูลค่าของดอลลาร์สหรัฐ การเคลื่อนไหวของอัตราแลกเปลี่ยน

และปัจจัยที่เกี่ยวข้องกับความเสี่ยงทางเศรษฐกิจที่ส่งผลต่อความต้องการในสินทรัพย์ปลอดภัย เช่น ทองคำ

งานวิจัย The Relationship Between Crude Oil Prices, EUR/USD Exchange Rate, and Gold Prices (Houcine et al., 2020) ได้ทำการศึกษาความสัมพันธ์ระหว่างราคาน้ำมันดิบ, อัตราแลกเปลี่ยนระหว่างยูโรและดอลลาร์สหรัฐและราคาทองคำ โดยใช้การวิเคราะห์เชิงปริมาณเพื่อทดสอบความสัมพันธ์ทั้งในระยะสั้นและระยะยาวของปัจจัยเหล่านี้กับการเคลื่อนไหวของราคาทองคำ ซึ่งผลการวิจัยพบว่าอัตราแลกเปลี่ยนยูโรและดอลลาร์สหรัฐมีความสัมพันธ์ในทิศทางตรงข้ามกับราคาทองคำ กล่าวคือ เมื่อค่าเงินยูโรอ่อนค่าลงเมื่อเทียบกับดอลลาร์สหรัฐ ราคาทองคำมักจะเพิ่มสูงขึ้น เนื่องจากทองคำถูกซื้อขายในดอลลาร์สหรัฐ การอ่อนค่าของยูโรทำให้ราคาทองคำสูงขึ้นในตลาดโลก ซึ่งสามารถอธิบายได้ว่าเมื่อดอลลาร์สหรัฐแข็งค่าขึ้น หรือยูโรอ่อนค่าลง การซื้อขายทองคำในหน่วยดอลลาร์สหรัฐก็จะมีแนวโน้มสูงขึ้น เพราะทองคำถูกมองว่าเป็นสินทรัพย์ที่ปลอดภัยในช่วงที่เกิดความไม่แน่นอนทางเศรษฐกิจและการเงิน นอกจากนี้ การเคลื่อนไหวของอัตราแลกเปลี่ยนยูโรและดอลลาร์สหรัฐ จึงสามารถใช้เป็นตัวชี้วัดในการคาดการณ์ความผันผวนของราคาทองคำได้

ข้อมูลอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ ถูกรวบรวมโดย investing.com ซึ่งเป็นแพลตฟอร์มออนไลน์ที่ให้บริการข้อมูลและเครื่องมือด้านการเงินและการลงทุนอย่างครบวงจร โดยข้อมูลอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,609 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 10 ตารางแสดงโครงสร้างข้อมูลของอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ
2	Price	ราคาปิดตลาดของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ
3	Open	ราคาเปิดตลาดของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ
4	High	ราคาสูงสุดภายในวันของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ

ตาราง 10 (ต่อ) ตารางแสดงโครงสร้างข้อมูลของอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
5	Low	ราคาต่ำสุดภายในวันของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ
6	Vol	ปริมาณการซื้อขายภายในวันของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ
7	Change%	อัตราการเปลี่ยนแปลงภายในวันของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐ โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับมีคอลัมน์ปริมาณการซื้อขายภายในวันของอัตราแลกเปลี่ยนระหว่างสกุลเงินยูโรและดอลลาร์สหรัฐเป็นค่าว่างทั้งคอลัมน์ และในส่วนของคอลัมน์อื่น ๆ ไม่มีค่าว่าง

การกำจัดค่าว่างของคอลัมน์ปริมาณการซื้อขายใช้วิธีการกำจัดคอลัมน์ โดยใช้คำสั่ง drop(columns=['Vol']) เพื่อกำจัดคอลัมน์ปริมาณการซื้อขายเนื่องจากข้อมูลทั้งหมดเป็นค่าว่าง ไม่สามารถวิธีการทางคณิตศาสตร์ในการเติมข้อมูลได้ หลังจากทำการกำจัดคอลัมน์ปริมาณการซื้อขายแล้ว ทำให้ชุดข้อมูลไม่มีค่าว่าง

3.2.1.11. ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา (US Dollar Index)

ดัชนีค่าเงินดอลลาร์สหรัฐ (US Dollar Index, USDX) เป็นตัวชี้วัดความแข็งแกร่งของเงินดอลลาร์สหรัฐเมื่อเปรียบเทียบกับกลุ่มสกุลเงินหลัก 6 สกุล ได้แก่ ยูโร, เยนญี่ปุ่น, ปอนด์สเตอร์ลิงอังกฤษ, ดอลลาร์แคนาดา, โครนาสวีเดน และฟรังก์สวิส โดยดัชนีนี้สะท้อนสถานะทางเศรษฐกิจของสหรัฐอเมริกา และแนวโน้มของตลาดเงินตราโลก ความสัมพันธ์ระหว่าง ดัชนีค่าเงินดอลลาร์สหรัฐกับราคาทองคำมักเป็นไปในทิศทางตรงข้าม กล่าวคือ เมื่อดัชนีค่าเงินดอลลาร์สหรัฐเพิ่มขึ้น ราคาทองคำมักลดลง เนื่องจากดอลลาร์แข็งค่าทำให้ทองคำมีราคาสูงขึ้นในสายตาของผู้ซื้อที่ใช้สกุลเงินอื่น ในทางกลับกัน เมื่อ ดัชนีค่าเงินดอลลาร์สหรัฐลดลง ราคาทองคำมักปรับตัวเพิ่มขึ้น เนื่องจากดอลลาร์อ่อนค่าทำให้ทองคำมีราคาถูกลงสำหรับผู้ซื้อในต่างประเทศ

ความสัมพันธ์นี้สะท้อนบทบาทของทองคำในฐานะสินทรัพย์ปลอดภัยที่ได้รับความนิยมในช่วงที่เกิดความไม่แน่นอนทางเศรษฐกิจ

บทความ เจาะลึกราคาทองคำหลัง FED ลดดอกเบี้ย ของชาญณรงค์ มีชัยเจริญยิ่ง (มีชัยเจริญยิ่ง, 2024) ผู้ช่วยกรรมการผู้อำนวยการ ฝ่ายวิเคราะห์การลงทุนเพื่อบริหารความมั่งคั่ง บล.บัวหลวง กล่าวว่า การลดอัตราดอกเบี้ยโดยธนาคารกลางสหรัฐอเมริกา ส่งผลกระทบโดยตรงต่อราคาทองคำ ซึ่งถือเป็นสินทรัพย์ที่ได้รับความนิยมในฐานะตัวช่วยกระจายความเสี่ยงและรักษามูลค่าในช่วงที่เศรษฐกิจมีความผันผวน การลดดอกเบี้ยมีผลต่อการถือครองทองคำ เนื่องจากช่วยลดต้นทุนในการถือครองทองคำที่ไม่มีดอกเบี้ยจ่าย เช่นเดียวกับพันธบัตรหรือเงินฝาก ทั้งนี้ ยังส่งผลให้ค่าเงินดอลลาร์อ่อนค่าลง ซึ่งทำให้ทองคำที่กำหนดราคาในสกุลดอลลาร์สหรัฐถูกลงสำหรับนักลงทุนต่างชาติ และช่วยกระตุ้นความต้องการในตลาดโลก นอกจากนี้ ความสัมพันธ์ระหว่างค่าเงินดอลลาร์กับราคาทองคำมีลักษณะเชิงลบ กล่าวคือ เมื่อค่าเงินดอลลาร์อ่อนค่าลง ราคาทองคำมักจะปรับตัวสูงขึ้น โดยเฉพาะในช่วงที่เศรษฐกิจเผชิญความไม่แน่นอน เช่น ความกังวลเกี่ยวกับเงินเฟ้อหรือภาวะเศรษฐกิจถดถอย ปัจจัยเสริมอื่นๆ เช่น การสะสมทองคำของธนาคารกลางทั่วโลก และความต้องการในภาคอุตสาหกรรมและเครื่องประดับ ยังช่วยสนับสนุนให้ราคาทองคำเพิ่มสูงขึ้น

ข้อมูลดัชนีค่าเงินดอลลาร์สหรัฐ ถูกรวบรวมโดย investing.com ซึ่งเป็นแพลตฟอร์มออนไลน์ที่ให้บริการข้อมูลและเครื่องมือด้านการเงินและการลงทุน โดยข้อมูลดัชนีค่าเงินดอลลาร์สหรัฐ ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,608 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 11 ตารางแสดงโครงสร้างข้อมูลของดัชนีค่าเงินดอลลาร์สหรัฐ

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของดัชนีค่าเงินดอลลาร์สหรัฐ
2	Price	ราคาปิดตลาดของดัชนีค่าเงินดอลลาร์สหรัฐ
3	Open	ราคาเปิดตลาดของดัชนีค่าเงินดอลลาร์สหรัฐ
4	High	ราคาสูงสุดภายในวันของดัชนีค่าเงินดอลลาร์สหรัฐ
5	Low	ราคาต่ำสุดภายในวันของดัชนีค่าเงินดอลลาร์สหรัฐ
6	Vol	ปริมาณการซื้อขายภายในวันของดัชนีค่าเงินดอลลาร์สหรัฐ
7	Change%	อัตราการเปลี่ยนแปลงภายในวันของดัชนีค่าเงินดอลลาร์สหรัฐ

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลดัชนีค่าเงินดอลลาร์สหรัฐ โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับมีคอลัมน์ปริมาณการซื้อขายภายในวันของดัชนีค่าเงินดอลลาร์สหรัฐเป็นค่าว่างทั้งคอลัมน์ และในส่วนของคอลัมน์อื่น ๆ ไม่มีค่าว่าง

การกำจัดค่าว่างของคอลัมน์ปริมาณการซื้อขายใช้วิธีการกำจัดคอลัมน์ โดยใช้คำสั่ง drop(columns=['Vol']) เพื่อกำจัดคอลัมน์ปริมาณการซื้อขายเนื่องจากข้อมูลทั้งหมดเป็นค่าว่าง ไม่สามารถวิธีการทางคณิตศาสตร์ในการเติมข้อมูลได้ หลังจากทำการกำจัดคอลัมน์ปริมาณการซื้อขายแล้ว ทำให้ชุดข้อมูลไม่มีค่าว่าง

3.2.1.12. **สัญญาซื้อขายแร่เงินล่วงหน้า** (Silver Futures)

สัญญาซื้อขายแร่เงินล่วงหน้าเป็นตราสารทางการเงินที่ช่วยให้ผู้ซื้อและผู้ขายสามารถทำข้อตกลงล่วงหน้าในการซื้อหรือขายแร่เงินในราคาที่กำหนดไว้ล่วงหน้า โดยการส่งมอบสินทรัพย์จะเกิดขึ้นในช่วงเวลาที่ระบุในอนาคต สัญญาประเภทนี้ถูกใช้งานอย่างแพร่หลายเพื่อวัตถุประสงค์ในการป้องกันความเสี่ยงหรือเพื่อการเก็งกำไรจากความผันผวนของราคาสินทรัพย์ในตลาดโลหะมีค่า ในเชิงความสัมพันธ์ ราคาทองคำและราคาสัญญาซื้อขายแร่เงินล่วงหน้ามักมีแนวโน้มเคลื่อนไหวในทิศทางที่สอดคล้องกัน เนื่องจากทั้งทองคำและแร่เงินจัดอยู่ในกลุ่มโลหะมีค่าซึ่งได้รับอิทธิพลจากปัจจัยทางเศรษฐกิจและตลาดการเงินในลักษณะที่คล้ายคลึงกัน เช่น การเปลี่ยนแปลงของค่าเงิน การปรับตัวของอัตราดอกเบี้ย และความเสี่ยงทางเศรษฐกิจโลก ทั้งนี้แม้ว่าแร่เงินจะมีความผันผวนของราคาที่สูงกว่าทองคำในบางช่วงเวลา แต่ทั้งสองสินทรัพย์ยังคงแสดงความสัมพันธ์ในเชิงบวกอย่างมีนัยสำคัญ โดยเฉพาะในช่วงที่นักลงทุนหันมาใช้โลหะมีค่าเป็นสินทรัพย์ปลอดภัยเมื่อเผชิญกับความไม่แน่นอนทางเศรษฐกิจหรือการเงิน

งานวิจัย On the Long Run Relationship Between Gold and Silver Prices: A Note (Ciner, 2001) ได้ศึกษาความสัมพันธ์ระหว่างราคาทองคำและราคาเงินในระยะยาว โดยใช้ข้อมูลราคาทั้งสองในช่วงเวลาหลายทศวรรษ ผลการศึกษาชี้ให้เห็นว่า ราคาทองคำและราคาเงินมีความสัมพันธ์ในทางบวกในระยะยาว ซึ่งแสดงให้เห็นถึงการเชื่อมโยงระหว่างทั้งสองโลหะในตลาดการลงทุน แม้ว่าราคาเงินจะมีความผันผวนมากกว่าทองคำ แต่ทั้งสองยังคงมีความสัมพันธ์ที่มั่นคง โดยงานวิจัยนี้ยังได้เน้นถึงความสำคัญของการใช้ทองคำและเงินเป็นสินทรัพย์ในการลงทุนเพื่อการป้องกันความเสี่ยงและการเก็บมูลค่าทางการเงินในระยะยาว

ข้อมูลสัญญาซื้อขายแร่เงินล่วงหน้า ถูกรวบรวมโดย investing.com ซึ่งเป็นแพลตฟอร์มออนไลน์ที่ให้บริการข้อมูลและเครื่องมือด้านการเงินและการลงทุน โดยข้อมูลสัญญาซื้อขายแร่เงินล่วงหน้า ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 2,595 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 12 ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่เงินล่วงหน้า

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของสัญญาซื้อขายแร่เงินล่วงหน้า
2	Price	ราคาปิดตลาดของสัญญาซื้อขายแร่เงินล่วงหน้า
3	Open	ราคาเปิดตลาดของสัญญาซื้อขายแร่เงินล่วงหน้า
4	High	ราคาสูงสุดภายในวันสัญญาซื้อขายแร่เงินล่วงหน้า
5	Low	ราคาต่ำสุดภายในวันของสัญญาซื้อขายแร่เงินล่วงหน้า
6	Vol	ปริมาณการซื้อขายภายในวันของสัญญาซื้อขายแร่เงินล่วงหน้า
7	Change%	อัตราการเปลี่ยนแปลงภายในวันของสัญญาซื้อขายแร่เงินล่วงหน้า

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลสัญญาซื้อขายแร่เงินล่วงหน้า โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับมีคอลัมน์ปริมาณการซื้อขายภายในวันของสัญญาซื้อขายแร่เงินล่วงหน้าเป็นค่าว่างจำนวน 30 แถว และในส่วนของคอลัมน์อื่น ๆ ไม่มีค่าว่าง

การกำจัดค่าว่างของคอลัมน์ปริมาณการซื้อขายใช้วิธีการกำจัดคอลัมน์ โดยใช้คำสั่ง drop(columns=['Vol']) เพื่อกำจัดคอลัมน์ปริมาณการซื้อขายเนื่องจากข้อมูลปริมาณการซื้อขายของสัญญาซื้อขายแร่เงินล่วงหน้าไม่มีความสัมพันธ์กับราคาทองคำ หลังจากทำการกำจัดคอลัมน์ปริมาณการซื้อขายแล้ว ทำให้ชุดข้อมูลไม่มีค่าว่าง

3.2.1.13. **สัญญาซื้อขายแร่แพลทินัมล่วงหน้า** (Platinum Futures)

สัญญาซื้อขายแร่แพลทินัมล่วงหน้า คือ สัญญาที่ทำขึ้นระหว่างผู้ซื้อและผู้ขายเพื่อซื้อหรือขายแร่แพลทินัมในราคาที่ตกลงกันล่วงหน้า ณ วันที่กำหนดในอนาคต ซึ่งเป็นเครื่องมือทางการเงินที่ใช้ในการซื้อขายแร่แพลทินัมในตลาดฟิวเจอร์ส โดยมีการใช้ราคาตลาดในปัจจุบันเป็นฐานในการคำนวณราคาสัญญาล่วงหน้า ความสัมพันธ์ระหว่างสัญญาซื้อขายแร่แพลทินัมล่วงหน้าและราคาทองคำมีอยู่ในระดับหนึ่ง เนื่องจากทั้งทองคำและแพลทินัมเป็นโลหะมีค่าที่ใช้ใน

หลายอุตสาหกรรม เช่น อุตสาหกรรมการผลิตเครื่องประดับและอุตสาหกรรมยานยนต์ อีกทั้งยังเป็นสินทรัพย์ที่ใช้ในการเก็บมูลค่าในฐานะการลงทุน ดังนั้น เมื่อราคาทองคำมีการเปลี่ยนแปลงปัจจัยทางเศรษฐกิจที่มีผลต่อทองคำ เช่น อัตราดอกเบี้ยของธนาคารกลางหรือปัจจัยทางการเมืองมักจะส่งผลต่อการเคลื่อนไหวของราคาของโลหะมีค่าอื่น ๆ เช่น แพลทินัม โดยทั่วไปแล้ว ราคาของทองคำและแพลทินัมจะมีความสัมพันธ์ในทางบวกในบางช่วงเวลา เนื่องจากทั้งสองถูกมองว่าเป็นสินทรัพย์ที่ช่วยป้องกันความเสี่ยงในช่วงวิกฤติหรือความไม่แน่นอนทางเศรษฐกิจ

งานวิจัย Gold and Platinum: Toward Solving the Price Puzzle (Kearney & Lombra, 2009) มุ่งเน้นการวิเคราะห์ความสัมพันธ์ระหว่างราคาทองคำและราคาแพลทินัม พร้อมทั้งสำรวจปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงราคาของทั้งสองโลหะมีค่าดังกล่าวในตลาดโลก การศึกษาเน้นการแก้ไข "ปริศนาราคาทองคำ" ซึ่งหมายถึงการวิเคราะห์ว่าทำไมราคาทองคำมักจะแตกต่างจากแพลทินัม ทั้งที่ทั้งสองมีลักษณะทางเคมีและการใช้งานในอุตสาหกรรมบางประการที่คล้ายคลึงกัน งานวิจัยพบว่า ราคาของทองคำและแพลทินัมมีความสัมพันธ์กันในบางมิติ แต่ก็มีความแตกต่างกันอย่างชัดเจนในบางกรณี โดยเฉพาะในช่วงที่เกิดความไม่แน่นอนทางเศรษฐกิจหรือวิกฤตการณ์ที่ส่งผลกระทบต่อสินทรัพย์ที่มีความเสี่ยงต่ำ ทั้งนี้ การลงทุนในทองคำและแพลทินัมมีความสำคัญในแง่ของการป้องกันความเสี่ยง ซึ่งเป็นกลยุทธ์ที่นักลงทุนใช้ในการรักษามูลค่าของเงินลงทุนในช่วงที่สภาวะเศรษฐกิจไม่แน่นอน อีกประเด็นที่งานวิจัยนี้ได้นำเสนอคือการศึกษาความสัมพันธ์ระหว่างตลาดแพลทินัมและทองคำผ่านทางปัจจัยต่างๆ เช่น อุปสงค์ในอุตสาหกรรม และอัตราแลกเปลี่ยนสกุลเงิน ซึ่งทั้งหมดนี้ช่วยให้เข้าใจลักษณะของราคาของทั้งสองโลหะมีค่าในระยะยาว

ข้อมูลสัญญาซื้อขายแร่แพลทินัมล่วงหน้า ถูกรวบรวมโดย investing.com ซึ่งเป็นแพลตฟอร์มออนไลน์ที่ให้บริการข้อมูลและเครื่องมือด้านการเงินและการลงทุน โดยข้อมูลสัญญาซื้อขายแร่แพลทินัมล่วงหน้า ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 3,092 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 13 ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า
2	Price	ราคาปิดตลาดของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า

ตาราง 13 (ต่อ) ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า

ลำดับ	ชื่อตัวแปร (Variable)	คำอธิบาย (Description)
3	Open	ราคาเปิดตลาดของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า
4	High	ราคาสูงสุดภายในวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า
5	Low	ราคาต่ำสุดภายในวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า
6	Vol	ปริมาณการซื้อขายภายในวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า
7	Change%	อัตราการเปลี่ยนแปลงภายในวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลสัญญาซื้อขายแร่แพลทินัมล่วงหน้าโดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับมีคอลัมน์ปริมาณการซื้อขายภายในวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้าเป็นค่าว่างจำนวน 1,998 แถว และในส่วนของคอลัมน์อื่น ๆ ไม่มีค่าว่าง

การกำจัดค่าว่างของคอลัมน์ปริมาณการซื้อขายใช้วิธีการกำจัดคอลัมน์ โดยใช้คำสั่ง drop(columns=['Vol']) เพื่อกำจัดคอลัมน์ปริมาณการซื้อขายเนื่องจากข้อมูลปริมาณการซื้อขายของสัญญาซื้อขายแร่แพลทินัมล่วงหน้าไม่มีความสัมพันธ์กับราคาทองคำ หลังจากทำการกำจัดคอลัมน์ปริมาณการซื้อขายแล้ว ทำให้ชุดข้อมูลไม่มีค่าว่าง

3.2.1.14. **สัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า** (Palladium Futures)

สัญญาซื้อขายแร่พัลลาเดียมล่วงหน้าคือสัญญาทางการเงินที่ให้สิทธิ์แก่ผู้ซื้อในการซื้อแร่พัลลาเดียมในราคาที่ตกลงกันล่วงหน้าสำหรับการส่งมอบในอนาคต โดยการซื้อขายผ่านตลาดฟิวเจอร์ส ซึ่งการซื้อขายฟิวเจอร์สนี้สามารถใช้เป็นเครื่องมือในการเก็งกำไรหรือป้องกันความเสี่ยงจากการเปลี่ยนแปลงของราคาพัลลาเดียมในตลาดโลก โดยพัลลาเดียมเป็นโลหะมีค่าชนิดหนึ่งที่มีการใช้ในอุตสาหกรรมมากมาย โดยเฉพาะในอุตสาหกรรมยานยนต์ ซึ่งถูกใช้ในการผลิตตัวกรองไอเสียที่ช่วยลดมลพิษจากการเผาไหม้ของเครื่องยนต์ นอกจากนี้ยังมีการใช้ในอุตสาหกรรมเครื่องประดับและอิเล็กทรอนิกส์อีกด้วย ราคาพัลลาเดียมจึงมีความผันผวนจากปัจจัยต่าง ๆ เช่น ความต้องการในอุตสาหกรรมยานยนต์ การเปลี่ยนแปลงทางเศรษฐกิจทั่วโลก และอุปทานจากแหล่งผลิตหลักเช่น รัสเซียและแอฟริกาใต้ ซึ่งความสัมพันธ์ระหว่างราคาพัลลาเดียมและราคา

ทองคำมีความน่าสนใจ เนื่องจากทั้งสองเป็นโลหะมีค่าและมีการซื้อขายในตลาดที่คล้ายกัน อย่างไรก็ตาม การเคลื่อนไหวของราคาของพัลลาเดียมและทองคำมักไม่ได้เคลื่อนไหวในทิศทางเดียวกันเสมอไป ในบางช่วงที่ตลาดทองคำมีความผันผวนสูง เช่น ในช่วงของวิกฤตเศรษฐกิจหรือตลาดหุ้นตกต่ำ ราคาทองคำอาจปรับตัวสูงขึ้นเพราะนักลงทุนมองว่าเป็นแหล่งเก็บมูลค่าที่ปลอดภัย ขณะที่ราคาพัลลาเดียมอาจไม่ได้รับผลกระทบในลักษณะเดียวกัน เนื่องจากราคาพัลลาเดียมมักจะได้รับผลกระทบจากปัจจัยด้านอุปสงค์ในอุตสาหกรรมมากกว่า นอกจากนี้ราคาพัลลาเดียมและทองคำยังสามารถมีความสัมพันธ์ทางบวกหรือทางลบขึ้นอยู่กับสถานการณ์ทางเศรษฐกิจและการเงินในแต่ละช่วงเวลา ในบางช่วง ราคาของทั้งสองอาจเคลื่อนไหวไปในทิศทางเดียวกัน เช่น ในกรณีที่มีความไม่แน่นอนทางเศรษฐกิจสูง และนักลงทุนต้องการกระจายความเสี่ยงด้วยการลงทุนในสินทรัพย์มีค่าหลายประเภท ขณะที่ในบางช่วงที่มีการเปลี่ยนแปลงอุปทานหรือความต้องการในตลาดอุตสาหกรรม ราคาพัลลาเดียมอาจได้รับผลกระทบอย่างรุนแรง ทำให้เกิดความแตกต่างจากราคาทองคำ

งานวิจัย Short-term and long-term relationships between gold prices and precious metal (palladium, silver and platinum) and energy (crude oil and gasoline) prices (Eryigit, 2017) ได้ศึกษาความสัมพันธ์ระหว่างราคาทองคำกับโลหะมีค่าอื่นๆ เช่น พัลลาเดียม รวมถึงราคาพลังงาน เช่น น้ำมันดิบ และน้ำมันเบนซิน ในระยะสั้นและระยะยาว โดยเฉพาะในส่วนของความสัมพันธ์ระหว่างราคาทองคำและราคาพัลลาเดียม ในส่วนของความสัมพันธ์ระหว่างราคาทองคำและราคาพัลลาเดียม ผลการศึกษาแสดงให้เห็นว่า ทั้งสองโลหะมีค่ามีความสัมพันธ์ในระยะสั้นและระยะยาว ซึ่งสัมพันธ์กันในบางช่วงเวลา แต่ก็มีความแตกต่างในปัจจัยที่ส่งผลกระทบต่อราคา ราคาพัลลาเดียมและทองคำต่างได้รับอิทธิพลจากปัจจัยด้านอุปทานและอุปสงค์ในอุตสาหกรรมที่ต่างกัน โดยพัลลาเดียมมีการใช้ที่สำคัญในอุตสาหกรรมยานยนต์และอิเล็กทรอนิกส์ ขณะที่ทองคำได้รับความนิยมจากการเก็บรักษามูลค่าและการลงทุนในสินทรัพย์ที่ปลอดภัยในช่วงที่เกิดความไม่แน่นอนทางเศรษฐกิจ ในระยะสั้น ราคาของทั้งสองอาจมีความสัมพันธ์ในทิศทางเดียวกันหรือสวนทางกัน ขึ้นอยู่กับปัจจัยต่างๆ เช่น ความไม่แน่นอนในตลาดการเงินหรือการเปลี่ยนแปลงในอุปทานของโลหะมีค่าทั้งสอง อย่างไรก็ตาม ในระยะยาว ความสัมพันธ์ระหว่างราคาทองคำและราคาพัลลาเดียมมีแนวโน้มที่จะแตกต่างกันบ้างเนื่องจากทั้งสองโลหะมีลักษณะการใช้งานและการตอบสนองต่อภาวะเศรษฐกิจที่แตกต่างกัน

ข้อมูลสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า ถูกรวบรวมโดย investing.com ซึ่งเป็นแพลตฟอร์มออนไลน์ที่ให้บริการข้อมูลและเครื่องมือด้านการเงินและการลงทุน โดยข้อมูล

สัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า ถูกรวบรวมข้อมูลตั้งแต่วันที่ 5 มกราคม 2015 ถึง 2 มกราคม 2025 รวมทั้งสิ้น 3,092 แถวในรูปแบบ Comma-Separated Values (CSV)

ตาราง 14 ตารางแสดงโครงสร้างข้อมูลของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า

ลำดับ	**ชื่อตัวแปร** (Variable)	**คำอธิบาย** (Description)
1	Date	วันที่เก็บข้อมูลของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า
2	Price	ราคาปิดตลาดของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า
3	Open	ราคาเปิดตลาดของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า
4	High	ราคาสูงสุดภายในวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า
5	Low	ราคาต่ำสุดภายในวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า
6	Vol	ปริมาณการซื้อขายภายในวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า
7	Change%	อัตราการเปลี่ยนแปลงภายในวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า

เมื่อทำการตรวจสอบความถูกต้องและความครบถ้วนของข้อมูลสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า โดยใช้คำสั่ง isnull().sum() เพื่อนับจำนวนค่าว่างของแต่ละคอลัมน์ของข้อมูล พบว่าข้อมูลที่ได้รับมีคอลัมน์ปริมาณการซื้อขายภายในวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้าเป็นค่าว่างจำนวน 1,998 แถว และในส่วนของคอลัมน์อื่น ๆ ไม่มีค่าว่าง

การกำจัดค่าว่างของคอลัมน์ปริมาณการซื้อขายใช้วิธีการกำจัดคอลัมน์ โดยใช้คำสั่ง drop(columns=['Vol']) เพื่อกำจัดคอลัมน์ปริมาณการซื้อขายเนื่องจากข้อมูลปริมาณการซื้อขายของสัญญาซื้อขายแร่แพลทินัมล่วงหน้าไม่มีความสัมพันธ์กับราคาทองคำ หลังจากทำการกำจัดคอลัมน์ปริมาณการซื้อขายแล้ว ทำให้ชุดข้อมูลไม่มีค่าว่าง

3.2.2. **การจัดเก็บข้อมูล** (Data Storage)

ข้อมูลที่ได้รับมาจากหลายหลายแหล่งข้อมูลซึ่งทำให้ข้อมูลไม่มีความสอดคล้องกันและการเตรียมข้อมูลเพื่อใช้ในการเรียนรู้ของเครื่อง จำเป็นต้องรวมข้อมูลจากหลากหลายนี้เข้าด้วยกัน ในส่วนของการรวมข้อมูลนี้ ชื่อคอลัมน์เป็นสิ่งสำคัญที่ช่วยให้สามารถแยกแยะได้ว่าข้อมูลใดมีความสัมพันธ์กับราคาทองคำได้ ซึ่งข้อมูลที่ได้รับมาทุกชุดข้อมูลมีชื่อคอลัมน์ที่เหมือนกัน เช่น ราคาเปิด ใช้ชื่อคอลัมน์ว่า Open เป็นต้น ซึ่งหากรวมข้อมูลเข้าด้วยกันแล้ว จะทำให้ไม่สามารถ

รับรู้ได้ว่าข้อมูลราคาเปิดนี้เป็นราคาของข้อมูลใด การกำหนดชื่อคอลัมน์ของแต่ละข้อมูล ใช้วิธีการกำหนดชื่อแทนของแต่ละชุดข้อมูล เช่น ข้อมูลกองทุนทองคำ ใช้ชื่อแทนว่า “GLD” โดยชื่อคอลัมน์ในชุดข้อมูลกองทุนทองคำจะใช้ชื่อแทน ตามด้วยเครื่องหมายขีดใต้ (_) และชื่อคอลัม เช่น ชื่อราคาปิดของกองทุนทองคำ จะใช้ว่า “GLD_Open” เป็นต้น โดยการกำหนดชื่อแทนจะอ้างอิงจากตาราง 15 ตารางแสงดการใช้ชื่อแทนของแต่ละชุดข้อมูล

ตาราง 15 ตารางแสดงการใช้ชื่อแทนของแต่ละชุดข้อมูล

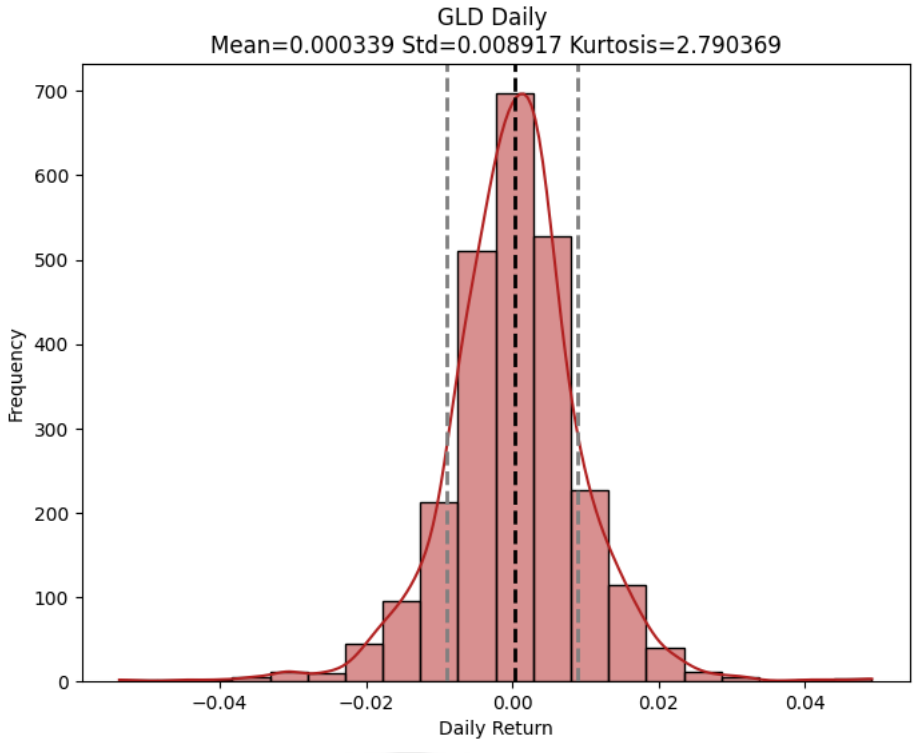
ลำดับ	ชื่อแทน	ชื่อชุดข้อมูล
1	GLD	กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ
2	SPX	ดัชนีตลาดหุ้นสหรัฐอเมริกา
3	DJI	ดัชนีดาวโจนส์
4	EGO	ดัชนีหุ้นเอลโดราโด โกลด์
5	GDX	กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ
6	WTI	ดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์
7	BNO	กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา
8	USO	กองทุนรวมราคาน้ำมันสหรัฐอเมริกา
9	USB	อัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว
10	EUR	อัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ
11	USD	ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา
12	SIL	สัญญาซื้อขายแร่เงินล่วงหน้า
13	PLT	สัญญาซื้อขายแร่แพลทินัมล่วงหน้า
14	PLD	สัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า

การรวมข้อมูลใช้วิธีการรวมข้อมูลโดยใช้วันที่เป็นคีย์ในการรวมข้อมูล และรวมข้อมูลแบบภายใน (Inner) ซึ่งเมื่อทำการรวมข้อมูลแล้วจะทำให้ข้อมูลที่มีวันที่ไม่ตรงกันหายไป โดยข้อมูลมีจำนวน 2,512 แถว และมีคอลัมน์ 68 คอลัมน์ ซึ่งข้อมูลหลังจากการรวมแล้วไม่พบค่าว่าง

3.2.3. **การสำรวจข้อมูล** (Data Exploratory)

3.2.3.1. **การพิจารณาคุณลักษณะเดี่ยว** (Univariant Data)

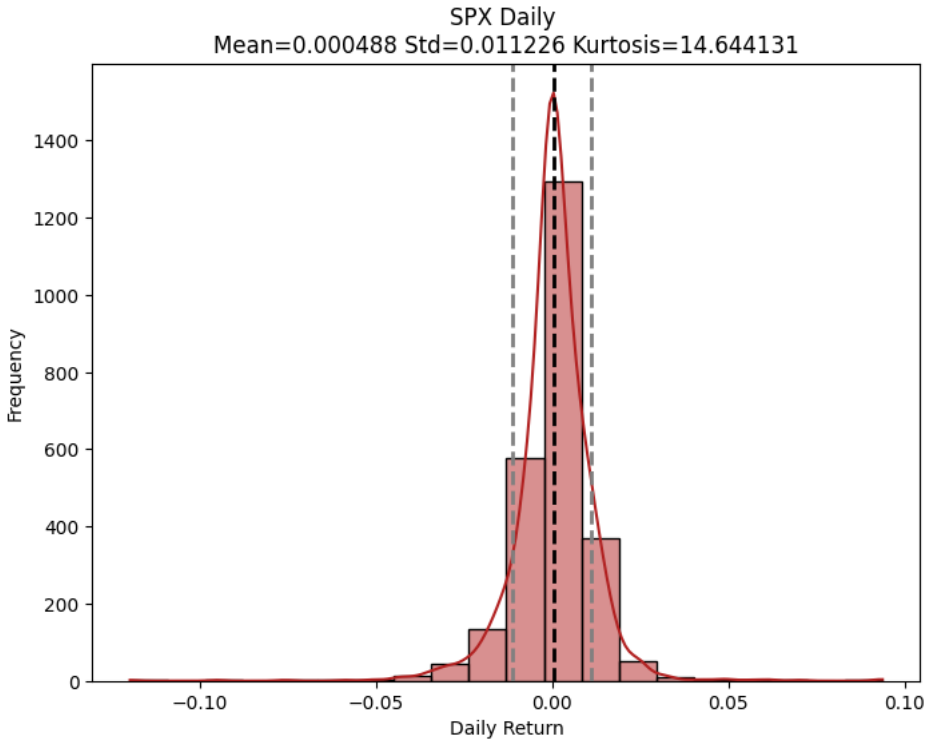
การพิจารณาตามคุณลักษณะจะใช้ราคาปิดของแต่ละกลุ่มข้อมูลเป็นตัวแทนของกลุ่มข้อมูลในการพิจารณาคุณลักษณะ โดยใช้ค่าอัตราการเปลี่ยนแปลงระหว่างวัน ซึ่งคำนวณจากส่วนต่างของข้อมูลระหว่างวัน ซึ่งการพิจารณาอัตราการเปลี่ยนแปลงระหว่างวัน สามารถทำให้ทราบถึงแนวโน้มการเปลี่ยนแปลงของข้อมูลในแต่ละช่วงเวลาได้ โดยเมื่อนำอัตราการเปลี่ยนแปลงระหว่างวัน มาหาค่าทางสถิติ โดยการใช้ค่าเฉลี่ย (Mean), ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) และค่าความโด่ง (Kurtosis)



ภาพประกอบ 9 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ

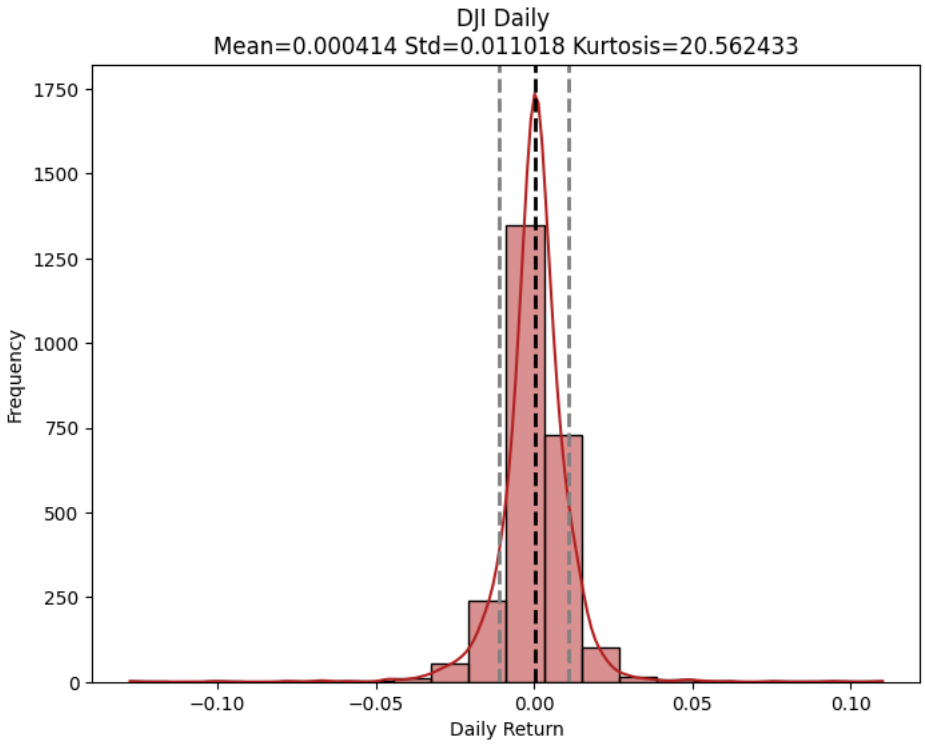
จากภาพประกอบที่ 9 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำพบว่ามีค่าเฉลี่ยอยู่ที่ 0.000339 แสดงที่เส้นปะสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.008917 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างน้อย จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบน

มาตรฐาน และค่าความโด่งอยู่ที่ 2.790369 ซึ่งใกล้เคียงกับการแจกแจงปกติ (Mesokurtic Distribution)



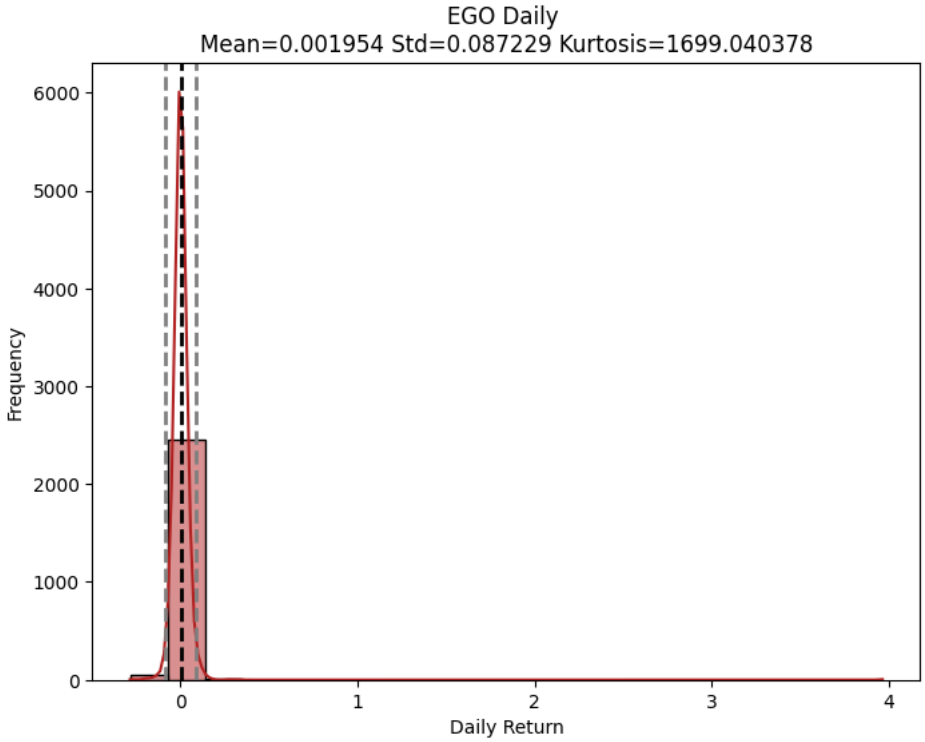
ภาพประกอบ 10 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีตลาดหุ้นสหรัฐอเมริกา

จากภาพประกอบที่ 10 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีตลาดหุ้นสหรัฐอเมริกาพบว่ามีค่าเฉลี่ยอยู่ที่ 0.000488 แสดงที่เส้นประสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.011226 แสดงที่เส้นประสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างน้อย จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน และค่าความโด่งอยู่ที่ 14.644131 ซึ่งมีค่าข้อนข้างสูง เป็นการกระจายตัวที่มียอดแหลม (Leptokurtic Distribution) อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



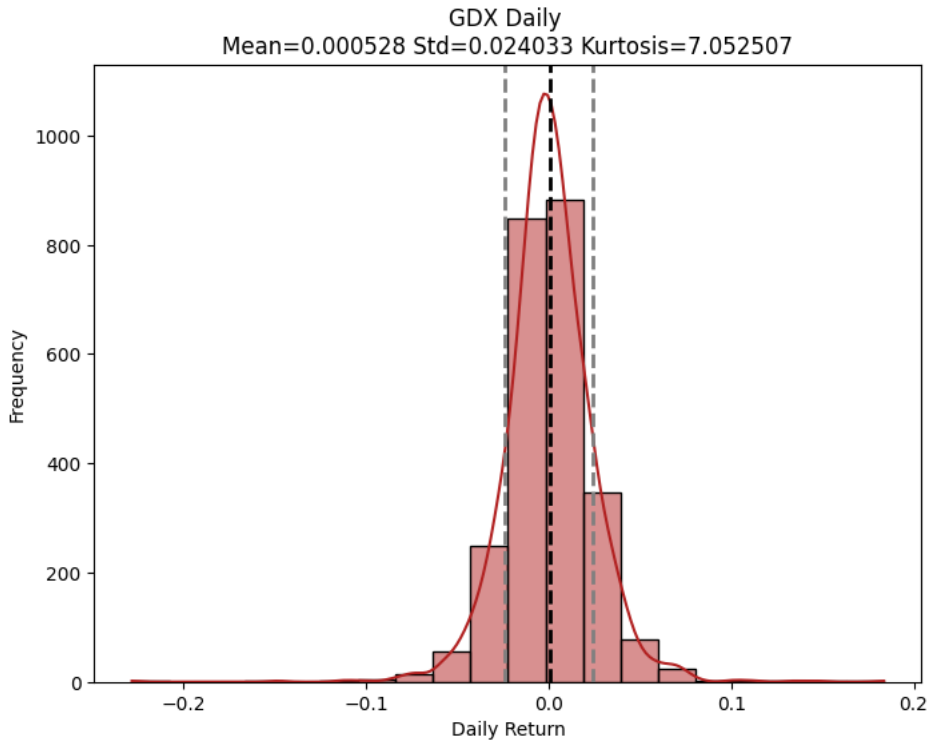
ภาพประกอบ 11 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีดาวโจนส์

จากภาพประกอบที่ 11 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีดาวโจนส์ พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000414 แสดงที่เส้นประสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.011018 แสดงที่เส้นประสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างน้อย จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน และค่าความโด่งอยู่ที่ 20.562433 ซึ่งมีค่าข้อนข้างสูง เป็นการกระจายตัวที่มียอดแหลมได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



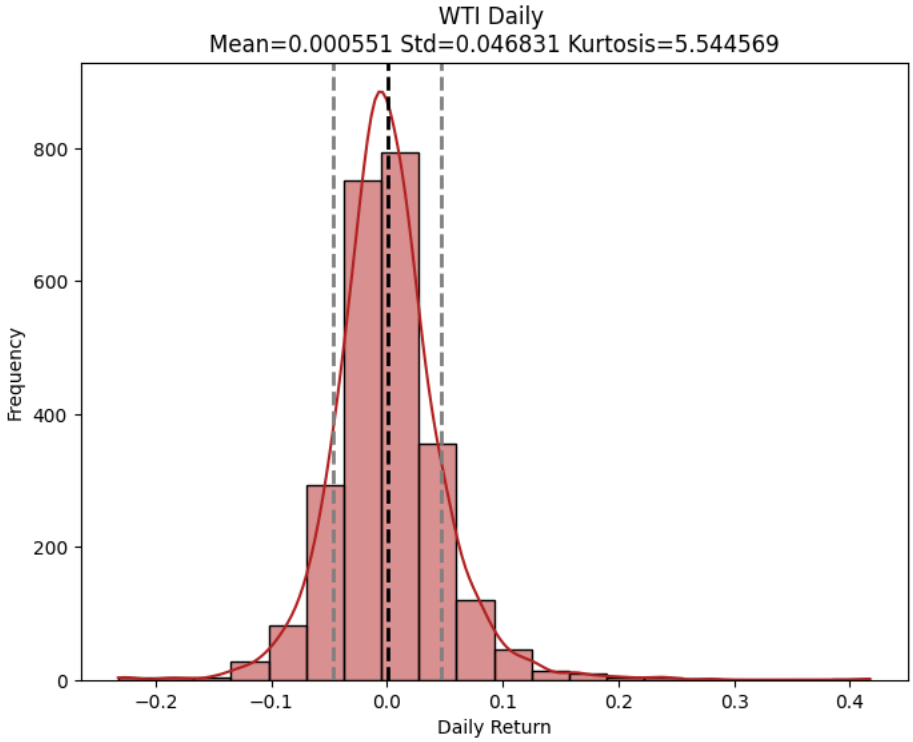
ภาพประกอบ 12 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีหุ้นเอลโดราโด โกลด์

จากภาพประกอบที่ 12 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีหุ้นเอลโดราโด โกลด์ (EGO) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.001954 แสดงที่เส้นปะสีดำ มีค่าที่เป็นบวก ซึ่งเป็นสัญญาณเชิงบวกต่อการลงทุน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.087229 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างสูง จากข้อมูลมีส่วนของข้อมูลเกือบครึ่งหนึ่งหลุดจากเส้นส่วนเบี่ยงเบนมาตรฐานในทิศทางบวก ซึ่งแสดงถึงความผันผวนของราคาที่ค่อนข้างสูง และค่าความโด่งอยู่ที่ 1699.040378 ซึ่งมีค่าสูงมาก เป็นการกระจายตัวที่มียอดแหลม อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมาก ซึ่งอาจหมายถึงการเพิ่มขึ้นหรือลดลงของราคาที่รุนแรงผิดปกติ โดยหางของการกระจายตัวมีความยาวมาก บ่งชี้ว่าผลตอบแทนบางค่าอยู่ในระดับที่แตกต่างจากค่าเฉลี่ยอย่างมาก



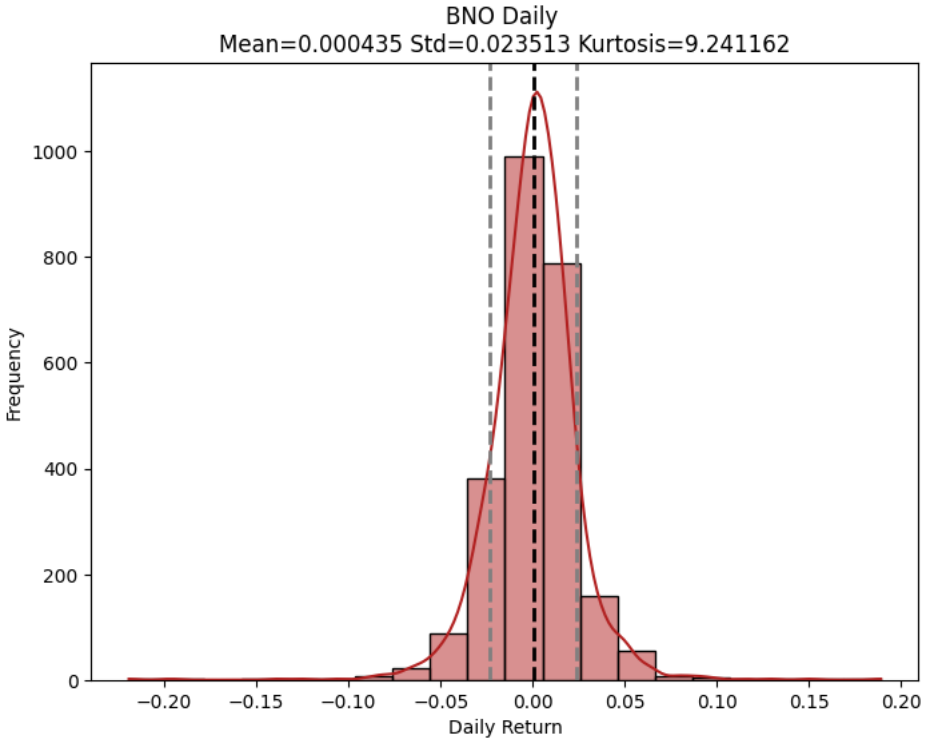
ภาพประกอบ 13 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ

จากภาพประกอบที่ 13 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ (GDX) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000528 แสดงที่เส้นปะสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.024033 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างต่ำถึงปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่งอยู่ที่ 7.052501 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติมากนัก แต่มีการกระจายตัวที่มียอดแหลม โดยหางของการกระจายตัวมีความยาวมากกว่าการกระจายตัว อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



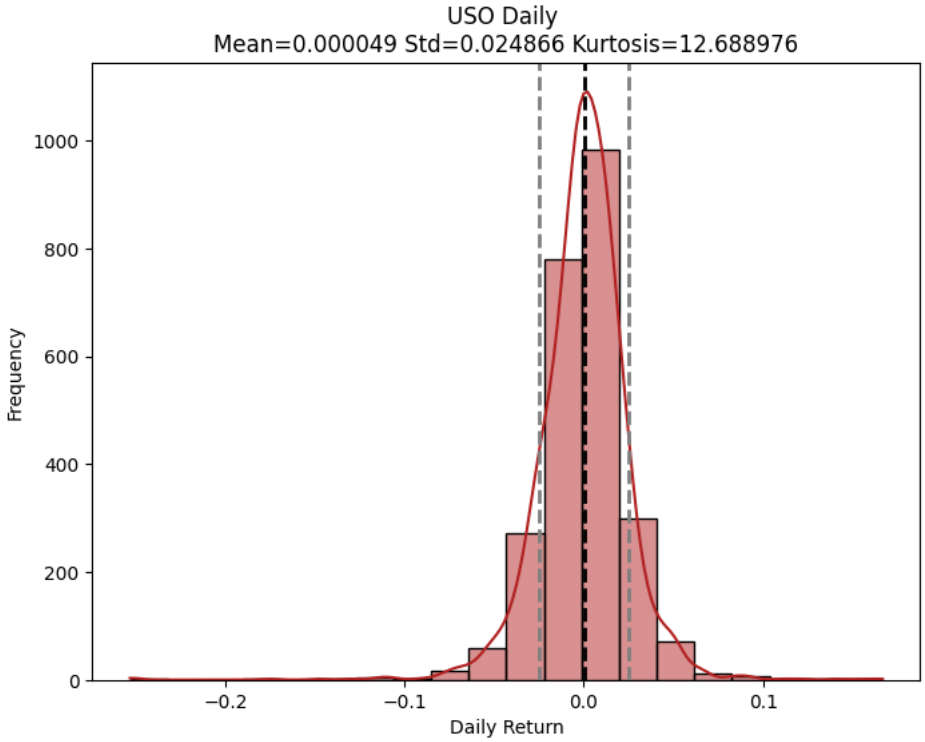
ภาพประกอบ 14 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์

จากภาพประกอบที่ 14 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ (WTI) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000551 แสดงที่เส้นปะสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.046831 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างต่ำปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่ง อยู่ที่ 5.544569 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติมากนัก แต่มีการกระจายตัวที่มียอดแหลมโดยหางของการกระจายตัวมีความยาวมากกว่าการกระจายตัว อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



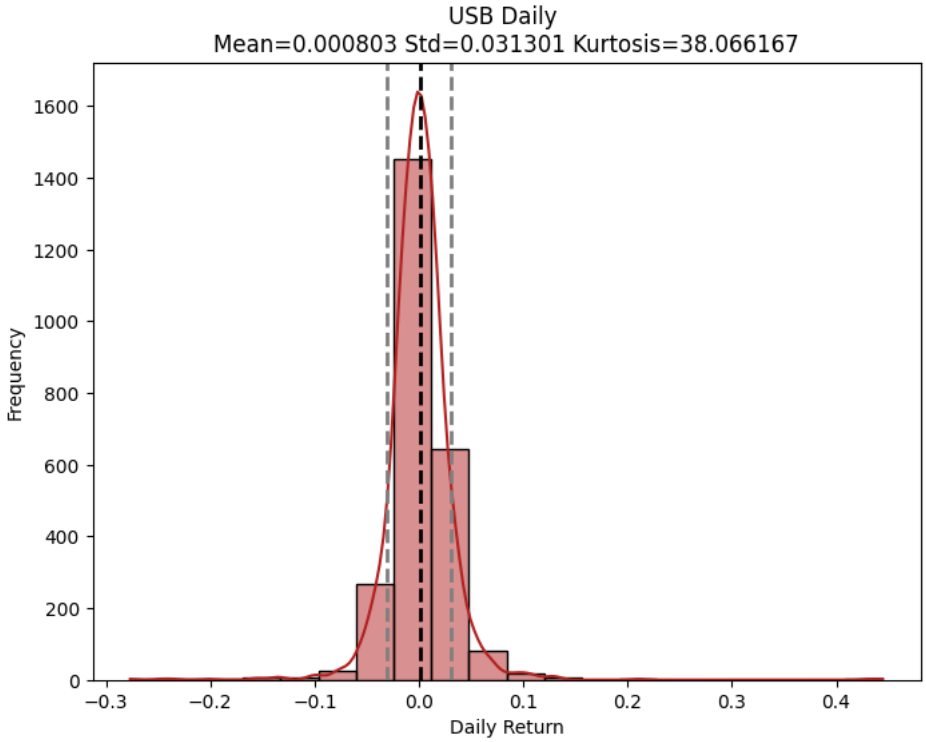
ภาพประกอบ 15 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา

จากภาพประกอบที่ 15 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา (BNO) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000435 แสดงที่เส้นปะสีดำซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.023513 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่ง อยู่ที่ 9.241162 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติมากนัก แต่มีการกระจายตัวที่มียอดแหลมโดยหางของการกระจายตัวมีความยาวมากกว่าการกระจายตัว อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



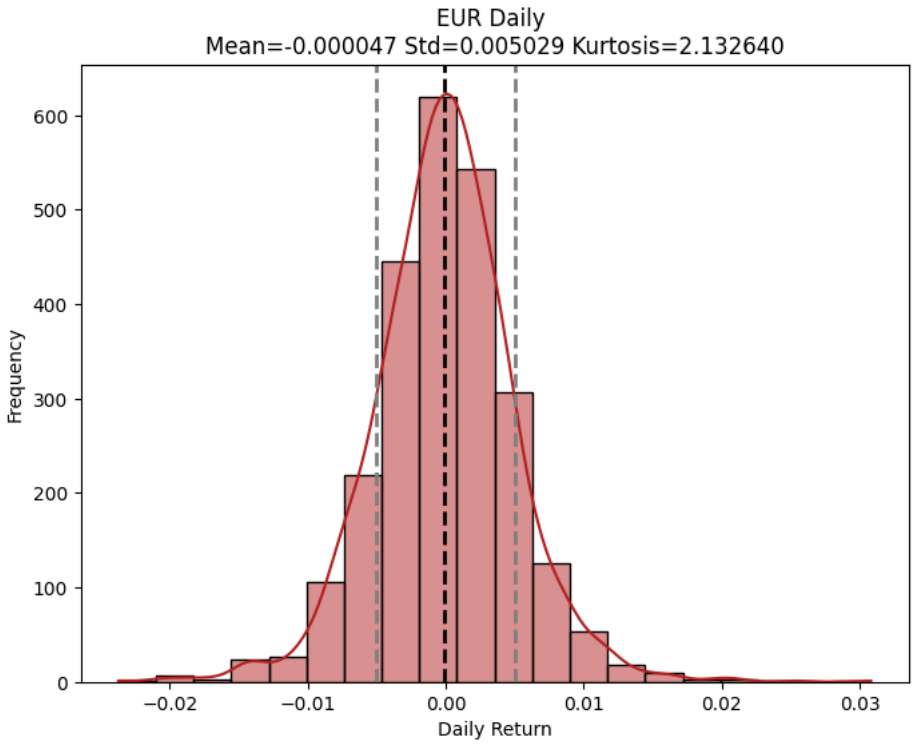
ภาพประกอบ 16 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา

จากภาพประกอบที่ 16 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของกองทุนรวมราคาน้ำมันสหรัฐอเมริกา (USO) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000049 แสดงที่เส้นปะสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.024866 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่ง อยู่ที่ 12.688976 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติมากนัก แต่มีการกระจายตัวที่มียอดแหลมโดยหางของการกระจายตัวมีความยาวมากกว่าการกระจายตัว อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



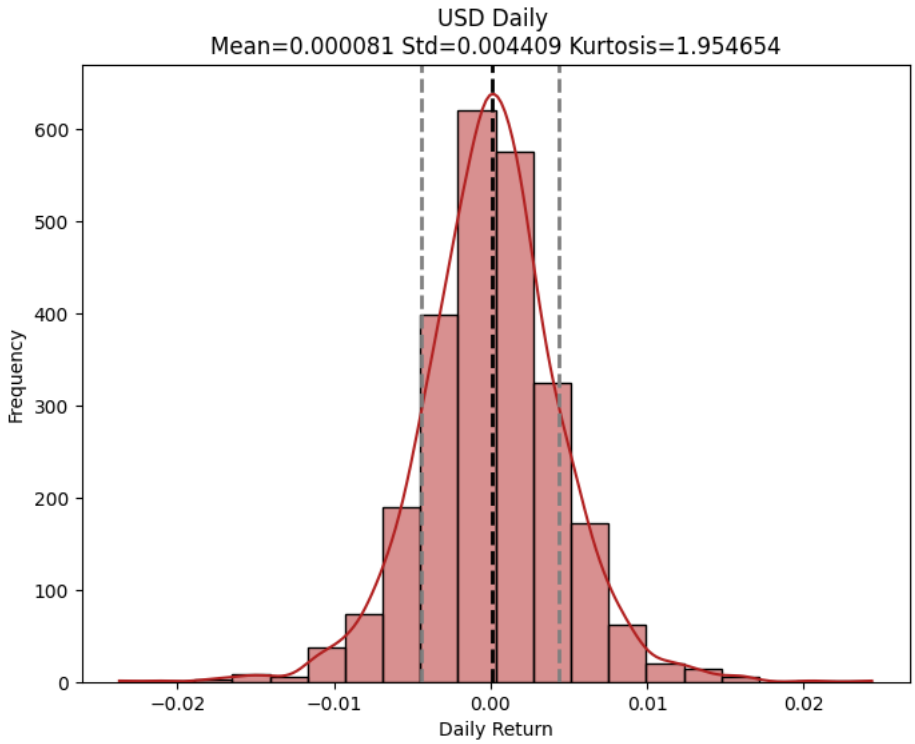
ภาพประกอบ 17 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว

จากภาพประกอบที่ 17 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว (USB) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000803 แสดงที่เส้นปะสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.031301 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างสูง จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายจนเกินขอบของส่วนเบี่ยงเบนมาตรฐานในด้านบวก และค่าความโด่งอยู่ที่ 38.066167 ซึ่งมีค่าข้อนข้างสูงงจากการแจกแจงปกติและมีการกระจายตัวที่มียอดแหลมโดยหางของการกระจายตัวมีความยาวมากกว่าการกระจายตัว อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ค่อนข้างบ่อย



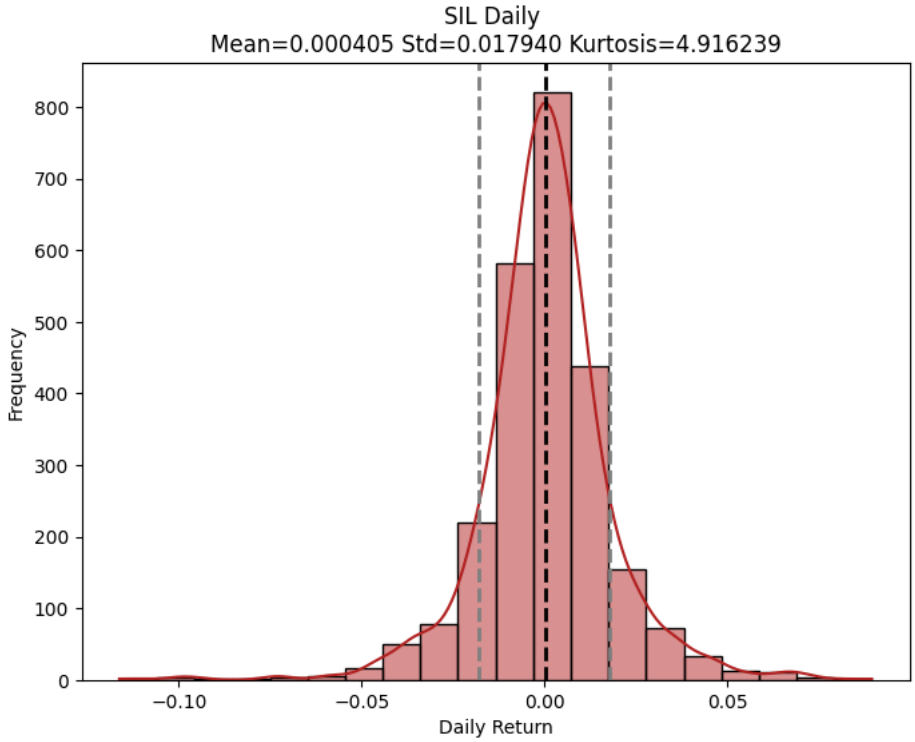
ภาพประกอบ 18 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ

จากภาพประกอบที่ 18 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ (EUR) พบว่ามีค่าเฉลี่ยอยู่ที่ -0.000047 แสดงที่เส้นประสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.005029 แสดงที่เส้นประสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างน้อย จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน และค่าความโด่งอยู่ที่ 2.132640 ซึ่งใกล้เคียงกับการแจกแจงปกติ



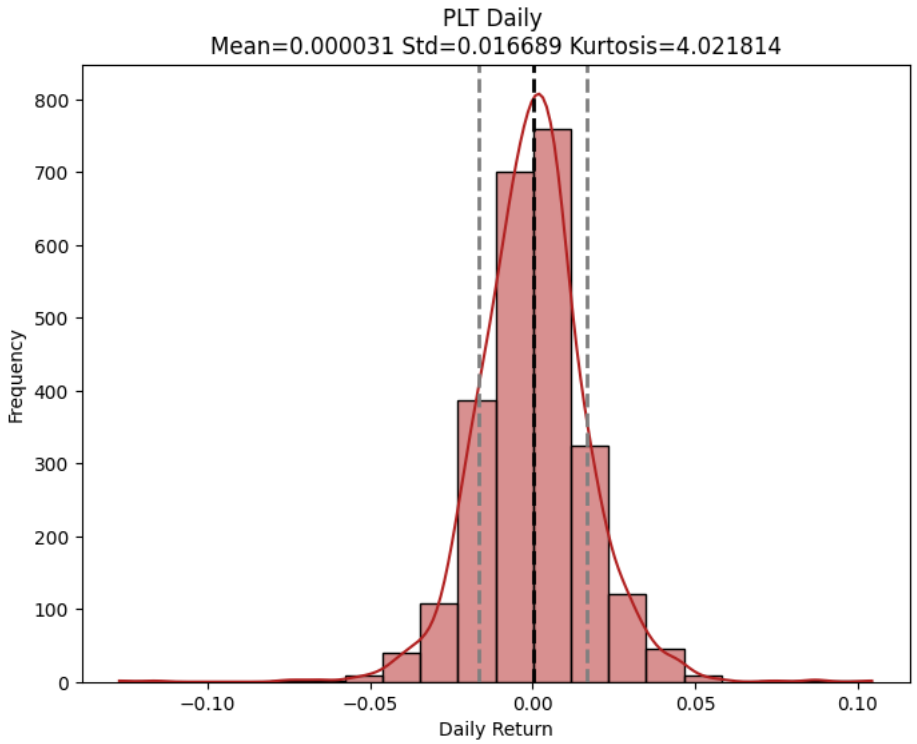
ภาพประกอบ 19 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา

จากภาพประกอบที่ 19 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา (USD) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000081 แสดงที่เส้นประสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.004409 แสดงที่เส้นประสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างน้อย จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน และค่าความโด่งอยู่ที่ 1.954654 ซึ่งใกล้เคียงกับการแจกแจงปกติ



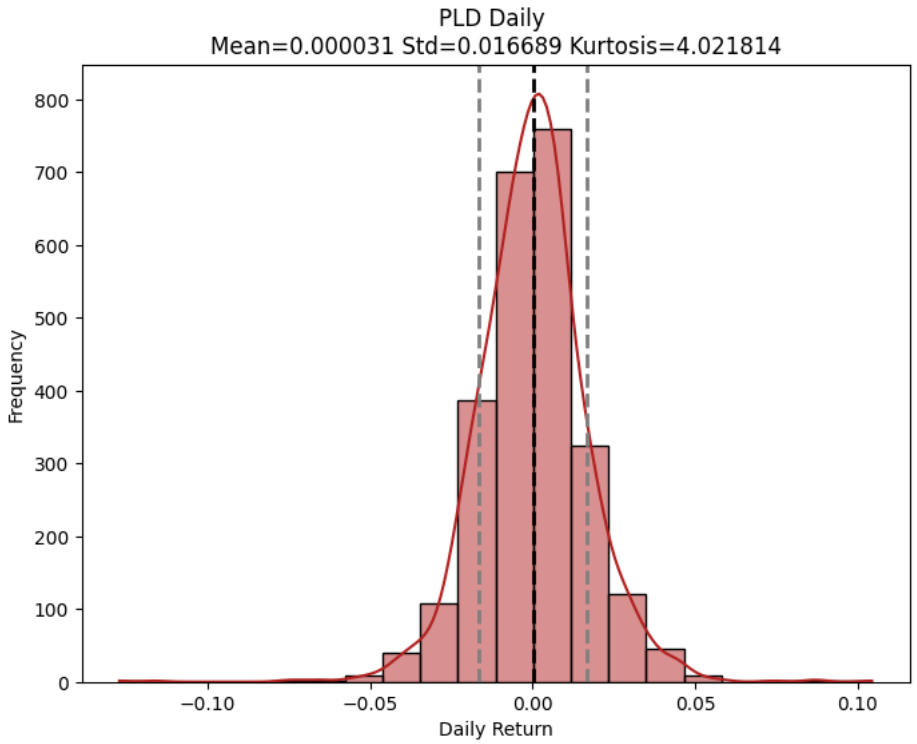
ภาพประกอบ 20 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่เงินล่วงหน้า

จากภาพประกอบที่ 20 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่เงินล่วงหน้า (SIL) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000405 แสดงที่เส้นปะสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.017940 แสดงที่เส้นปะสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่งอยู่ที่ 4.916239 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติมากนัก แต่มีการกระจายตัวที่มียอดแหลม อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



ภาพประกอบ 21 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า

จากภาพประกอบที่ 21 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่แพลทินัมล่วงหน้า (PLT) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000031 แสดงที่เส้นประสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.016689 แสดงที่เส้นประสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่ง อยู่ที่ 4.021814 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติ มากนัก แต่มีการกระจายตัวที่มียอดแหลม อธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา



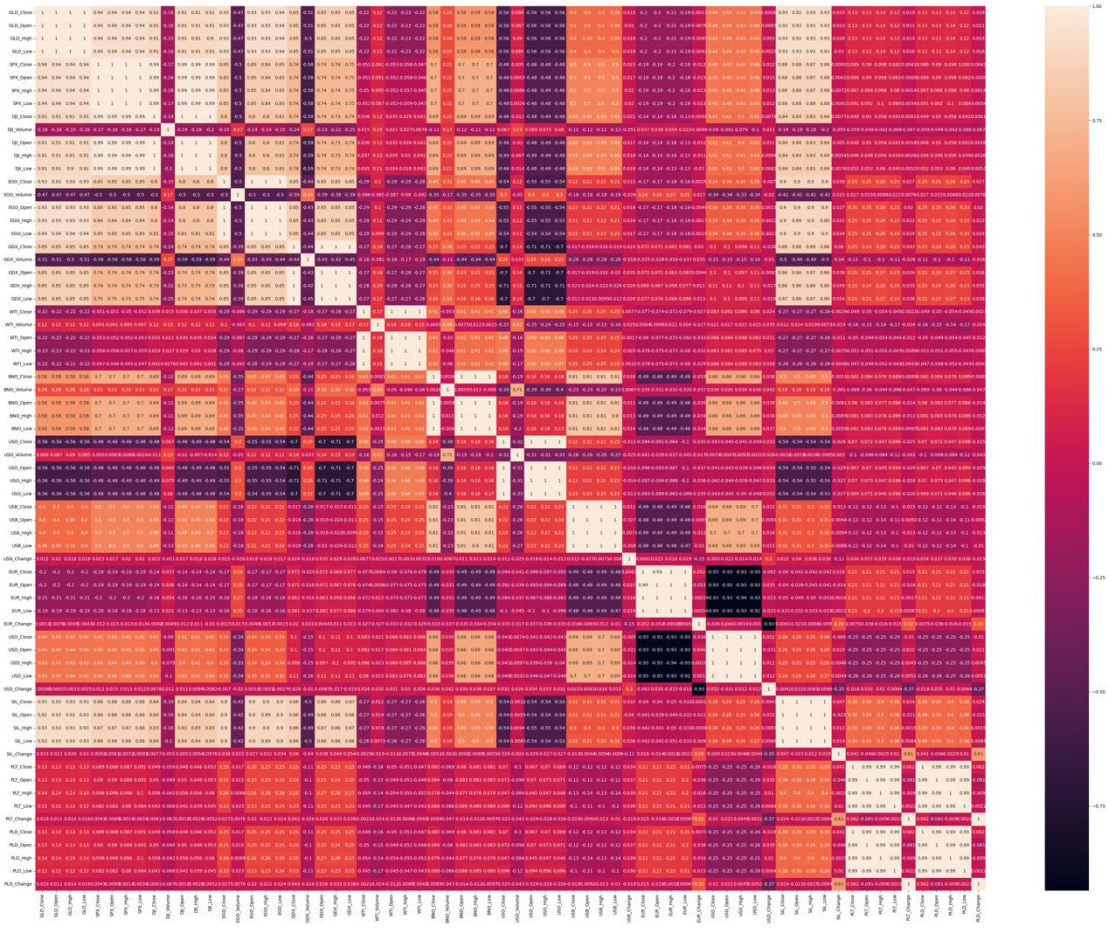
ภาพประกอบ 22 ภาพแสดงข้อมูลทางสถิติของอัตราการเปลี่ยนแปลงระหว่างวันของสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า

จากภาพประกอบที่ 22 การพิจารณาอัตราการเปลี่ยนแปลงระหว่างวันสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า (PLD) พบว่ามีค่าเฉลี่ยอยู่ที่ 0.000031 แสดงที่เส้นประสีดำ ซึ่งมีค่าที่เข้าใกล้ศูนย์ แสดงถึงทิศทางของข้อมูลไม่มีทิศทางบวกหรือลบที่ชัดเจน ในด้านส่วนเบี่ยงเบนมาตรฐานอยู่ที่ 0.016689 แสดงที่เส้นประสีเทาทั้งด้านบวกและลบ สามารถอธิบายได้ถึงการกระจายตัวของข้อมูลที่ค่อนข้างปกติ จากข้อมูลส่วนใหญ่ที่เกาะกลุ่มอยู่ในเส้นของค่าเบี่ยงเบนมาตรฐาน แต่กระจายเกือบเต็มขอบของส่วนเบี่ยงเบนมาตรฐานทั้งสองด้าน และค่าความโด่ง อยู่ที่ 4.021814 ซึ่งมีค่าข้อนข้างสูงแต่ไม่สูงจากการแจกแจงปกติมากนัก แต่มีการกระจายตัวที่มียอดแหลมอธิบายได้ว่ามีความเป็นไปได้ที่จะมีความแตกต่างที่สุดขั้วมากกว่าปกติ และอาจเกิดความผันผวนรุนแรงได้ในบางช่วงเวลา

3.2.3.2. **การพิจารณาความสัมพันธ์** (Correlation)

การพิจารณาความสัมพันธ์ โดยใช้วิธีการหาความสัมพันธ์ (Correlation) ระหว่างคุณลักษณะ ซึ่งการพยากรณ์ราคาทองคำ จะมุ่งเน้นถึงคุณลักษณะที่มีความสัมพันธ์ที่ใกล้ชิดกับ

ราคาปิด (GLD_Close) ซึ่งการแสดงความสัมพันธ์จะใช้กราฟความร้อน (Heat Map) ในการแสดงความสัมพันธ์ของข้อมูล โดยการแปรผลของกราฟความร้อนจะแปรความไปใน 2 ลักษณะ คือ ความสัมพันธ์ของคุณลักษณะในทิศทางเดียวกัน ค่าความสัมพันธ์จะแสดงในทิศทางบวก โดยยิ่งมีค่ามากจะแสดงถึงความสัมพันธ์ที่ใกล้ชิดในทิศทางเดียวกันอย่างมาก และความสัมพันธ์ของคุณลักษณะในทิศทางตรงข้ามกัน ค่าความสัมพันธ์จะแสดงในทิศทางลบ โดยยิ่งมีค่าลบมากจะแสดงถึงความสัมพันธ์ที่ใกล้ชิดในทิศทางตรงข้ามกันอย่างมาก



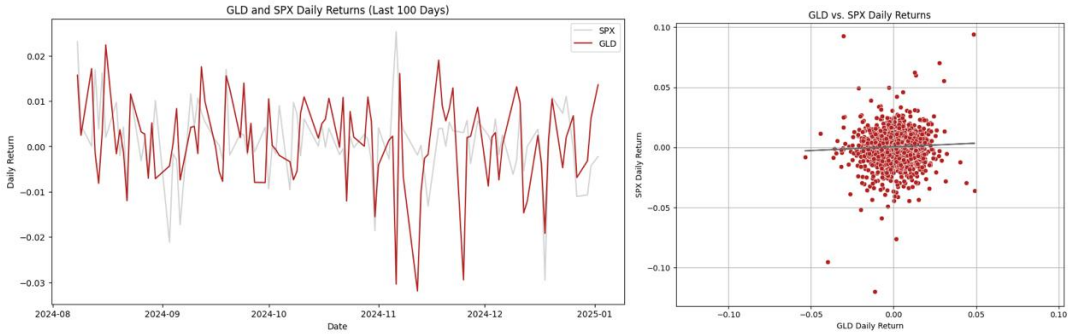
ภาพประกอบ 23 แสดงความสัมพันธ์ (Correlation) ระหว่างคุณลักษณะ

จากภาพประกอบที่ 23 แสดงความสัมพันธ์ ระหว่างคุณลักษณะ ซึ่งในกลุ่มของคุณลักษณะที่มีความสัมพันธ์กับราคาปิดตลาดของกองทุนทองคำในทิศทางเดียวกัน ได้แก่ ราคากองทุนทองคำ ได้ค่าความสัมพันธ์ที่ 1, ดัชนีตลาดหุ้นสหรัฐอเมริกาได้ค่าความสัมพันธ์ที่ 0.94, ดัชนีหุ้นเอลโดราโด โกลด์ ได้ค่าความสัมพันธ์ที่ 0.93, สัญญาซื้อขายแร่เงินล่วงหน้า ได้ค่า

ความสัมพันธ์ที่ 0.93, ดัชนีดาวโจนส์ ได้ค่าความสัมพันธ์ที่ 0.91,กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ ได้ค่าความสัมพันธ์ที่ 0.85, กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา ได้ค่าความสัมพันธ์ที่ 0.58, ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา ได้ค่าความสัมพันธ์ที่ 0.44, อัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว ได้ค่าความสัมพันธ์ที่ 0.40, อัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ ได้ค่าความสัมพันธ์ที่ 0.20, สัญญาซื้อขายแร่แพลทินัมล่วงหน้า ได้ค่าความสัมพันธ์ที่ 0.13 และสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า ได้ค่าความสัมพันธ์ที่ 0.13 และคุณลักษณะที่มีความสัมพันธ์กับราคาปิดตลาดของกองทุนทองคำในทิศทางตรงข้ามกัน ได้แก่ กองทุนรวมราคาน้ำมันสหรัฐอเมริกา ได้ค่าความสัมพันธ์ที่ -0.56 และดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ ได้ค่าความสัมพันธ์ที่ -0.22

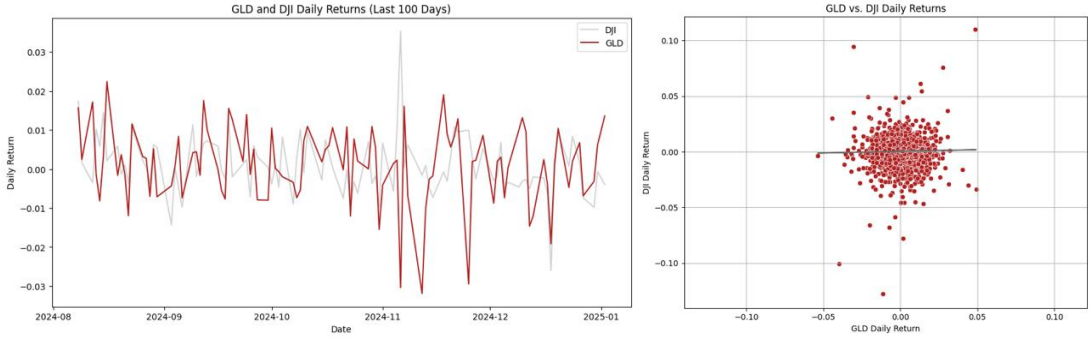
เมื่อพิจารณาความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างราคาปิดกองทุนทองคำ กับอัตราการเปลี่ยนแปลงของคุณลักษณะอื่นๆ โดยแสดงข้อมูลอัตราการเปลี่ยนแปลง 100 วันล่าสุดเพื่อสังเกตความเปลี่ยนแปลงในปัจจุบัน และการกระจายตัวของข้อมูล ซึ่งสามารถอธิบายความสัมพันธ์ของข้อมูลได้ในระดับหนึ่ง โดยการพิจารณาความสัมพันธ์สามารถอธิบายได้ดังนี้

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างราคาปิดกองทุนทองคำ และดัชนีตลาดหุ้นสหรัฐอเมริกา พบว่าในกราฟแสดงอัตราการเปลี่ยนแปลงในหลายจุดที่อัตราการเปลี่ยนแปลงของกองทุนทองคำ และดัชนีตลาดหุ้นสหรัฐอเมริกามีแนวโน้มที่ไปในทิศทางเดียวกันค่อนข้างมาก ซึ่งแสดงได้จากเส้นแนวโน้ม แต่หากที่สังเกตในช่วงเดือนพฤศจิกายน 2024 เกิดเหตุการณ์ที่ความสัมพันธ์มีทิศทางตรงข้ามกันที่ค่อนข้างรุนแรง ซึ่งเป็นผลกระทบจากการเลือกตั้งใหม่ในประเทศสหรัฐอเมริกา ส่งผลให้ตลาดหุ้นสหรัฐอเมริกาเป็นที่ต้องการของนักลงทุนมากยิ่งขึ้น จึงมีการปรับเปลี่ยนการลงทุนไปที่ดัชนีตลาดหุ้นสหรัฐอเมริกา ซึ่งเป็นผลกระทบภายนอกที่ทำให้ความสัมพันธ์ของข้อมูลไม่ไปในทิศทางเดียวกันได้ ในด้านการจะจายตัวของอัตราการเปลี่ยนแปลงจะพบว่ามีหลายจุดความสัมพันธ์ที่ไม่เกาะกลุ่มกับความสัมพันธ์ข้อมูลส่วนใหญ่ ซึ่งอาจใช้เป็นสัญญาณบอกความเปลี่ยนแปลงในแต่ละช่วงได้ โดยข้อมูลที่กระจายออกมานั้นมีความสัมพันธ์ที่หลากหลาย แต่สามารถสังเกตได้ว่าเมื่อราคากองทุนทองคำจะเพิ่มขึ้นไม่เกิน 0.05 และจะไม่ลดลงเกิน 0.5 เช่นกัน โดยมีเพียงจุดเดียวเท่านั้นที่ลดลงเกิน 0.05 ซึ่งในจุดดังกล่าวดัชนีตลาดหุ้นสหรัฐอเมริกาไม่สามารถอธิบายเหตุการณ์ที่เกิดขึ้นได้



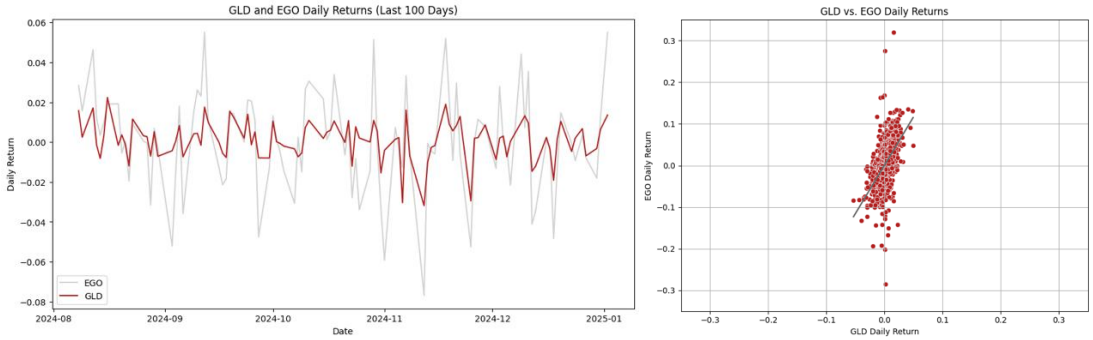
ภาพประกอบ 24 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันและการกระจายตัวของราคาปิดของกองทุนทองคำและดัชนีตลาดหุ้นสหรัฐอเมริกา

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างราคาปิดกองทุนทองคำ และดัชนีดาวโจนส์ พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะเหมือนกับดัชนีดาวโจนส์ แต่ดัชนีดาวโจนส์การกระจายข้อมูลที่ต่ำกว่า หรืออัตราการเปลี่ยนแปลงของดัชนีดาวโจนส์นั้นไม่รุนแรงเท่าดัชนีตลาดหุ้นสหรัฐอเมริกา ซึ่งสังเกตได้จากเส้นแนวโน้ม



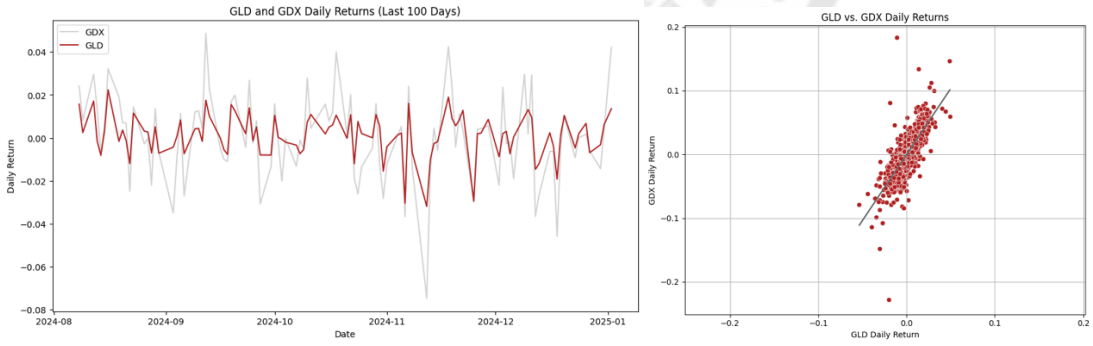
ภาพประกอบ 25 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีดาวโจนส์

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างราคาปิดกองทุนทองคำ และดัชนีหุ้นเอลโดราโด โกลด์ พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกันแต่เนื่องจากอัตราการเปลี่ยนแปลงของดัชนีหุ้นเอลโดราโด โกลด์ มีความรุนแรงที่มากกว่ากองทุนทองคำ จึงส่งผลให้เส้นแนวโน้มมีทิศทางความสัมพันธ์ในเชิงบวกอย่างชัดเจน



ภาพประกอบ 26 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีหุ้นเอลโดราโด โกลด์

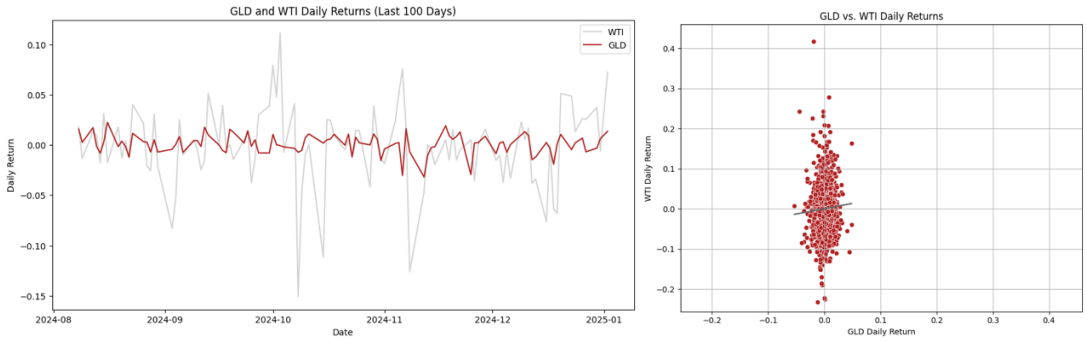
ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำและกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน แต่เมื่อสังเกตในช่วงเดือนพฤศจิกายน 2024 ซึ่งเกิดเหตุการณ์การเลือกตั้งใหม่ในประเทศสหรัฐอเมริกา อธิบายได้อย่างชัดเจนว่านักลงทุนเปลี่ยนทิศทางการลงทุนอย่างรุนแรงทำให้กองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำได้รับผลกระทบที่รุนแรงกว่ากองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ



ภาพประกอบ 27 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและกองทุนซื้อขายหุ้นกลุ่มบริษัทเหมืองทองคำ

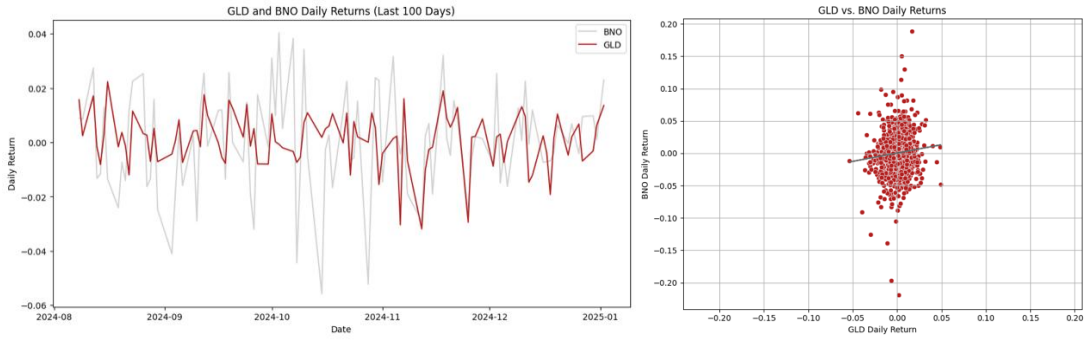
ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และดัชนีหุ้นดับบลิว แอนด์ ที ออฟชอร์ พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน แต่หากสังเกตในจุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 นั้น ดัชนีหุ้นดับบลิว แอนด์ ที

ออฟซอร์ มีอัตราการเปลี่ยนแปลงที่เป็นบวก ซึ่งเป็นสัญญาณได้ว่าดัชนี หรือกองทุนในกลุ่มของน้ำมัน และพลังงานอาจมีสัญญาณที่ส่งผลให้กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีแนวโน้มลดลงอย่างรุนแรงได้



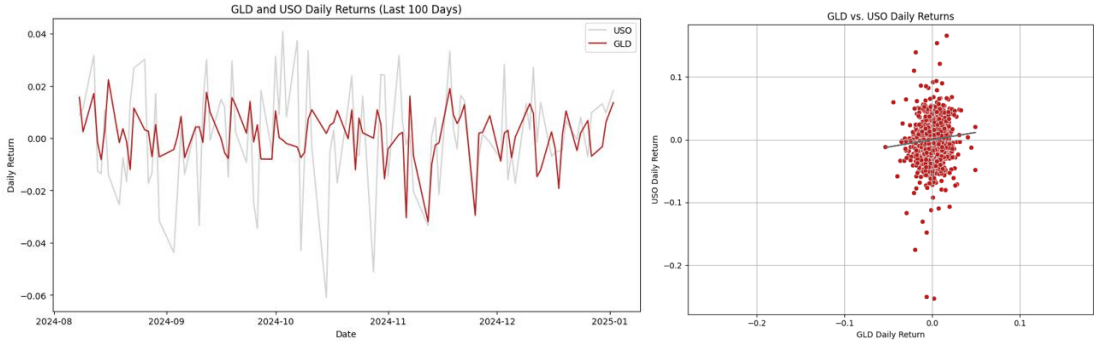
ภาพประกอบ 28 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีหุ้นดับบลิว แอนด์ ที ออฟซอร์

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา พบว่าลักษณะอัตราการเปลี่ยนแปลงและการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน เหมือนกับดัชนีหุ้นดับบลิว แอนด์ ที ออฟซอร์ แต่จุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 นั้น กองทุนน้ำมันเบรนต์ของสหรัฐอเมริกาไม่สามารถอธิบายได้เนื่องจากอัตราการเปลี่ยนแปลงยังคงไปในทิศทางเดียวกัน และไม่รุนแรงจนสามารถอธิบายได้



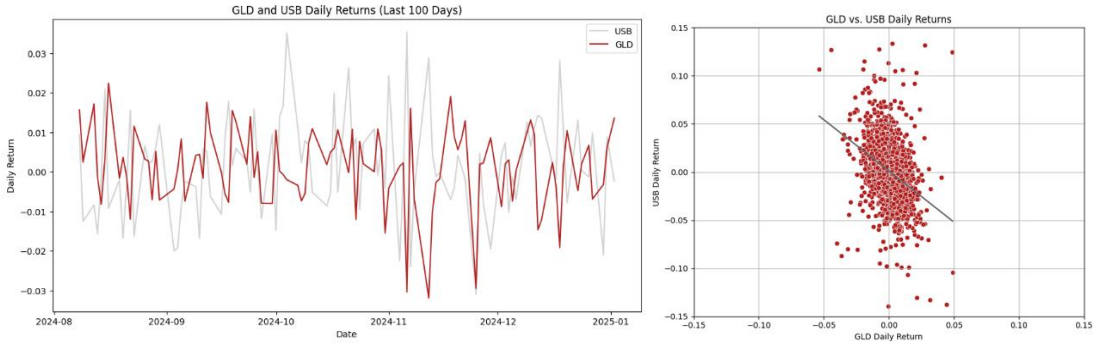
ภาพประกอบ 29 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและกองทุนน้ำมันเบรนต์ของสหรัฐอเมริกา

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และกองทุนรวมราคาน้ำมันสหรัฐอเมริกา พบว่าลักษณะอัตราการเปลี่ยนแปลงและการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน เหมือนกับกองทุนรวมราคาน้ำมันสหรัฐอเมริกา และไม่สามารถอธิบายจุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 ได้



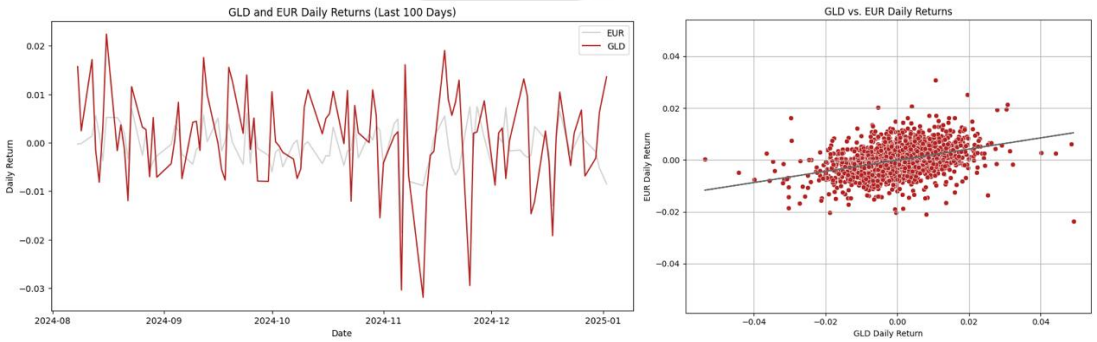
ภาพประกอบ 30 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและกองทุนรวมราคาน้ำมันสหรัฐอเมริกา

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางตรงกันข้ามกัน ซึ่งทำให้จุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 มีอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาวไปในทิศทางบวก ซึ่งตรงข้ามกับอัตราการเปลี่ยนแปลงของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ



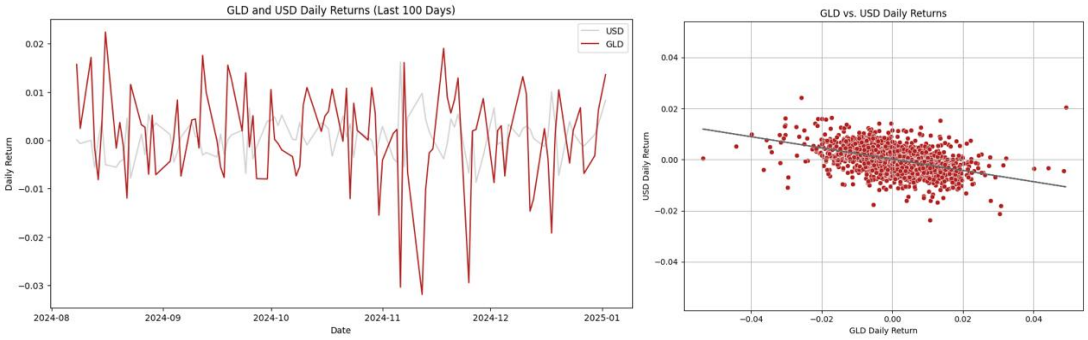
ภาพประกอบ 31 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและอัตราผลตอบแทนพันธบัตรสหรัฐแบบระยะยาว

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน โดยหากอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐเพิ่มขึ้น กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำก็มีแนวโน้มที่จะเพิ่มขึ้นด้วย โดยหากสังเกตุในจุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 อัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐไม่มีการเปลี่ยนแปลงเลย



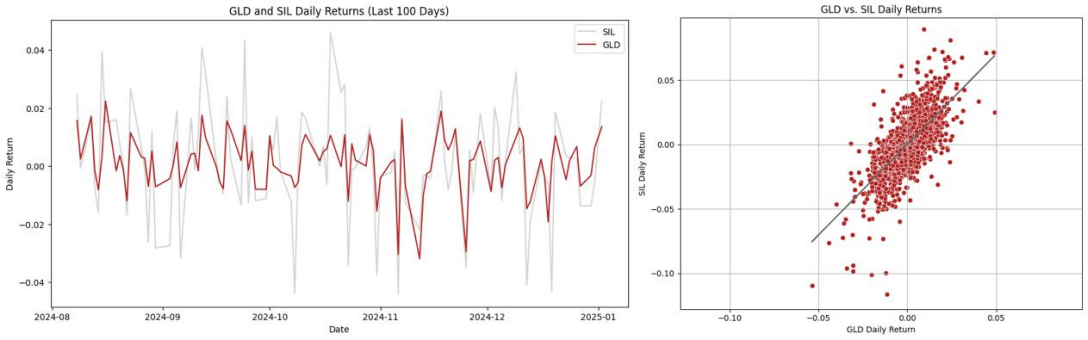
ภาพประกอบ 32 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางตรงข้ามกัน แต่ในจุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกาไม่มีการเปลี่ยนแปลงเลย ดังนั้นข้อมูลในกลุ่มค่าเงิน คือ ดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา และอัตราแลกเปลี่ยนสกุลเงินยูโรเป็นดอลลาร์สหรัฐ อาจสามารถใช้เพื่ออธิบายได้ว่าในจุดที่กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำมีอัตราการเปลี่ยนแปลงน้อยกว่า 0.05 นั้นข้อมูลในกลุ่มค่าเงินจะไม่มีการเปลี่ยนแปลงเลย



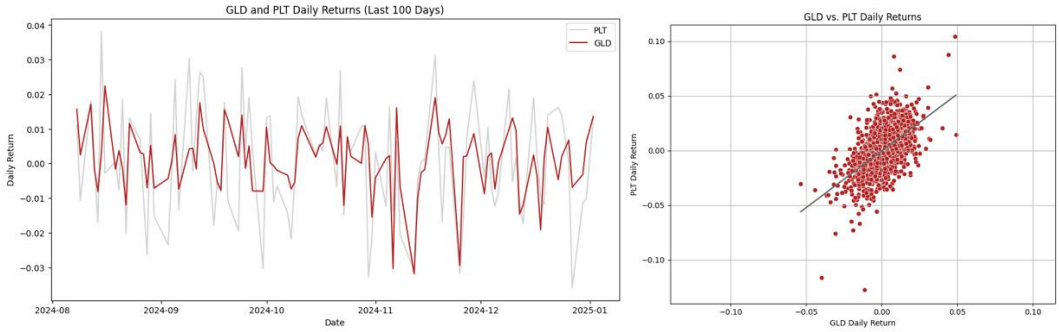
ภาพประกอบ 33 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและดัชนีค่าเงินดอลล่าร์สหรัฐอเมริกา

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และสัญญาซื้อขายแร่เงินล่วงหน้า พบว่าลักษณะอัตราการเปลี่ยนแปลง และการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน กล่าวคือหากสัญญาซื้อขายแร่เงินล่วงหน้ามีอัตราการเปลี่ยนแปลงเป็นบวก กองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำก็จะมีแนวโน้มเป็นบวกด้วย



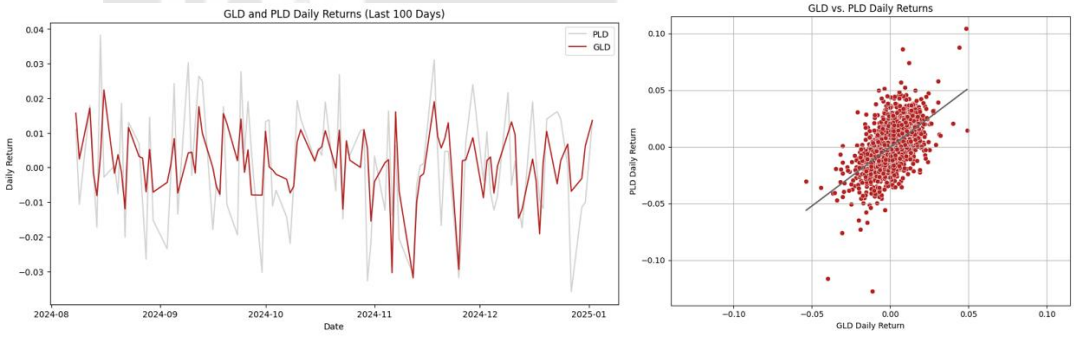
ภาพประกอบ 34 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและสัญญาซื้อขายแร่เงินล่วงหน้า

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และสัญญาซื้อขายแร่แพลทินัมล่วงหน้า พบว่าลักษณะอัตราการเปลี่ยนแปลงและการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน ซึ่งมีลักษณะเดียวกับสัญญาซื้อขายแร่เงินล่วงหน้า



ภาพประกอบ 35 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและสัญญาซื้อขายแร่แพลทินัมล่วงหน้า

ความสัมพันธ์ระหว่างอัตราการเปลี่ยนแปลงระหว่างกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ และสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า พบว่าลักษณะอัตราการเปลี่ยนแปลงและการกระจายตัวของข้อมูลมีลักษณะที่ไปในทิศทางเดียวกัน ซึ่งมีลักษณะเดียวกับสัญญาซื้อขายแร่เงินล่วงหน้า และสัญญาซื้อขายแร่แพลทินัมล่วงหน้า



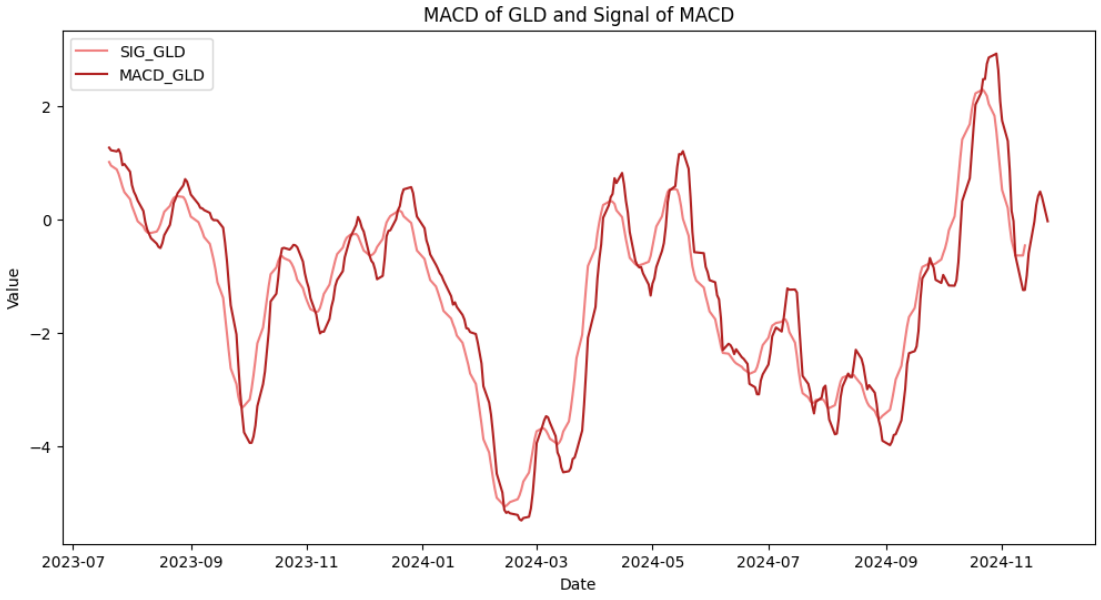
ภาพประกอบ 36 ภาพแสดงอัตราการเปลี่ยนแปลงระหว่างวันของราคาปิดของกองทุนทองคำและสัญญาซื้อขายแร่พัลลาเดียมล่วงหน้า

3.2.3. **กระบวนการสร้างและปรับปรุงคุณลักษณะ** (Feature Engineering)

คุณลักษณะที่มีความสัมพันธ์ใกล้ชิดกับราคาปิดกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ คือคุณลักษณะในกลุ่มของกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำ ซึ่งหมายถึงคุณลักษณะของตัวเองมีความสัมพันธ์ที่ดีที่สุดกับตัวเอง ดังนั้นการใช้ประโยชน์จากข้อมูลในกลุ่มกองทุนซื้อขายแลกเปลี่ยนที่ลงทุนในทองคำจึงเป็นสิ่งที่น่าสนใจในการนำมาผ่านกระบวนการทางสถิติ เพื่อใช้เป็นตัวเลือกและการปรับปรุงการทำงานของการเรียนรู้ของเครื่องให้ดียิ่งขึ้น

การสร้างและการปรับปรุงคุณลักษณะจะใช้วิธีการทางสถิติ ได้แก่ Moving Average Convergence Divergence (MACD), ดัชนีความแข็งแกร่งสัมพัทธ์ (Relative Strength Index : RSI), ค่าเฉลี่ยเคลื่อนที่แบบง่าย (Simple Moving Average : SMA), ขอบบนของค่าเฉลี่ยเคลื่อนที่ (Upper Band of Simple Moving Average), ขอบล่างของค่าเฉลี่ยเคลื่อนที่ (Lower Band of Simple Moving Average) , ส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่ (Moving Standard Deviation), ผลต่างของราคาปิดและราคาเปิด (Open-Close) และผลต่างของราคาสูงสุดและราคาต่ำสุด (High-Low)

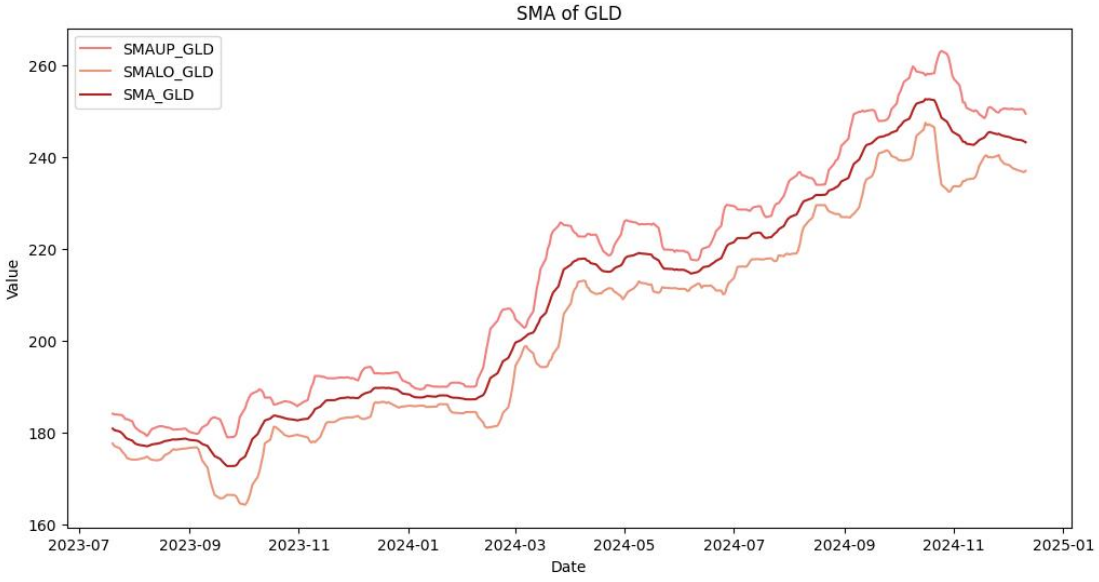
การคำนวณ MACD จะใช้ค่าเฉลี่ยเคลื่อนที่แบบเอ็กโปเนนเชียลแบบระยะสั้นจำนวน 12 วัน และแบบระยะยาวจำนวน 26 วัน หาค่าส่วนต่างระหว่างแบบระยะสั้น และแบบระยะยาว แทนชื่อการคำนวณนี้ด้วยชื่อ MACD_GLD และเมื่อได้ส่วนต่างแล้ว จะทำการหาเส้นสัญญาณ (Signal Line) โดยการใช้ค่าเฉลี่ยเคลื่อนที่แบบเอ็กโปเนนเชียลแบบระยะสั้นจำนวน 9 วัน แทนชื่อการคำนวณนี้ด้วยชื่อ SIG_GLD ซึ่งการใช้หาค่าเฉลี่ยเคลื่อนที่แบบเอ็กโปเนนเชียลใช้ฟังก์ชัน ewm() จาก Pandas Library โดยผลที่ได้จากการคำนวณ สามารถแสดงได้ดังภาพประกอบที่ 37 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณ MACD ของราคาปิดกองทุนทองคำ



ภาพประกอบ 37 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณ MACD ของราคาปิดกองทุนทองคำ

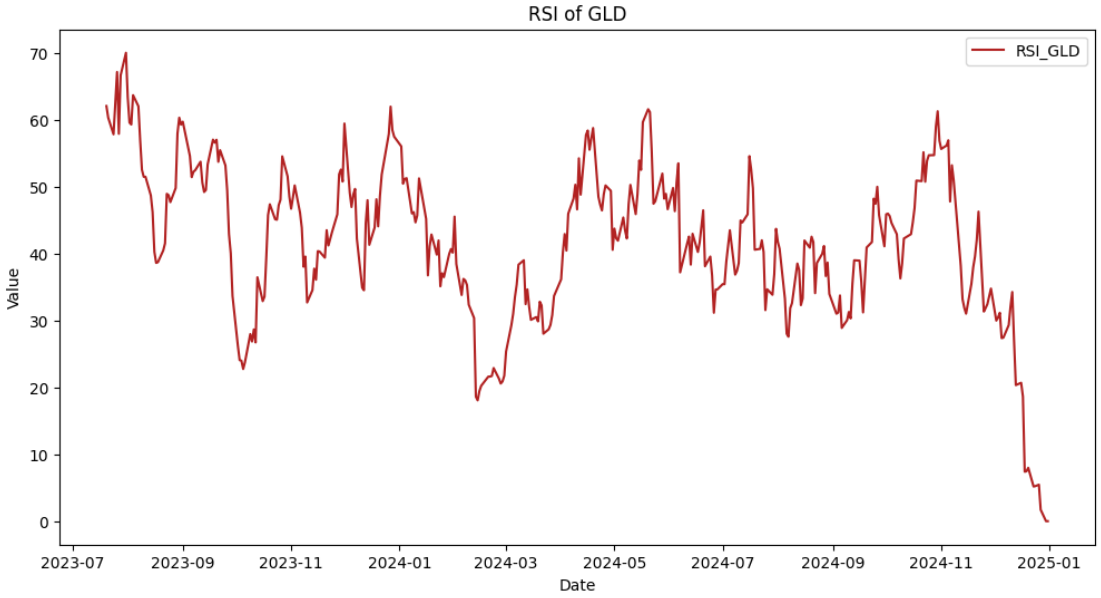
การคำนวณค่าเฉลี่ยเคลื่อนที่แบบง่าย จะใช้การคำนวนค่าเฉลี่ยเคลื่อนที่แบบง่ายจากข้อมูล 15 วัน แทนชื่อการคำนวณนี้ด้วยชื่อ SMA_GLD โดยใช้ฟังก์ชัน rolling().mean() จาก

Pandas Library ซึ่งการคำนวณค่าเฉลี่ยนเคลื่อนที่แบบง่าย จะสามารถหาขอบบน (Upper Band) และขอบล่าง (Lower Band) โดยได้จากผลรวมของค่าเฉลี่ยเคลื่อนแบบง่ายที่กับส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่แบบง่าย (Simple Moving Standard Deviation) ในกรณีของการหาของบน (Upper Band) แทนชื่อการคำนวณนี้ด้วยชื่อ SMAUP_GLD และผลต่างของค่าเฉลี่ยเคลื่อนแบบง่ายที่กับส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่แบบง่าย (Simple Moving Standard Deviation) ในกรณีของการหาของขอบล่าง (Lower Band) แทนชื่อการคำนวณนี้ด้วยชื่อ SMALO_GLD โดยผลที่ได้จากการคำนวณ สามารถแสดงได้ดังภาพประกอบที่ 38



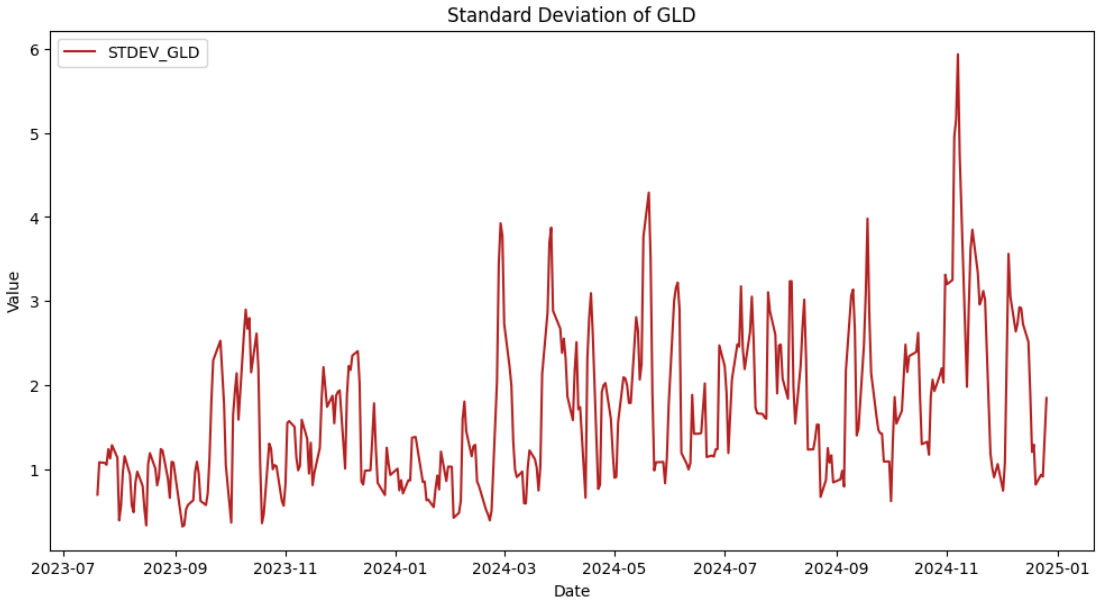
ภาพประกอบ 38 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณค่าเฉลี่ยเคลื่อนที่แบบง่าย

การคำนวณดัชนีความแข็งแกร่งสัมพัทธ์ จะใช้การคำนวนจากค่าเฉลี่ยเคลื่อนที่แบบเอ็กโปเนนเชียลจำนวน 14 วัน แทนชื่อการคำนวณนี้ด้วยชื่อ RSI_GLD โดยผ่านฟังก์ชัน ewm() จาก Pandas Library สามารถแสดงได้ดังภาพประกอบที่ 39 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณดัชนีความแข็งแกร่งสัมพัทธ์



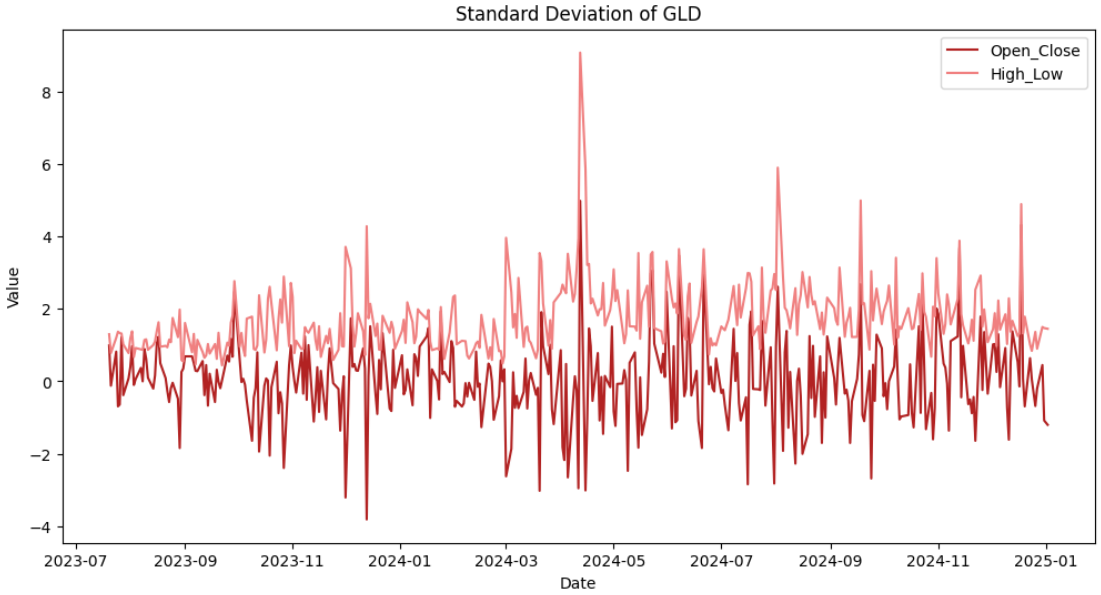
ภาพประกอบ 39 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณดัชนีความแข็งแกร่งสัมพัทธ์

การคำนวณส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่ จะใช้การคำนวณส่วนเบี่ยงเบนมาตรฐานแบบง่ายจำนวณ 5 วัน แทนชื่อการคำนวณนี้ด้วยชื่อ STDEV_GLD โดยผ่านฟังก์ชัน rolling().std() จาก Pandas Library สามารถแสดงได้ดังภาพประกอบที่ 40 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่



ภาพประกอบ 40 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่

การคำนวณผลต่างของราคาปิดและราคาเปิด และผลต่างของราคาสูงสุดและราคาต่ำสุดใช้การคำนวณผลต่างระหว่างวันในการคำนวณ สามารถแสดงได้ดังภาพประกอบที่ 41 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณผลต่างของราคาปิดและราคาเปิด, ราคาสูงสุดและราคาต่ำสุด



ภาพประกอบ 41 ภาพแสดงตัวอย่างข้อมูลจากการคำนวณผลต่างของราคาปิดและราคาเปิด, ราคาสูงสุดและราคาต่ำสุด

3.2.4. **กระบวนการแบ่งชุดข้อมูล** (Splitting Dataset)

งานวิจัยการพยากรณ์ราคาทองคำ จะทำการพยากรณ์ราคาปิดของกองทุนทองคำล่วงหน้า โดยการพยากรณ์จะใช้คุณลักษณะเป้าหมาย (Target) ที่แตกต่างกัน ซึ่งจะแบ่งคุณลักษณะเป้าหมายออกเป็น 1, 3, 5 และ 7 วัน

กระบวนการแบ่งชุดข้อมูลเริ่มจากการปรับขนาดของข้อมูล (Data Normalization) ให้อยู่ในช่วง [0, 1] ผ่านการใช้งานฟังก์ชัน MinMaxScaler() จาก Sklearn Libarry เพื่อลดความแปรปรวนของข้อมูลเพื่อไม่ให้คุณลักษณะบางประการมีผลมากจนเกินไป ซึ่งทำให้สามารถเพิ่มประสิทธิในการพยากรณ์ และเร่งกระบวนการทำงานให้รวดเร็วยิ่งขึ้น ภาพประกอบที่ 42 และภาพประกอบที่ 43 แสดงข้อมูลตัวอย่างหลังจากผ่านการปรับขนาดของข้อมูล

Date	GLD_Close	GLD_Open	GLD_High	GLD_Low	SPX_Close	SPX_Open	SPX_High	SPX_Low	DJI_Close	DJI_Volume	DJI_Open
2024-11-13 00:00:00	0.8734394904458599	0.9008598562628338	0.896375701888719	0.8823201811437106	0.997561479049361	0.995719629753968	0.9989773013720005	0.9973959958335934	0.998942246668076	0.05856560153198102	0.9872427697441605
2024-11-12 00:00:00	0.8888535031847135	0.9007315195071	0.8975880551301686	0.8938825456472513	0.9972278625019801	1.0	0.9993926226458125	0.9959967935948697	0.997461392003384	0.0832741056930834	1.0
2024-11-08 00:00:00	0.939235668789809	0.9506545174537988	0.9466564573762125	0.9477760113054229	1.0	0.9935638578485444	1.0	1.0000000000000002	1.0	0.1276849029796295	0.9816462736373751
2024-11-07 00:00:00	0.9500000000000001	0.9456493839835731	0.9523353751914244	0.9499600134894252	0.9946141328609901	0.9864778667689797	0.9931315944255723	0.9929079886527818	0.990409703123898	0.1060361963027742	0.9789696885428253
2024-11-06 00:00:00	0.9248407643312101	0.928612679671458	0.9338948443083207	0.9296012590130237	0.984039208344734	0.9667378063402234	0.9816802506331851	0.9731511570418514	0.9898455680135392	0.1931541751230454	0.9689933259176865

ภาพประกอบ 42 แสดงข้อมูลตัวอย่างหลังจากผ่านการปรับขนาดของข้อมูล

DJI_High	DJI_Low	EGO_Close	EGO_Volume	EGO_Open	EGO_High	EGO_Low	GDX_Close	GDX_Volume
0.9906708595387841	0.9978851594068463	0.817629179331307	0.027410958239551 56	0.8163944459319893	0.8188131449711183	0.8215711713323863	0.7202121023081721	0.0743101795201 1144
1.0	1.0	0.8181818181818182	0.027426387734265062	0.8021868732957742	0.8090777318650888	0.8188142398586303	0.7373674360573922	0.1307127671876425 8
0.9911949685534589	0.9942801205330882	0.8889195910472506	0.015196489305025701	0.871585401172671	0.8739804859052853	0.8792094932390868	0.8281347473487212	0.065670743403282
0.9791303284416 49	0.9909177285066022	0.8938933407018513	0.0179105975400456	0.8590171638406345	0.878848192458 3	0.8685844023665991	0.8452900810979412	0.0837896439163301 6
0.9775157232704403	0.9814891652983118	0.864050842774247	0.032170897243753 21	0.8240446773514898	0.8533143673063193	0.8311169766583529	0.8162819713038053	0.1611993683816070 3

ภาพประกอบ 43 แสดงข้อมูลตัวอย่างหลังจากผ่านการปรับขนาดของข้อมูล (ต่อ)

การเตรียมชุดข้อมูลสำหรับการพยากรณ์แบบ 1, 3, 5 และ 7 วัน จะใช้วิธีการขยับคุณลักษณะเป้าหมายขึ้นโดยใช้คำสั่ง np.roll() ตามจำนวณวันที่ต้องการขยับ เช่น การขยับคุณลักษณะเป้าหมายของ 1 วัน จะใช้คำสั่ง np.roll(y, -1) เป็นต้น ซึ่งกระบวนการนี้จะได้รับชุดข้อมูลที่แบ่งออกมา 4 ชุดข้อมูล สำหรับการพยากรณ์ 1, 3, 5 และ 7 วัน

การแบ่งชุดข้อมูลจะทำการแบ่งชุดข้อมูลจากข้อมูลที่ถูกขยับคุณลักษณะเป้าหมายแล้ว โดยงานวิจัยการพยากรณ์ราคาทองคำจะใช้การแบ่งชุดข้อมูลออกเป็นสำหรับการเรียนรู้ 80% และสำหรับการทดสอบ 20% ผ่านฟังก์ชัน train_test_split จาก Sklearn Libarry โดยเมื่อผ่านกระบวนการทั้งหมดแล้ว ทำให้ได้การแบ่งชุดข้อมูลตามรูปภาพที่ 44 แสดงรูปร่างของข้อมูลหลังจากทำการแบ่งชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

```
Shape of X_train_1day for 1-day forecast: (1982, 76)
Shape of y_train_1day for 1-day forecast: (1982,)
Shape of X_test_1day for 1-day forecast: (496, 76)
Shape of y_test_1day for 1-day forecast: (496,)

Shape of X_train_3day for 3-day forecast: (1980, 76)
Shape of y_train_3day for 3-day forecast: (1980,)
Shape of X_test_3day for 3-day forecast: (496, 76)
Shape of y_test_3day for 3-day forecast: (496,)

Shape of X_train_5day for 5-day forecast: (1979, 76)
Shape of y_train_5day for 5-day forecast: (1979,)
Shape of X_test_5day for 5-day forecast: (495, 76)
Shape of y_test_5day for 5-day forecast: (495,)

Shape of X_train_7day for 7-day forecast: (1977, 76)
Shape of y_train_7day for 7-day forecast: (1977,)
Shape of X_test_7day for 7-day forecast: (495, 76)
Shape of y_test_7day for 7-day forecast: (495,)
```

ภาพประกอบ 44 แสดงรูปร่างของข้อมูลหลังจากทำการแบ่งชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

3.2.5. **กระบวนการคัดเลือกคุณลักษณะ** (Feature Selection)

เนื่องจากคุณลักษณะที่มีมากเกินไปในชุดข้อมูลซึ่งอาจจะส่งผลให้การเรียนรู้ของเครื่องไม่มีประสิทธิภาพเท่าที่ควร โดยกระบวนการคัดเลือกคุณลักษณะจะคัดเลือกคุณลักษณะที่มีผลต่อการพยากรณ์ราคาปิดของกองทุนทองคำ

การคัดเลือกคุณลักษณะเพื่อคัดเลือกคุณลักษณะบางประการที่มีผลต่อการพยากรณ์ราคาปิดของกองทุนทองคำ โดยงานวิจัยนี้ใช้เครื่องมือในการคัดเลือกคุณลักษณะ 3 เครื่องมือ ได้แก่ การคัดเลือกโดยการถดถอยลาสโซ , การถดถอยของป่าแบบสุ่ม และการค้นหาแบบขั้นตอนเชิงพันธุกรรม เพื่อค้นหาเครื่องมือการคัดเลือกคุณลักษณะที่ดีที่สุด ที่สามารถทำให้การเรียนรู้ของเครื่องสามารถเรียนรู้ได้ดีที่สุดเช่นกัน

การคัดเลือกคุณลักษณะโดยการถดถอยลาสโซ สามารถคัดเลือกคุณลักษณะจากการที่ใช้ L1 Regularization ในการเพิ่มน้ำหนักให้กับคุณลักษณะที่มีสำคัญกับเป้าหมาย และการลดน้ำหนักจนเป็นศูนย์ให้กับคุณลักษณะที่ไม่มีความสำคัญกับเป้าหมาย ซึ่งงานวิจัยนี้ใช้ฟังก์ชัน SelectFormModel จาก Sklearn Libarry โดยกระบวนการคัดเลือกคุณลักษณะจะคัดเลือกคุณลักษณะจากการขยับข้อมูลและแบ่งข้อมูลเรียบร้อยแล้ว ส่งผลให้ได้คุณลักษณะที่ถูกคัดเลือกตามภาพประกอบที่ 45 แสดงคุณลักษณะที่ถูกคัดเลือกจากการถดถอยลาสโซ ของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

```
Selected features for 1-day forecast:
 ['GLD_Close', 'GLD_Open', 'SPX_Open', 'GDX_Close', 'USO_High', 'EUR_Change', 'USD_Low', 'SIL_Close', 'SIL_Low', 'PLT_High', 'SIG_GLD', 'SMAUP_GLD']
Number of selected features for 1day: 12

Selected features for 3-day forecast:
 ['GLD_Close', 'GLD_Open', 'SPX_Open', 'GDX_Close', 'USO_High', 'EUR_Change', 'USD_Low', 'SIL_Close', 'SIL_Low', 'PLT_High', 'SIG_GLD', 'SMAUP_GLD']
Number of selected features for 1day: 12

Selected features for 5-day forecast:
 ['GLD_Close', 'GLD_Open', 'SPX_Open', 'GDX_Close', 'USO_High', 'EUR_Change', 'USD_Low', 'SIL_Close', 'SIL_Low', 'PLT_High', 'SIG_GLD', 'SMAUP_GLD']
Number of selected features for 1day: 12

Selected features for 7-day forecast:
 ['GLD_Close', 'GLD_Open', 'SPX_Open', 'GDX_Close', 'USO_High', 'EUR_Change', 'USD_Low', 'SIL_Close', 'SIL_Low', 'PLT_High', 'SIG_GLD', 'SMAUP_GLD']
Number of selected features for 1day: 12
```

ภาพประกอบ 45 แสดงคุณลักษณะที่ถูกคัดเลือกจากการถดถอยลาสโซของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

การคัดเลือกคุณลักษณะโดยการถดถอยของป่าแบบสุ่มใช้การคัดเลือกแบบต้นไม้หลายต้นช่วยกันตัดสินใจ ซึ่งงานวิจัยนี้ต้นไม้ในการตัดสินใจจำนวน 100 ต้น (n_estimators = 100) โดยค่าที่ได้รับจะเป็นค่าความแม่นยำที่ได้จากการพยากรณ์ และใช้ค่ามัธยฐาน ในการคัดเลือกคุณลักษณะที่มีค่าความแม่นยำสูงกว่าค่ามัธยฐาน ซึ่งงานวิจัยนี้ใช้ฟังก์ชัน SelectFormModel จาก Sklearn Libarry โดยกระบวนการคัดเลือกคุณลักษณะจะคัดเลือก

คุณลักษณะจากการขยับข้อมูลและแบ่งข้อมูลเรียบร้อยแล้ว ส่งผลให้ได้คุณลักษณะที่ถูกคัดเลือกตามภาพประกอบที่ 46 แสดงคุณลักษณะที่ถูกคัดเลือกจากการถดถอยของป่าแบบสุ่มของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

```
Total features selected by Random Forest for 1 day: 38
Features for 1 day:
 Index(['GLD_Close', 'GLD_Open', 'GLD_High', 'SPX_Close', 'SPX_Open', 'SPX_Low',
       'DJI_Volume', 'DJI_Open', 'DJI_High', 'DJI_Low', 'EGO_Volume',
       'EGO_Open', 'EGO_High', 'EGO_Low', 'GDX_Close', 'GDX_Volume',
       'GDX_Open', 'GDX_High', 'USB_Close', 'USB_Open', 'USB_High', 'USB_Low',
       'EUR_Low', 'USD_Close', 'USD_Low', 'USD_Change', 'SIL_Close',
       'SIL_Open', 'SIL_High', 'SIL_Low', 'PLT_Low', 'PLD_Low', 'PLD_Change',
       'MACD_GLD', 'SIG_GLD', 'SMA_GLD', 'SMAUP_GLD', 'SMALO_GLD'],
      dtype='object')

Total features selected by Random Forest for 3 days: 38
Features for 3 days:
 Index(['GLD_Close', 'GLD_Open', 'GLD_High', 'SPX_Close', 'SPX_Open', 'SPX_Low',
       'DJI_Close', 'DJI_Volume', 'DJI_Open', 'DJI_High', 'DJI_Low',
       'EGO_Close', 'EGO_Volume', 'EGO_Open', 'EGO_High', 'EGO_Low',
       'GDX_Close', 'GDX_Volume', 'GDX_Open', 'GDX_High', 'WTI_Close',
       'USB_High', 'USB_Low', 'EUR_Close', 'USD_Open', 'USD_Low', 'USD_Change',
       'SIL_Close', 'SIL_Open', 'SIL_High', 'SIL_Low', 'PLD_Change',
       'MACD_GLD', 'SIG_GLD', 'SMA_GLD', 'SMAUP_GLD', 'SMALO_GLD', 'RSI_GLD'],
      dtype='object')

Total features selected by Random Forest for 5 days: 38
Features for 5 days:
 Index(['GLD_Close', 'GLD_Open', 'GLD_High', 'SPX_Close', 'SPX_Open',
       'SPX_High', 'SPX_Low', 'DJI_Close', 'DJI_Volume', 'DJI_Open',
       'DJI_High', 'DJI_Low', 'EGO_Volume', 'EGO_Open', 'EGO_High',
       'GDX_Close', 'GDX_Volume', 'GDX_Low', 'WTI_Open', 'USO_Low', 'USB_Open',
       'EUR_High', 'USD_Close', 'USD_Open', 'USD_High', 'USD_Change',
       'SIL_Close', 'SIL_Open', 'SIL_High', 'SIL_Low', 'PLT_Open',
       'PLD_Change', 'MACD_GLD', 'SIG_GLD', 'SMA_GLD', 'SMAUP_GLD',
       'SMALO_GLD', 'RSI_GLD'],
      dtype='object')

Total features selected by Random Forest for 7 days: 38
Features for 7 days:
 Index(['GLD_Close', 'GLD_Open', 'GLD_High', 'GLD_Low', 'SPX_Close', 'SPX_Open',
       'SPX_High', 'SPX_Low', 'DJI_Volume', 'DJI_Open', 'DJI_High', 'DJI_Low',
       'EGO_Volume', 'EGO_Open', 'EGO_High', 'GDX_Low', 'WTI_Volume',
       'WTI_Open', 'WTI_High', 'USO_Low', 'USB_Close', 'USB_Open', 'USB_High',
       'EUR_Change', 'USD_Close', 'USD_High', 'USD_Change', 'SIL_Close',
       'SIL_Open', 'SIL_High', 'SIL_Change', 'PLD_Change', 'MACD_GLD',
       'SIG_GLD', 'SMA_GLD', 'SMAUP_GLD', 'SMALO_GLD', 'RSI_GLD'],
      dtype='object')
```

ภาพประกอบ 46 แสดงคุณลักษณะที่ถูกคัดเลือกจากการถดถอยของป่าแบบสุ่มของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

การคัดเลือกคุณลักษณะโดยการค้นหาแบบขั้นตอนเชิงพันธุกรรมโดยใช้การเรียนรู้ของเครื่องแบบซัพพอร์ตเวกเตอร์รีเกสเป็นการเรียนรู้พื้นฐาน ซึ่งเทคนิคการค้นหาแบบขั้นตอนเชิงพันธุกรรมจะคัดเลือกคุณลักษณะที่มีค่าความแม่นยำเป็นเครื่องมือในการคัดลือกคุณลักษณะ ซึ่งงานวิจัยนี้ใช้ฟังก์ชัน GAFeatureSelectionCV จาก Sklearn Libarry โดยกระบวนการคัดเลือกคุณลักษณะจะคัดเลือกคุณลักษณะจากการขยับข้อมูลและแบ่งข้อมูลเรียบร้อยแล้ว ส่งผลให้ได้คุณลักษณะที่ถูกคัดเลือกตามภาพประกอบที่ 47 แสดงคุณลักษณะที่ถูกคัดเลือกจากการค้นหาแบบขั้นตอนเชิงพันธุกรรมของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

```
Total features selected by Random Forest for 1 days: 42
Features for 1 day:
 Index(['GLD_Close', 'SPX_Low', 'DJI_Volume', 'DJI_Open', 'DJI_High',
       'EGO_Close', 'EGO_High', 'EGO_Low', 'GDX_Volume', 'GDX_High', 'GDX_Low',
       'WTI_Close', 'WTI_Volume', 'WTI_Open', 'WTI_High', 'WTI_Low',
       'BNO_Close', 'BNO_Open', 'USO_Close', 'USO_Open', 'USO_High',
       'USB_Close', 'USB_High', 'USB_Low', 'USB_Change', 'EUR_Close',
       'EUR_Change', 'USD_Close', 'USD_High', 'SIL_Close', 'SIL_High',
       'SIL_Low', 'PLT_Close', 'PLT_Low', 'PLD_High', 'PLD_Low', 'PLD_Change',
       'MACD_GLD', 'SMA_GLD', 'SMAUP_GLD', 'STDEV_GLD', 'Open_Close'],
      dtype='object')
Total features selected by Random Forest for 3 days: 38
Features for 3 day:
 Index(['GLD_Close', 'GLD_Open', 'GLD_High', 'SPX_Close', 'SPX_High',
       'DJI_Close', 'DJI_Volume', 'DJI_Open', 'EGO_Close', 'EGO_Open',
       'EGO_High', 'EGO_Low', 'GDX_Volume', 'GDX_High', 'GDX_Low',
       'WTI_Volume', 'WTI_Open', 'WTI_Low', 'BNO_Close', 'BNO_Volume',
       'BNO_High', 'BNO_Low', 'USO_Volume', 'USO_Open', 'USO_High', 'USB_High',
       'EUR_Close', 'EUR_High', 'USD_Close', 'USD_Change', 'SIL_High',
       'SIL_Low', 'PLT_High', 'PLD_Low', 'PLD_Change', 'MACD_GLD', 'RSI_GLD',
       'STDEV_GLD'],
      dtype='object')
Total features selected by Random Forest for 5 days: 42
Features for 5 day:
 Index(['GLD_Close', 'GLD_High', 'GLD_Low', 'SPX_Open', 'SPX_Low', 'DJI_Close',
       'DJI_Volume', 'DJI_High', 'EGO_Close', 'EGO_Open', 'WTI_Volume',
       'WTI_Low', 'BNO_Close', 'BNO_Low', 'USO_Close', 'USO_Low', 'USB_Open',
       'USB_Low', 'USB_Change', 'EUR_Open', 'EUR_High', 'USD_Close',
       'USD_Open', 'USD_High', 'USD_Low', 'SIL_Close', 'SIL_Open', 'SIL_High',
       'SIL_Change', 'PLT_Close', 'PLT_Open', 'PLT_High', 'PLT_Low',
       'PLT_Change', 'PLD_Close', 'PLD_Open', 'PLD_High', 'PLD_Low',
       'PLD_Change', 'SIG_GLD', 'RSI_GLD', 'Open_Close'],
      dtype='object')
Total features selected by Random Forest for 7 days: 44
Features for 7 day:
 Index(['GLD_Close', 'GLD_Open', 'GLD_High', 'GLD_Low', 'SPX_Close', 'SPX_Open',
       'SPX_Low', 'DJI_Close', 'DJI_Volume', 'DJI_High', 'DJI_Low',
       'EGO_Close', 'EGO_Volume', 'EGO_High', 'GDX_Volume', 'GDX_Low',
       'WTI_Close', 'WTI_Volume', 'WTI_Open', 'WTI_Low', 'BNO_Close',
       'BNO_Open', 'USO_Close', 'USO_Volume', 'USO_Open', 'USB_Close',
       'USB_Open', 'USB_Low', 'USB_Change', 'EUR_Close', 'EUR_Open',
       'EUR_High', 'EUR_Change', 'USD_Open', 'USD_High', 'USD_Low', 'SIL_Open',
       'SIL_High', 'SIL_Change', 'PLT_Close', 'PLT_Open', 'PLD_High',
       'RSI_GLD', 'Open_Close'],
      dtype='object')
```

ภาพประกอบ 47 แสดงคุณลักษณะที่ถูกคัดเลือกจากการค้นหาแบบขั้นตอนเชิงพันธุกรรมของชุดข้อมูลแบบพยากรณ์ 1, 3, 5 และ 7 วัน ตามลำดับ

ตาราง 16 ตารางสรุปการคัดเลือกคุณลักษณะแต่ละเทคนิค

Feature Name	1 Day			3 Day			5 Day			7 Day		
	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic
GLD_Close	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
GLD_Open	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	
GLD_High		✓	✓		✓	✓		✓	✓		✓	
GLD_Low			✓			✓			✓		✓	
SPX_Close		✓	✓		✓	✓		✓	✓		✓	✓
SPX_Open	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SPX_High			✓					✓			✓	
SPX_Low		✓			✓	✓		✓			✓	✓
DJI_Close					✓	✓						✓
DJI_Volume		✓			✓	✓			✓		✓	✓

ตาราง 16 (ต่อ) ตารางสรุปการคัดเลือกคุณลักษณะแต่ละเทคนิค

Feature Name	1 Day			3 Day			5 Day			7 Day		
	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic
DJI_Open		✓	✓		✓	✓					✓	✓
DJI_High		✓			✓						✓	
DJI_Low		✓	✓		✓	✓			✓		✓	
EGO_Close					✓				✓		✓	
EGO_Volume		✓			✓			✓				✓
EGO_Open		✓			✓	✓		✓	✓		✓	
EGO_High		✓			✓	✓		✓	✓		✓	
EGO_Low		✓			✓				✓			
GDX_Close	✓	✓		✓	✓		✓	✓	✓	✓		
GDX_Volume		✓			✓			✓	✓			✓
GDX_Open		✓			✓				✓			
GDX_High		✓			✓	✓			✓			
GDX_Low						✓		✓	✓		✓	✓
WTI_Close					✓	✓		✓				
WTI_Volume						✓			✓		✓	✓
WTI_Open			✓						✓		✓	✓
WTI_High			✓								✓	
WTI_Low												✓
BNO_Close			✓		✓	✓						✓
BNO_Volume									✓			✓
BNO_Open			✓			✓						✓
BNO_High						✓						
BNO_Low			✓			✓			✓			
USO_Close												✓
USO_Volume			✓									✓
USO_Open						✓						✓
USO_High	✓		✓	✓			✓			✓		
USO_Low			✓					✓			✓	✓
USB_Close		✓	✓			✓					✓	✓
USB_Open		✓						✓	✓		✓	✓

ตาราง 16 (ต่อ) ตารางสรุปการคัดเลือกคุณลักษณะแต่ละเทคนิค

Feature Name	1 Day			3 Day			5 Day			7 Day		
	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic
USB_High		✓									✓	✓
USB_Low		✓			✓							✓
USB_Change						✓						✓
EUR_Close					✓							
EUR_Open						✓			✓			
EUR_High			✓			✓		✓				✓
EUR_Low		✓	✓			✓						
EUR_Change	✓		✓	✓			✓			✓	✓	✓
USD_Close		✓						✓			✓	✓
USD_Open			✓		✓			✓	✓			✓
USD_High								✓	✓		✓	✓
USD_Low	✓	✓	✓	✓	✓	✓	✓			✓		✓
USD_Change		✓			✓	✓		✓	✓		✓	
SIL_Close	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓
SIL_Open		✓			✓			✓	✓		✓	✓
SIL_High		✓	✓		✓	✓		✓	✓		✓	✓
SIL_Low	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓
SIL_Change			✓								✓	
PLT_Close						✓			✓			✓
PLT_Open			✓					✓	✓			✓
PLT_High	✓			✓			✓			✓		✓
PLT_Low		✓										✓
PLT_Change					✓	✓		✓			✓	
PLD_Close			✓						✓			
PLD_Open									✓			✓
PLD_High						✓						
PLD_Low		✓	✓			✓						
PLD_Change		✓	✓			✓			✓			✓
MACD_GLD		✓	✓		✓	✓		✓			✓	✓
SIG_GLD	✓	✓		✓	✓		✓	✓		✓	✓	✓

ตาราง 16 (ต่อ) ตารางสรุปการคัดเลือกคุณลักษณะแต่ละเทคนิค

Feature Name	1 Day			3 Day			5 Day			7 Day		
	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic	Lasso	Random Forest	Genetic
SMA_GLD		✓			✓			✓	✓		✓	
SMAUP_GLD	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
SMALO_GLD		✓	✓		✓	✓		✓			✓	✓
RSI_GLD					✓			✓	✓		✓	✓
STDEV_GLD			✓			✓			✓			
Open_Close			✓			✓			✓			✓
High_Low												
Total	12	38	34	12	38	41	12	38	40	12	38	45

3.2.5. **การตั้งค่าและการฝึกฝนแบบจำลอง** (Model Setup and Training)

งานวิจัยการพยากรณ์ราคาทองคำใช้การเรียนรู้ของเครื่อง 8 อัลกอริทึม ได้แก่ ซัพพอร์ตเวกเตอร์รีเกสชัน, การถดถอยแบบต้นไม้ตัดสินใจ, การถดถอยของป่าแบบสุ่ม, การถดถอยลาสโซ, การถดถอยแบบริดจ์, การถดถอยของริดจ์แบบเบส์, การถดถอยแบบกราเดียนส์บูสติงและสโคแคสติกกราเดียนเดสเซนท์

3.2.5.1. **ซัพพอร์ตเวกเตอร์รีเกสชัน** (Support Vector Regression)

ซัพพอร์ตเวกเตอร์รีเกรสชันเป็นอัลกอริทึมการเรียนรู้ของเครื่องที่นิยมใช้สำหรับปัญหาการพยากรณ์และการถดถอย เนื่องจากความสามารถในการจัดการข้อมูลเชิงซ้อนและการปรับให้เหมาะสมในกรณีที่ข้อมูลมีมิติมาก อย่างไรก็ตาม ประสิทธิภาพของซัพพอร์ตเวกเตอร์รีเกรสชันขึ้นอยู่กับค่าพารามิเตอร์สำคัญ เช่น C, epsilon และ kernel การปรับปรุงค่าพารามิเตอร์เหล่านี้จึงเป็นกระบวนการที่สำคัญในการเพิ่มความแม่นยำของแบบจำลอง

หนึ่งในเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของซัพพอร์ตเวกเตอร์รีเกรสชัน ได้แก่ C, gamma, kernel และ epsilon ซึ่งค่าที่ใช้ในการการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 48 แสดงพารามิเตอร์ของ GridSearchCV สำหรับซัพพอร์ตเวกเตอร์รีเกรสชัน

```
# Define the parameter grid for SVR
param_grid = {
    'C': [0.1, 1, 10, 100],
    'gamma': ['scale', 'auto', 0.01, 0.1, 1],
    'epsilon': [0.01, 0.1, 1]
}
```

ภาพประกอบ 48 แสดงพารามิเตอร์ของ GridSearchCV สำหรับซัพพอร์ตเวกเตอร์รีเกรสชัน

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับค่าพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลแต่ละชุดข้อมูล ทั้งหมด 4 ชุดข้อมูล ได้แก่ ชุดข้อมูลต้นแบบ , ชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะจากการถดถอยลาสโซ, ชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะจากการถดถอยของป่าแบบสุ่ม และชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะจากการค้นหาแบบขั้นตอนเชิงพันธุกรรม โดยแต่ละชุดข้อมูลจะแบ่งชุดข้อมูลย่อยตามการพยากรณ์แบบ 1, 3, 5 และ 7 วันตามลำดับ ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตารางที่ 17 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับซัพพอร์ตเวกเตอร์รีเกรสชัน

ตาราง 17 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับซัพพอร์ตเวกเตอร์รีเกรสชัน

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
ต้นแบบ	C	100	10	10	10
	gamma	0.01	0.01	0.01	0.01
	epsilon	0.01	0.01	0.01	0.01
การถดถอยลาสโซ	C	100	10	10	10
	gamma	0.01	0.01	0.01	0.01
	epsilon	0.01	0.01		0.01
การถดถอยของป่าแบบสุ่ม	C	100	10	10	10
	gamma	0.01	0.01	0.01	0.01
	epsilon	0.01	0.01	0.01	0.01
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	C	100	10	10	10
	gamma	0.01	0.01	0.01	0.01
	epsilon	0.01	0.01	0.01	0.01

3.2.5.2. **การถดถอยแบบต้นไม้ตัดสินใจ** (Decision Tree Regression)

การถดถอยแบบต้นไม้ตัดสินใจเป็นอัลกอริทึมการเรียนรู้ของเครื่องที่นิยมใช้สำหรับปัญหาการพยากรณ์และการถดถอย เนื่องจากมีความสามารถในการจัดการข้อมูลที่ซับซ้อนและรองรับข้อมูลที่มีมิติสูงได้อย่างดี โดยอัลกอริทึมนี้ทำงานผ่านกระบวนการแบ่งข้อมูลออกเป็นกลุ่มย่อย เพื่อหาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรเป้าหมายอย่างไรก็ตาม ประสิทธิภาพของการถดถอยแบบต้นไม้ตัดสินใจ ขึ้นอยู่กับค่าพารามิเตอร์สำคัญ max_depth, min_samples_split, min_samples_leaf เป็นต้น การปรับปรุงค่าพารามิเตอร์เหล่านี้จึงเป็นกระบวนการที่สำคัญในการเพิ่มความแม่นยำของแบบจำลอง

หนึ่งในเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของการถดถอยแบบต้นไม้ตัดสินใจ ได้แก่ max_depth, min_samples_split และ min_samples_leaf ซึ่งค่าที่ใช้ในการการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 49 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบต้นไม้ตัดสินใจ

```
param_grid_dt = {
    'max_depth': [None, 10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}
```

ภาพประกอบ 49 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบต้นไม้ตัดสินใ

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับค่าพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลแต่ละชุดข้อมูล ทั้งหมด 4 ชุดข้อมูล ได้แก่ ชุดข้อมูลต้นแบบ , ชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะจากการถดถอยลาสโซ, ชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะจากการถดถอยของป่าแบบสุ่ม และชุดข้อมูลที่ผ่านการคัดเลือกคุณลักษณะจากการค้นหาแบบขั้นตอนเชิงพันธุกรรม โดยแต่ละชุดข้อมูลจะแบ่งชุดข้อมูลย่อยตามการพยากรณ์แบบ 1, 3, 5 และ 7 วันตามลำดับ ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตารางที่ 18 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบต้นไม้ตัดสินใจ

ตาราง 18 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบต้นไม้ตัดสินใจ

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
ต้นแบบ	max_dept	None	10	10	None
	min_samples_split	5	10	2	10
	min_samples_leaf	1	1	1	2
การถดถอยลาสโซ	max_dept	10	None	10	10
	min_samples_split	10	10	10	10
	min_samples_leaf	1	2	4	1
การถดถอยของป่าแบบสุ่ม	max_dept	10	None	10	20
	min_samples_split	5	2	2	2
	min_samples_leaf	2	4	4	1
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	max_dept	10	None	10	10
	min_samples_split	10	5	10	2
	min_samples_leaf	4	2	1	1

3.2.5.3. **การถดถอยของป่าแบบสุ่ม** (Random Forest Regression)

การถดถอยของป่าแบบสุ่มเป็นหนึ่งในอัลกอริทึมการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสำหรับงานพยากรณ์ค่าต่อเนื่องคุณสมบัติของการถดถอยของป่าแบบสุ่มสามารถจัดการข้อมูลที่ซับซ้อนและมีตัวแปรหลายตัวได้ดี รวมถึงความสามารถในการรับมือกับความไม่เชิงเส้นทำให้อัลกอริทึมนี้เป็นตัวเลือกที่เหมาะสมสำหรับการพยากรณ์ราคาปิดของกองทุนทองคำในสถานการณ์ที่มีความซับซ้อน

โดยเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของการถดถอยของป่าแบบสุ่ม ได้แก่ max_depth, min_samples_split, min_samples_leaf และ n_estimators ซึ่งค่าที่ใช้ในการการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 50 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยของป่าแบบสุ่ม

```
param_grid_rf = {
    'n_estimators': [50, 100, 200],
    'max_depth': [None, 10, 20],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}
```

ภาพประกอบ 50 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยของป่าแบบสุ่ม

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV โดยใช้วิธีการปรับค่าพารามิเตอร์ที่เหมาะสมกับชุดข้อมูลแต่ละชุดข้อมูล ทั้งหมด 4 ชุดข้อมูล โดยแต่ละชุดข้อมูลจะแบ่งชุดข้อมูลย่อยตามการพยากรณ์แบบ 1, 3, 5 และ 7 วันตามลำดับ ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตารางที่ 19 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของป่าแบบสุ่ม

ตาราง 19 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของป่าแบบสุ่ม

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
ต้นแบบ	max_dept	10	10	20	10
	min_samples_split	2	5	2	5
	min_samples_leaf	2	2	1	1
	n_estimators	50	50	50	50
การถดถอยลาสโซ	max_dept	None	10	None	10
	min_samples_split	2	10	10	2
	min_samples_leaf	1	2	4	1
	n_estimators	50	50	200	100
การถดถอยของป่าแบบสุ่ม	max_dept	None	10	10	10
	min_samples_split	2	10	10	5
	min_samples_leaf	1	2	4	1
	n_estimators	50	50	50	50
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	max_dept	10	None	20	10
	min_samples_split	2	10	2	5
	min_samples_leaf	2	2	1	2
	n_estimators	50	50	50	50

3.2.5.4. **การถดถอยลาสโซ** (Lasso Regression)

การถดถอยลาสโซเป็นเทคนิคการถดถอยเชิงเส้นที่ได้รับการพัฒนาขึ้นเพื่อลดปัญหา Multicollinearity และการ Overfitting โดยใช้การปรับค่าพารามิเตอร์ Regularization แบบ L1 ซึ่งสามารถเลือกตัวแปรที่สำคัญในการพยากรณ์ได้ งานวิจัยนี้นำเสนอการประยุกต์ใช้การถดถอยลาสโซ ในการพยากรณ์ราคาปิดของกองทุนทองคำ ซึ่งเป็นงานที่ต้องวิเคราะห์ข้อมูลเชิงซ้อนและปัจจัยที่มีผลกระทบหลายประการ เช่น ราคาน้ำมัน อัตราแลกเปลี่ยนเงินตรา และดัชนีเศรษฐกิจโลก

โดยเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของการถดถอยลาสโซ ได้แก่ alpha ซึ่งค่าที่ใช้ในการการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 51 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบริดจ์

```
param_grid_lasso = {
    'alpha': [0.001, 0.01, 0.1, 1, 10, 100]
}
```

ภาพประกอบ 51 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยลาสโซ

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับเช่นเดียวกับการถดถอยของป่าแบบสุ่ม ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตารางที่ 20 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยลาสโซ

ตาราง 20 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยลาสโซ

ชุดข้อมูล	**พารามิเตอร์**	**จำนวนวันในการพยากรณ์**			
		1	3	5	7
ต้นแบบ	alpha	0.001	0.001	0.001	0.001
การถดถอยลาสโซ	alpha	0.001	0.001	0.001	0.001
การถดถอยของป่าแบบสุ่ม	alpha	0.001	0.001	0.001	0.001
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	alpha	0.001	0.001	0.001	0.001

3.2.5.5. **การถดถอยแบบริดจ์** (Ridge Regression)

การถดถอยอบบริดจ์เป็นอัลกอริทึมการถดถอยเชิงเส้นที่พัฒนาเพื่อจัดการกับปัญหา Multicollinearity และการ Overfitting ซึ่งเป็นข้อจำกัดสำคัญในโมเดลถดถอยทั่วไป งานวิจัยนี้นำเสนอการประยุกต์ใช้การถดถอยอบบริดจ์ ในการพยากรณ์ราคาปิดของกองทุนทองคำ โดยเน้นถึงการแก้ปัญหาความซับซ้อนของข้อมูลและการประเมินผลกระทบจากตัวแปรหลายมิติที่ส่งผลต่อราคาทองคำ เช่น อัตราแลกเปลี่ยนค่าเงินดอลลาร์ ราคาน้ำมัน และดัชนีตลาดหุ้น

โดยเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของการถดถอยแบบริดจ์ ได้แก่ alpha ซึ่งค่าที่ใช้ในการการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 51 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบริดจ์

```
param_grid_ridge = {
    'alpha': [0.001, 0.01, 0.1, 1, 10, 100]
}
```

ภาพประกอบ 52 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบริดจ์

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับค่าพารามิเตอร์ที่เหมือนกับการถดถอยของป่าแบบสุ่ม ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตารางที่ 21 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบริดจ์

ตาราง 21 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบริดจ์

ชุดข้อมูล	**พารามิเตอร์**	**จำนวนวันในการพยากรณ์**			
		1	3	5	7
ต้นแบบ	alpha	0.01	0.1	1	1
การถดถอยลาสโซ	alpha	0.001	0.001	0.01	0.001
การถดถอยของป่าแบบสุ่ม	alpha	0.001	0.1	0.001	0.01
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	alpha	0.001	0.01	0.01	1

3.2.5.6. **การถดถอยของริดจ์แบบเบส์** (Bayesian Ridge Regression)

การถดถอยของริดจ์แบบเบส์เป็นหนึ่งในเทคนิคการถดถอยที่พัฒนาจากการถดถอยแบบริดจ์ โดยเพิ่มมุมมองทางสถิติในกรอบของกฎของเบย์ (Bayes' Theorem) เพื่อสร้างโมเดลที่สามารถจัดการกับความไม่แน่นอนในข้อมูลได้อย่างมีประสิทธิภาพ งานวิจัยนี้มุ่งเน้นการประยุกต์ใช้ การถดถอยของริดจ์แบบเบส์ ในการพยากรณ์ราคาปิดกองทุนทองคำ โดยพิจารณาความซับซ้อนของข้อมูล เช่น ความสัมพันธ์เชิงเส้นระหว่างตัวแปรต้นและความผันผวนของราคาทองคำ

โดยเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของการถดถอยแบบริดจ์ ได้แก่ alpha_1, alpha_2, lambda_1 และ lambda_2 ซึ่งค่าที่ใช้ในการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 53 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยของริดจ์แบบเบส์

```
param_grid_bayesian = {
    'alpha_1': [1e-6, 1e-5, 1e-4, 1e-3, 1e-2],
    'alpha_2': [1e-6, 1e-5, 1e-4, 1e-3, 1e-2],
    'lambda_1': [1e-6, 1e-5, 1e-4, 1e-3, 1e-2],
    'lambda_2': [1e-6, 1e-5, 1e-4, 1e-3, 1e-2]
}
```

ภาพประกอบ 53 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยของริดจ์แบบเบส์

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับค่าพารามิเตอร์ที่เหมือนกับการถดถอยของป่าแบบสุ่ม ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตาราง 22 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของริดจ์แบบเบส์

ตาราง 22 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของริดจ์แบบเบส์

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
ต้นแบบ	alpha_1	1e-06	1e-06	1e-06	1e-06
	alpha_2	0.001	0.001	0.001	0.001
	lambda_1	0.001	0.001	0.001	0.001
	lambda_2	1e-06	1e-06	1e-06	1e-06

ตาราง 22 (ต่อ) ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยของริดจ์แบบเบส์

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
การถดถอยลาสโซ	alpha_1	0.001	0.001	1e-06	0.001
	alpha_2	1e-06	1e-06	0.001	1e-06
	lambda_1	1e-06	1e-06	0.001	1e-06
	lambda_2	0.001	0.001	0.001	0.001
การถดถอยของป่าแบบสุ่ม	alpha_1	1e-06	1e-06	0.001	1e-06
	alpha_2	0.001	0.001	1e-06	0.001
	lambda_1	0.001	0.001	1e-06	0.001
	lambda_2	1e-06	1e-06	0.001	0.0001
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	alpha_1	0.001	1e-06	1e-06	1e-06
	alpha_2	1e-06	0.001	0.001	0.001
	lambda_1	1e-06	0.001	0.001	0.001
	lambda_2	0.001	1e-06	1e-06	1e-06

3.2.5.7. **การถดถอยแบบกราเดียนส์บูสติง** (Gradient Boosting Regression)

การถดถอยแบบกราเดียนส์บูสติงเป็นวิธีการของการเรียนรู้ที่ใช้การรวมโมเดลหลายๆ ตัวเพื่อเพิ่มความแม่นยำในการพยากรณ์ ซึ่งในที่นี้จะใช้ ต้นไม้ตัดสินใจหลายๆ ต้นไม้ในการเรียนรู้จากข้อผิดพลาดที่เกิดขึ้นในโมเดลก่อนหน้า กระบวนการนี้จะช่วยให้สามารถจับความสัมพันธ์เชิงซ้อนที่อาจจะไม่สามารถมองเห็นได้ด้วยโมเดลเชิงเส้นแบบดั้งเดิม เช่น การถดถอยเชิงเส้น

โดยเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของการถดถอยแบบริดจ์ ได้แก่ learning_rate, max_dept, min_samples_split, min_sample_leaf และ n_estimators ซึ่งค่าที่ใช้ในการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 54 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบกราเดียนส์บูสติง

```
param_grid_gb = {
    'n_estimators': [50, 100, 200],
    'learning_rate': [0.01, 0.1, 0.5],
    'max_depth': [3, 5, 7],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}
```

ภาพประกอบ 54 แสดงพารามิเตอร์ของ GridSearchCV สำหรับการถดถอยแบบกราเดียนส์บูสติง

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับค่าพารามิเตอร์ที่เหมือนกับการถดถอยของป่าแบบสุ่ม ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็นดังตาราง 23 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบกราเดียนส์บูสติง

ตาราง 23 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบกราเดียนส์บูสติง

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
ต้นแบบ	learning_rate	0.1	0.5	0.5	0.5
	max_dept	7	3	5	3
	min_samples_split	10	5	10	5
	min_samples_leaf	1	1	1	1
	n_estimators	100	50	100	200
การถดถอยลาสโซ	learning_rate	0.5	0.5	0.1	0.1
	max_dept	3	3	3	3
	min_samples_split	10	2	5	10
	min_samples_leaf	1	2	1	4
	n_estimators	100	200	200	200
การถดถอยของป่าแบบสุ่ม	learning_rate	0.1	0.5	0.1	0.5
	max_dept	7	3	3	3
	min_samples_split	10	5	2	5
	min_samples_leaf	2	2	4	2
	n_estimators	100	50	50	100

ตาราง 23 (ต่อ) ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับการถดถอยแบบกราเดียนส์บูสติง

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
การค้นหาแบบขั้นตอนเชิงพันธุกรรม (Genetic Algorithms)	learning_rate	0.1	0.5	0.1	0.1
	max_dept	7	3	3	3
	min_samples_split	10	2	10	2
	min_samples_leaf	4	1	2	2
	n_estimators	200	100	200	200

3.2.5.8. **สโคแคสติกกราเดียนเดสเซนท์** (Stochastic Gradient Decent)

สโตแคสติกกราเดียนเดสเซนท์เป็นเทคนิคหนึ่งในกระบวนการกราเดียนเดสเซนท์ (Gradient Descent) ที่ใช้ในการหาค่าต่ำสุดของฟังก์ชันค่าเสียหาย (Loss Function) โดยจะทำการปรับปรุงพารามิเตอร์ของโมเดลเพื่อให้ผลลัพธ์ที่ได้มีความแม่นยำที่สุด กระบวนการนี้ใช้ในการฝึกโมเดลที่มีจำนวนพารามิเตอร์หลายตัว

โดยเครื่องมือที่มีประสิทธิภาพสำหรับการค้นหาค่าพารามิเตอร์ที่เหมาะสมคือ GridSearchCV ซึ่งให้บริการโดย Sklearn Library ใน Python โดยงานวิจัยนี้ใช้พารามิเตอร์ของสโตแคสติกกราเดียนเดสเซนท์ ได้แก่ alpha, eta0, learning_rate, loss และ penalty ซึ่งค่าที่ใช้ในการกำหนดพารามิเตอร์เป็นดังภาพประกอบที่ 55 แสดงพารามิเตอร์ของ GridSearchCV สำหรับสโตแคสติกกราเดียนเดสเซนท์

```
param_grid_sgd = {
    'loss': ['squared_error', 'huber', 'epsilon_insensitive'],
    'penalty': ['l2', 'l1', 'elasticnet'],
    'alpha': [0.0001, 0.001, 0.01],
    'learning_rate': ['constant', 'optimal', 'invscaling'],
    'eta0': [0.01, 0.1, 1]
}
```

ภาพประกอบ 55 แสดงพารามิเตอร์ของ GridSearchCV สำหรับสโตแคสติกกราเดียนเดสเซนท์

จากการเลือกพารามิเตอร์ผ่านฟังก์ชัน GridSearchCV จะใช้วิธีการปรับค่าพารามิเตอร์ที่เหมือนกับการถดถอยของป่าแบบสุ่ม ซึ่งผลการค้นหาพารามิเตอร์ที่เหมาะสมเป็น

ดังตาราง 24 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับสโตแคสติกกราเดียนเดสเซนท์

ตาราง 24 ตารางแสดงผลการค้นหาพารามิเตอร์ที่เหมาะสมสำหรับสโตแคสติกกราเดียนเดสเซนท์

ชุดข้อมูล	พารามิเตอร์	จำนวนวันในการพยากรณ์			
		1	3	5	7
ต้นแบบ	alpha	0.0001	0.0001	0.001	0.001
	eta0	0.01	1	0.01	1
	learning_rate	optimal	invscaling	optimal	invscaling
	loss	huber	huber	huber	huber
	penalty	L1	elasticnet	elasticnet	L1
การถดถอยลาสโซ	alpha	0.0001	0.001	0.0001	0.001
	eta0	1	0.01	1	1
	learning_rate	invscaling	optimal	invscaling	invscaling
	loss	squared_error	huber	huber	huber
	penalty	elasticnet	L2	L2	L2
การถดถอยของป่าแบบสุ่ม	alpha	0.0001	0.0001	0.0001	0.001
	eta0	0.01	0.1	1	1
	learning_rate	optimal	constant	invscaling	invscaling
	loss	huber	huber	epsilon_insensitive	huber
	penalty	L1	L1	L1	L1
การค้นหาแบบขั้นตอนเชิงพันธุกรรม	alpha	0.0001	0.0001	0.001	0.001
	eta0	0.01	0.1	0.01	0.01
	learning_rate	optimal	constant	optimal	optimal
	loss	huber	huber	huber	huber
	penalty	L1	L1	L1	L1

3.2.6. **การประเมินผลแบบจำลอง** (Model Evaluation)

การประเมินผลแบบจำลอง ใช้เครื่องมือการวัดประสิทธิภาพของโมเดล 3 เครื่องมือ คือ ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง หรือ RMSE โดยค่าดังกล่าวหากมีค่าต่ำหรือเข้าใกล้ศูนย์ จะเป็นโมเดลที่มีความแม่นยำสูง ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ หรือ R^2 โดย

ค่าดังกล่าวหากมีค่าสูง หรือเข้าใกล้หนึ่ง จะเป็นโมเดลที่มีความแม่นยำสูง และค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย หรือ MAPE โดยค่าดังกล่าวแสดงเป็นร้อยละ โดยยิ่งต่ำแสดงว่ายิ่งมีความแม่นยำที่สูง

การเปรียบเทียบประสิทธิภาพจากการประเมินผลแบบจำลองเหล่านี้ จะทำการเปรียบเทียบเฉพาะในกลุ่มการพยากรณ์เดียวกัน คือในการพยากรณ์ข้อมูลแบบล่วงหน้า 1 วัน จะเปรียบเทียบเฉพาะในกลุ่มข้อมูลแบบล่วงหน้า 1 วันที่ได้รับจากการเรียนรู้ของเครื่องในเทคนิคต่างๆ เป็นต้น

3.2.7. **การบวนการใช้เทคนิคการรวมโมเดล** (Ensemble Algorithms)

ในกระบวนการประเมินผลและพัฒนาแบบจำลองด้วยเทคนิคการรวมโมเดล งานวิจัยนี้ได้นำเสนอการใช้เทคนิคการรวมโมเดล 2 รูปแบบ ได้แก่ Voting และ Stacking โดยพิจารณาเลือกแบบจำลองจากลำดับของค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย หรือ MAPE ที่ได้จากการประเมินผลเบื้องต้น โดยคัดเลือกเฉพาะแบบจำลองที่มีค่า MAPE ต่ำที่สุด 4 อันดับแรก

และในการรวมโมเดล จะเริ่มต้นจากการรวมโมเดลที่มีประสิทธิภาพสูงสุดตามลำดับ เช่น การรวมแบบจำลองจำนวน 2 โมเดล จะประกอบด้วยแบบจำลองที่มีค่า MAPE ต่ำที่สุดอันดับที่ 1 และอันดับที่ 2 ส่วนการรวมแบบจำลองจำนวน 3 โมเดล จะประกอบด้วยแบบจำลองที่มีค่า MAPE ต่ำที่สุดอันดับที่ 1, 2 และ 3 ตามลำดับ และดำเนินการในลักษณะเดียวกันสำหรับจำนวนแบบจำลองที่เพิ่มขึ้น เพื่อเปรียบเทียบและประเมินผลลัพธ์ที่ได้จากการรวมโมเดลแต่ละกรณี

บทที่ 4
ผลการดำเนินการวิจัย

งานวิจัยฉบับนี้มุ่งเน้นการพัฒนาแบบจำลองเพื่อการพยากรณ์ราคาปิดของกองทุนทองคำ โดยอาศัยเทคนิคการเรียนรู้ของเครื่องซึ่งประกอบด้วยกระบวนการคัดเลือกคุณลักษณะเพื่อระบุชุดข้อมูลที่มีอิทธิพลสูงสุดต่อผลลัพธ์ และกระบวนการพยากรณ์ราคาปิดของกองทุนทองคำด้วยแบบจำลองที่ได้รับการฝึกฝน โดยในการประเมินประสิทธิภาพของแบบจำลอง จะใช้ตัวชี้วัด 3 เครื่องมือ ได้แก่ ค่ารากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ และค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย เพื่อเปรียบเทียบระดับความแม่นยำและประสิทธิภาพของแต่ละแบบจำลองที่พัฒนา ซึ่งรูปแบบการพยากรณ์ในงานวิจัยนี้แบ่งออกเป็น 4 ระยะเวลา ได้แก่ การพยากรณ์ล่วงหน้า 1 วัน 3 วัน 5 วัน และ 7 วัน ตามลำดับ ทั้งนี้ ได้มีการจัดทำแบบจำลองในแต่ละช่วงระยะเวลาพยากรณ์โดยอิงตามชุดคุณลักษณะที่ผ่านการคัดเลือกแล้ว ผ่านเทคนิคการเรียนรู้ของเครื่อง เพื่อเพิ่มความแม่นยำและประสิทธิผลของแบบจำลองในการพยากรณ์ราคาปิดของกองทุนทองคำอย่างมีประสิทธิภาพสูงสุด

4.1. การพยากรณ์แบบล่วงหน้า 1 วัน

การพยากรณ์โดยใช้การคัดเลือกคุณลักษณะผ่านเรียนรู้ของเครื่องในเทคนิคต่างๆ ได้ผลการทำงานดังนี้

4.1.1. Support Vector Machine

จากการทดลองโดยใช้แบบจำลอง SVR ภายใต้ชุดคุณลักษณะที่แตกต่างกัน พบว่าเทคนิคการเลือกคุณลักษณะมีผลโดยตรงต่อประสิทธิภาพของแบบจำลอง ทั้งในด้าน ความแม่นยำในการพยากรณ์ และ ความสามารถในการอธิบายความแปรปรวนของข้อมูล โดยสามารถสรุปผลการเปรียบเทียบได้ดังนี้

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.9076 ให้ค่า RMSE อยู่ที่ 2.2826 และให้ค่า MAPE อยู่ที่ 0.1797 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.9878 ให้ค่า RMSE อยู่ที่ 0.8284 และให้ค่า MAPE อยู่ที่ 0.0055 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.9475 ให้ค่า RMSE อยู่ที่ 1.7204 และให้ค่า MAPE อยู่ที่ 0.0127 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่

0.9867 ให้ค่า RMSE อยู่ที่ 0.8628 และให้ค่า MAPE อยู่ที่ 0.0058 ดังตาราง 25 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 25 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9076	2.2826	0.1797
Lasso Selection	0.9878	0.8284	0.0055
Random Forest Selection	0.9475	1.7204	0.0127
Genetic Selection	0.9867	0.8628	0.0058

4.1.2. Decision Tree Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.7721 ให้ค่า RMSE อยู่ที่ 3.5851 และให้ค่า MAPE อยู่ที่ 0.0210 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.8844 ให้ค่า RMSE อยู่ที่ 2.5529 และให้ค่า MAPE อยู่ที่ 0.0148 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.8738 ให้ค่า RMSE อยู่ที่ 2.6671 และให้ค่า MAPE อยู่ที่ 0.0167 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.7439 ให้ค่า RMSE อยู่ที่ 3.8801 และให้ค่า MAPE อยู่ที่ 0.0216 ดังตาราง 26 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 26 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.7721	3.5851	0.0210
Lasso Selection	0.8844	2.5529	0.0148
Random Forest Selection	0.8738	2.6671	0.0167
Genetic Selection	0.7439	3.8001	0.0216

4.1.3. Random Forest Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.8338 ให้ค่า RMSE อยู่ที่ 3.0607 และให้ค่า MAPE อยู่ที่ 0.0175 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.9227 ให้ค่า RMSE อยู่ที่ 2.0190 และให้ค่า MAPE อยู่ที่ 0.0117 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.8782 ให้ค่า RMSE อยู่ที่ 2.6205 และให้ค่า MAPE อยู่ที่ 0.0152 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.8571 ให้ค่า RMSE อยู่ที่ 2.8379 และให้ค่า MAPE อยู่ที่ 0.0156 ดังตาราง 27 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 27 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8338	3.0607	0.0175
Lasso Selection	0.9277	2.0190	0.0117
Random Forest Selection	0.8782	2.6205	0.0152
Genetic Selection	0.8571	2.8379	0.0156

4.1.4. Lasso Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9612 ให้ค่า RMSE อยู่ที่ 1.4774 และให้ค่า MAPE อยู่ที่ 0.0109 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9607 ให้ค่า RMSE อยู่ที่ 1.4871 และให้ค่า MAPE อยู่ที่ 0.0110 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9696 ให้ค่า RMSE อยู่ที่ 1.3090 และให้ค่า MAPE อยู่ที่ 0.0094 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9696 ให้ค่า RMSE อยู่ที่ 1.3090 และให้ค่า MAPE อยู่ที่ 0.0094 ดังตาราง 28 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 28 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9612	1.4774	0.0109
Lasso Selection	0.9607	1.4871	0.0110
Random Forest Selection	0.9696	1.3090	0.0094
Genetic Selection	0.9696	1.3090	0.0094

4.1.5. Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9901 ให้ค่า RMSE อยู่ที่ 0.7459 และให้ค่า MAPE อยู่ที่ 0.0048 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9627 ให้ค่า RMSE อยู่ที่ 1.4871 และให้ค่า MAPE อยู่ที่ 0.0110 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9612 ให้ค่า RMSE อยู่ที่ 1.4774 และให้ค่า MAPE อยู่ที่ 0.0109 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9696 ให้ค่า RMSE อยู่ที่ 1.3090 และให้ค่า MAPE อยู่ที่ 0.0094 ดังตาราง 29 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 29 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9901	0.7459	0.0048
Lasso Selection	0.9627	1.4871	0.0110
Random Forest Selection	0.9612	1.4774	0.0109
Genetic Selection	0.9696	1.3090	0.0094

4.1.6. Baysian Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9933 ให้ค่า RMSE อยู่ที่ 0.6108 และให้ค่า MAPE อยู่ที่ 0.0038 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9914 ให้ค่า RMSE อยู่ที่ 0.6928 และให้ค่า MAPE อยู่ที่ 0.0042 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9940 ให้ค่า RMSE อยู่ที่ 0.5795 และให้ค่า MAPE อยู่ที่ 0.0036 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9932 ให้ค่า RMSE อยู่ที่ 0.6179 และให้ค่า MAPE อยู่ที่ 0.0038 ดังตาราง 30 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 30 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9933	0.6108	0.0038
Lasso Selection	0.9914	0.6928	0.0042
Random Forest Selection	0.9940	0.5795	0.0036
Genetic Selection	0.9932	0.6179	0.0038

4.1.7. Gradient Boosting Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.8578 ให้ค่า RMSE อยู่ที่ 2.8313 และให้ค่า MAPE อยู่ที่ 0.0158 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.9216 ให้ค่า RMSE อยู่ที่ 2.1020 และให้ค่า MAPE อยู่ที่ 0.0125 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.9123 ให้ค่า RMSE อยู่ที่ 2.2229 และให้ค่า MAPE อยู่ที่ 0.0129 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.8622 ให้ค่า RMSE อยู่ที่ 2.7867 และให้ค่า MAPE อยู่ที่ 0.0154 ดังตาราง 31 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 31 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8578	2.8313	0.0158
Lasso Selection	0.9216	2.1020	0.0125
Random Forest Selection	0.9123	2.2229	0.0129
Genetic Selection	0.8622	2.7867	0.0154

4.1.8. Stochastic Gradient Descent

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9813 ให้ค่า RMSE อยู่ที่ 1.0265 และให้ค่า MAPE อยู่ที่ 0.0067 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9813 ให้ค่า RMSE อยู่ที่ 1.0265 และให้ค่า MAPE อยู่ที่ 0.0067 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9671 ให้ค่า RMSE อยู่ที่ 1.3619 และให้ค่า MAPE อยู่ที่ 0.0103 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9801 ให้ค่า RMSE อยู่ที่ 1.0579 และให้ค่า MAPE อยู่ที่ 0.0072 ดังตาราง 32 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 1 วัน

ตาราง 32 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 1 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9813	1.0265	0.0067
Lasso Selection	0.9813	1.0265	0.0067
Random Forest Selection	0.9671	1.3619	0.0103
Genetic Selection	0.9801	1.0579	0.0072

4.2. การพยากรณ์แบบล่วงหน้า 3 วัน

การพยากรณ์โดยใช้การคัดเลือกคุณลักษณะผ่านเรียนรู้ของเครื่องในเทคนิคต่างๆ ได้ผลการทำงานดังนี้

4.2.1. Support Vector Machine

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.9065 ให้ค่า RMSE อยู่ที่ 2.2963 และให้ค่า MAPE อยู่ที่ 0.0164 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.9446 ให้ค่า RMSE อยู่ที่ 1.7660 และให้ค่า MAPE อยู่ที่ 0.0119 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.8922 ให้ค่า RMSE อยู่ที่ 2.4649 และให้ค่า MAPE อยู่ที่ 0.0180 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.9077ให้ค่า RMSE อยู่ที่ 2.2814 และให้ค่า MAPE อยู่ที่ 0.0160 ดังตาราง 33 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 33 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9065	2.2963	0.0164
Lasso Selection	0.9446	1.7660	0.0119
Random Forest Selection	0.8922	2.4649	0.0180
Genetic Selection	0.9077	2.2814	0.0160

4.2.2. Decision Tree Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5314 ให้ค่า RMSE อยู่ที่ 5.1405 และให้ค่า MAPE อยู่ที่ 0.0324 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5306 ให้ค่า RMSE อยู่ที่ 5.1465 และให้ค่า MAPE อยู่ที่ 0.0331 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5770 ให้ค่า RMSE อยู่ที่ 4.8841 และให้ค่า MAPE อยู่ที่ 0.0310 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.6968

ให้ค่า RMSE อยู่ที่ 4.1348 และให้ค่า MAPE อยู่ที่ 0.0271 ดังตาราง 34 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 34 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.5314	5.1405	0.0324
Lasso Selection	0.5306	5.1465	0.0331
Random Forest Selection	0.5770	4.8841	0.0310
Genetic Selection	0.6968	4.1348	0.0271

4.2.3. Random Forest Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.6013 ให้ค่า RMSE อยู่ที่ 4.7419 และให้ค่า MAPE อยู่ที่ 0.0303 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.7453 ให้ค่า RMSE อยู่ที่ 3.7899 และให้ค่า MAPE อยู่ที่ 0.0246 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.5937 ให้ค่า RMSE อยู่ที่ 4.7866 และให้ค่า MAPE อยู่ที่ 0.0308 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.6907 ให้ค่า RMSE อยู่ที่ 4.1762 และให้ค่า MAPE อยู่ที่ 0.0269 ดังตาราง 35 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 35 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.6013	4.7419	0.0303
Lasso Selection	0.7453	3.7899	0.0246
Random Forest Selection	0.5937	4.7866	0.0308
Genetic Selection	0.6907	4.1762	0.0269

4.2.4. Lasso Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9266 ให้ค่า RMSE อยู่ที่ 2.0344 และให้ค่า MAPE อยู่ที่ 0.0142 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9336 ให้ค่า RMSE อยู่ที่ 1.9346 และให้ค่า MAPE อยู่ที่ 0.0132 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9266 ให้ค่า RMSE อยู่ที่ 2.0341 และให้ค่า MAPE อยู่ที่ 0.0141 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.9266 ให้ค่า RMSE อยู่ที่ 2.0341 และให้ค่า MAPE อยู่ที่ 0.0141 ดังตาราง 36 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 36 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9266	2.0344	0.0142
Lasso Selection	0.9336	1.9346	0.0132
Random Forest Selection	0.9266	2.0341	0.0141
Genetic Selection	0.9266	2.0341	0.0141

4.2.5. Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9220 ให้ค่า RMSE อยู่ที่ 2.0972 และให้ค่า MAPE อยู่ที่ 0.0144 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9336 ให้ค่า RMSE อยู่ที่ 1.9346 และให้ค่า MAPE อยู่ที่ 0.0132 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9266 ให้ค่า RMSE อยู่ที่ 2.0343 และให้ค่า MAPE อยู่ที่ 0.0142 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.9266 ให้ค่า RMSE อยู่ที่ 2.0341 และให้ค่า MAPE อยู่ที่ 0.0141 ดังตาราง 37 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 37 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9220	2.0972	0.0144
Lasso Selection	0.9336	1.9346	0.0132
Random Forest Selection	0.9266	2.0343	0.0142
Genetic Selection	0.9266	2.0341	0.0141

4.2.6. Baysian Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9225 ให้ค่า RMSE อยู่ที่ 2.0902 และให้ค่า MAPE อยู่ที่ 0.0142 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9310 ให้ค่า RMSE อยู่ที่ 1.9726 และให้ค่า MAPE อยู่ที่ 0.0135 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9483 ให้ค่า RMSE อยู่ที่ 1.7061 และให้ค่า MAPE อยู่ที่ 0.0115 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.9310 ให้ค่า RMSE อยู่ที่ 1.9719 และให้ค่า MAPE อยู่ที่ 0.0134 ดังตาราง 38 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 38 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9225	2.0902	0.0142
Lasso Selection	0.9310	1.9726	0.0135
Random Forest Selection	0.9483	1.7061	0.0115
Genetic Selection	0.9310	1.9719	0.0134

4.2.7. Gradient Boosting Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.6560 ให้ค่า RMSE อยู่ที่ 4.4040 และให้ค่า MAPE อยู่ที่ 0.0322 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.6682 ให้ค่า RMSE อยู่ที่ 4.3253 และให้ค่า MAPE อยู่ที่ 0.0320 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.7823 ให้ค่า RMSE อยู่ที่ 3.5037 และให้ค่า MAPE อยู่ที่ 0.0240 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.6410 ให้ค่า RMSE อยู่ที่ 4.4992 และให้ค่า MAPE อยู่ที่ 0.0329 ดังตาราง 39 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 39 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.6560	4.4040	0.0322
Lasso Selection	0.6682	4.3253	0.0320
Random Forest Selection	0.7823	3.5037	0.0240
Genetic Selection	0.6410	4.4992	0.0329

4.2.8. Stochastic Gradient Descent

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9203 ให้ค่า RMSE อยู่ที่ 2.1194 และให้ค่า MAPE อยู่ที่ 0.0147 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9203 ให้ค่า RMSE อยู่ที่ 2.1194 และให้ค่า MAPE อยู่ที่ 0.0147 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9313 ให้ค่า RMSE อยู่ที่ 1.9678 และให้ค่า MAPE อยู่ที่ 0.0138 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.9398 ให้ค่า RMSE อยู่ที่ 1.8421 และให้ค่า MAPE อยู่ที่ 0.0124 ดังตาราง 40 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 3 วัน

ตาราง 40 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 3 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.9203	2.1194	0.0147
Lasso Selection	0.9203	2.1194	0.0147
Random Forest Selection	0.9313	1.9678	0.0138
Genetic Selection	0.9398	1.8421	0.0124

4.3. การพยากรณ์แบบล่วงหน้า 5 วัน

การพยากรณ์โดยใช้การคัดเลือกคุณลักษณะผ่านเรียนรู้ของเครื่องในเทคนิคต่างๆ ได้ผลการทำงานดังนี้

4.3.1. Support Vector Machine

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.8425 ให้ค่า RMSE อยู่ที่ 2.9802 และให้ค่า MAPE อยู่ที่ 0.0204 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.8708 ให้ค่า RMSE อยู่ที่ 2.6988 และให้ค่า MAPE อยู่ที่ 0.0191 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.8724 ให้ค่า RMSE อยู่ที่ 2.6815 และให้ค่า MAPE อยู่ที่ 0.0186 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.7763 ให้ค่า RMSE อยู่ที่ 3.5516 และให้ค่า MAPE อยู่ที่ 0.0257 ดังตาราง 41 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 41 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8425	2.9802	0.0204
Lasso Selection	0.8708	2.6988	0.0191
Random Forest Selection	0.8724	2.6815	0.0186
Genetic Selection	0.7763	3.5516	0.0257

4.3.2. Decision Tree Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5926 ให้ค่า RMSE อยู่ที่ 4.7931 และให้ค่า MAPE อยู่ที่ 0.3332 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.6823 ให้ค่า RMSE อยู่ที่ 4.2327 และให้ค่า MAPE อยู่ที่ 0.0295 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5213 ให้ค่า RMSE อยู่ที่ 5.1954 และให้ค่า MAPE อยู่ที่ 0.0357 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.4602 ให้ค่า RMSE อยู่ที่ 5.5175 และให้ค่า MAPE อยู่ที่ 0.0373 ดังตาราง 42 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 42 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.5926	4.7931	0.3332
Lasso Selection	0.6823	4.2327	0.0295
Random Forest Selection	0.5213	5.1954	0.0357
Genetic Selection	0.4602	5.5175	0.0373

4.3.3. Random Forest Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.5682 ให้ค่า RMSE อยู่ที่ 4.9344 และให้ค่า MAPE อยู่ที่ 0.0332 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.6514 ให้ค่า RMSE อยู่ที่ 4.4340 และให้ค่า MAPE อยู่ที่ 0.0303 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.5755 ให้ค่า RMSE อยู่ที่ 4.8328 และให้ค่า MAPE อยู่ที่ 0.0329 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.6070 ให้ค่า RMSE อยู่ที่ 4.7074 และให้ค่า MAPE อยู่ที่ 0.0317 ดังตาราง 35 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 43 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.5682	4.9344	0.0332
Lasso Selection	0.6514	4.4340	0.0303
Random Forest Selection	0.5755	4.8928	0.0329
Genetic Selection	0.6070	4.7074	0.0317

4.3.4. Lasso Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8894 ให้ค่า RMSE อยู่ที่ 2.4965 และให้ค่า MAPE อยู่ที่ 0.0172 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8949 ให้ค่า RMSE อยู่ที่ 2.4336 และให้ค่า MAPE อยู่ที่ 0.0166 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8897 ให้ค่า RMSE อยู่ที่ 2.4940 และให้ค่า MAPE อยู่ที่ 0.0172 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8897 ให้ค่า RMSE อยู่ที่ 2.4940 และให้ค่า MAPE อยู่ที่ 0.0172 ดังตาราง 44 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 44 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8894	2.4965	0.0172
Lasso Selection	0.8949	2.4336	0.0166
Random Forest Selection	0.8897	2.4940	0.0172
Genetic Selection	0.8897	2.4940	0.0172

4.3.5. Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8203 ให้ค่า RMSE อยู่ที่ 3.1834 และให้ค่า MAPE อยู่ที่ 0.0220 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8494 ให้ค่า RMSE อยู่ที่ 2.4336 และให้ค่า MAPE อยู่ที่ 0.0166 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8894 ให้ค่า RMSE อยู่ที่ 2.4965 และให้ค่า MAPE อยู่ที่ 0.0172 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8897 ให้ค่า RMSE อยู่ที่ 2.4940 และให้ค่า MAPE อยู่ที่ 0.0172 ดังตาราง 45 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 45 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8203	3.1834	0.0220
Lasso Selection	0.8494	2.4336	0.0166
Random Forest Selection	0.8894	2.4965	0.0172
Genetic Selection	0.8897	2.4940	0.0172

4.3.6. Baysian Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.8169 ให้ค่า RMSE อยู่ที่ 3.2126 และให้ค่า MAPE อยู่ที่ 0.0222 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.8373 ให้ค่า RMSE อยู่ที่ 3.0290 และให้ค่า MAPE อยู่ที่ 0.0213 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.8804 ให้ค่า RMSE อยู่ที่ 2.5970 และให้ค่า MAPE อยู่ที่ 0.0181 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.8279 ให้ค่า RMSE อยู่ที่ 3.1146 และให้ค่า MAPE อยู่ที่ 0.0217 ดังตาราง 46 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 46 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8169	3.2126	0.0222
Lasso Selection	0.8373	3.0290	0.0213
Random Forest Selection	0.8804	2.5970	0.0181
Genetic Selection	0.8279	3.1146	0.0217

4.3.7. Gradient Boosting Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.5879 ให้ค่า RMSE อยู่ที่ 4.8205 และให้ค่า MAPE อยู่ที่ 0.0334 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.7017 ให้ค่า RMSE อยู่ที่ 4.1015 และให้ค่า MAPE อยู่ที่ 0.0296 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.6373 ให้ค่า RMSE อยู่ที่ 4.5223 และให้ค่า MAPE อยู่ที่ 0.0301 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.6816 ให้ค่า RMSE อยู่ที่ 4.2370 และให้ค่า MAPE อยู่ที่ 0.0306 ดังตาราง 47 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 47 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.5879	4.8205	0.0334
Lasso Selection	0.7017	4.1015	0.0296
Random Forest Selection	0.6373	4.5223	0.0301
Genetic Selection	0.6816	4.2370	0.0306

4.3.8. Stochastic Gradient Descent

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.8612 ให้ค่า RMSE อยู่ที่ 2.7970 และให้ค่า MAPE อยู่ที่ 0.0200 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.8612 ให้ค่า RMSE อยู่ที่ 2.7970 และให้ค่า MAPE อยู่ที่ 0.0200 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.8823 ให้ค่า RMSE อยู่ที่ 2.5754 และให้ค่า MAPE อยู่ที่ 0.0183 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.8815 ให้ค่า RMSE อยู่ที่ 2.5847 และให้ค่า MAPE อยู่ที่ 0.0178 ดังตาราง 48 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 5 วัน

ตาราง 48 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 5 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8612	2.7970	0.0200
Lasso Selection	0.8612	2.7970	0.0200
Random Forest Selection	0.8823	2.5754	0.0183
Genetic Selection	0.8815	2.5847	0.0178

4.4. การพยากรณ์แบบล่วงหน้า 7 วัน

4.4.1. Support Vector Machine

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.7349 ให้ค่า RMSE อยู่ที่ 8.8665 และให้ค่า MAPE อยู่ที่ 0.0270 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.2085 ให้ค่า RMSE อยู่ที่ 6.6811 และให้ค่า MAPE อยู่ที่ 0.0528 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่ 0.7794 ให้ค่า RMSE อยู่ที่ 3.5269 และให้ค่า MAPE อยู่ที่ 0.0249 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SVR ให้ค่า R^2 อยู่ที่

0.0148 ให้ค่า RMSE อยู่ที่ 7.4540 และให้ค่า MAPE อยู่ที่ 0.0556 ดังตาราง 49 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 49 ตารางแสดงผลการทดสอบแบบจำลองของ Support Vector Regression (SVR) แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.7349	8.8665	0.0270
Lasso Selection	0.2085	6.6811	0.0528
Random Forest Selection	0.7794	3.5269	0.0249
Genetic Selection	0.0148	7.4540	0.0556

4.4.2. Decision Tree Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.6706 ให้ค่า RMSE อยู่ที่ 4.3099 และให้ค่า MAPE อยู่ที่ 0.0313 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5295 ให้ค่า RMSE อยู่ที่ 5.1511 และให้ค่า MAPE อยู่ที่ 0.0360 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.5931 ให้ค่า RMSE อยู่ที่ 4.1598 และให้ค่า MAPE อยู่ที่ 0.0306 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง DT ให้ค่า R^2 อยู่ที่ 0.3177 ให้ค่า RMSE อยู่ที่ 6.2031 และให้ค่า MAPE อยู่ที่ 0.0492 ดังตาราง 50 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 50 ตารางแสดงผลการทดสอบแบบจำลองของ Decision Tree Regression (DT) แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.6706	4.3099	0.0313
Lasso Selection	0.5295	5.1511	0.0360
Random Forest Selection	0.6931	4.1598	0.0306
Genetic Selection	0.3177	6.2031	0.0492

4.4.3. Random Forest Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.5815 ให้ค่า RMSE อยู่ที่ 4.8580 และให้ค่า MAPE อยู่ที่ 0.0338 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.6136 ให้ค่า RMSE อยู่ที่ 4.6679 และให้ค่า MAPE อยู่ที่ 0.0327 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.6378 ให้ค่า RMSE อยู่ที่ 4.1594 และให้ค่า MAPE อยู่ที่ 0.0318 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง RF ให้ค่า R^2 อยู่ที่ 0.4733 ให้ค่า RMSE อยู่ที่ 5.4500 และให้ค่า MAPE อยู่ที่ 0.0372 ดังตาราง 51 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 51 ตารางแสดงผลการทดสอบแบบจำลองของ Random Forest Regression (RF) แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.5815	4.8580	0.0338
Lasso Selection	0.6136	4.6679	0.0327
Random Forest Selection	0.6378	4.1594	0.0318
Genetic Selection	0.4733	5.4500	0.0372

4.4.4. Lasso Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8492 ให้ค่า RMSE อยู่ที่ 2.9170 และให้ค่า MAPE อยู่ที่ 0.0201 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8032 ให้ค่า RMSE อยู่ที่ 3.3313 และให้ค่า MAPE อยู่ที่ 0.0233 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8454 ให้ค่า RMSE อยู่ที่ 2.9526 และให้ค่า MAPE อยู่ที่ 0.0204 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Lasso ให้ค่า R^2 อยู่ที่ 0.8454 ให้ค่า RMSE อยู่ที่ 2.9526และให้ค่า MAPE อยู่ที่ 0.0204 ดังตาราง 52 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 52 ตารางแสดงผลการทดสอบแบบจำลองของ Lasso Regression แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.8492	2.9170	0.0205
Lasso Selection	0.8032	3.3313	0.0233
Random Forest Selection	0.8454	2.9526	0.0204
Genetic Selection	0.8454	2.9526	0.0204

4.4.5. Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.7282 ให้ค่า RMSE อยู่ที่ 3.9149 และให้ค่า MAPE อยู่ที่ 0.0269 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8032 ให้ค่า RMSE อยู่ที่ 3.3313 และให้ค่า MAPE อยู่ที่ 0.0233 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8491 ให้ค่า RMSE อยู่ที่ 2.9172 และให้ค่า MAPE อยู่ที่ 0.0205 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง Ridge ให้ค่า R^2 อยู่ที่ 0.8454 ให้ค่า RMSE อยู่ที่ 2.9526 และให้ค่า MAPE อยู่ที่ 0.0204 ดังตาราง 53 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 53 ตารางแสดงผลการทดสอบแบบจำลองของ Ridge Regression แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.7282	3.9149	0.0269
Lasso Selection	0.8032	3.3313	0.0233
Random Forest Selection	0.8491	2.9172	0.0201
Genetic Selection	0.8454	2.9526	0.0204

4.4.6. Baysian Ridge Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.7313 ให้ค่า RMSE อยู่ที่ 3.8921 และให้ค่า MAPE อยู่ที่ 0.0268 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.7402 ให้ค่า RMSE อยู่ที่ 3.8276 และให้ค่า MAPE อยู่ที่ 0.0265 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.8103 ให้ค่า RMSE อยู่ที่ 3.2701 และให้ค่า MAPE อยู่ที่ 0.0227 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง BR ให้ค่า R^2 อยู่ที่ 0.6894 ให้ค่า RMSE อยู่ที่ 4.1850 และให้ค่า MAPE อยู่ที่ 0.0290 ดังตาราง 54 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 54 ตารางแสดงผลการทดสอบแบบจำลองของ Baysian Ridge Regression (BR) แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.7313	3.8921	0.0268
Lasso Selection	0.7402	3.8276	0.0265
Random Forest Selection	0.8103	3.2701	0.0227
Genetic Selection	0.6894	4.1850	0.0290

4.4.7. Gradient Boosting Regression

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.5029 ให้ค่า RMSE อยู่ที่ 5.2946 และให้ค่า MAPE อยู่ที่ 0.0372 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.4803 ให้ค่า RMSE อยู่ที่ 5.4135 และให้ค่า MAPE อยู่ที่ 0.0393 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.5709 ให้ค่า RMSE อยู่ที่ 4.9188 และให้ค่า MAPE อยู่ที่ 0.0355 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง GB ให้ค่า R^2 อยู่ที่ 0.6010 ให้ค่า RMSE อยู่ที่ 4.7431 และให้ค่า MAPE อยู่ที่ 0.0339 ดังตาราง 55 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 55 ตารางแสดงผลการทดสอบแบบจำลองของ Gradient Boosting Regression (GB) แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.5029	5.2946	0.0372
Lasso Selection	0.4803	5.4135	0.0393
Random Forest Selection	0.5709	4.9188	0.0355
Genetic Selection	0.6010	4.7431	0.0339

4.4.8. Stochastic Gradient Descent

เมื่อทดสอบด้วยชุดคุณลักษณะทั้งหมดโดยไม่ผ่านการคัดเลือก พบว่า แบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.7266 ให้ค่า RMSE อยู่ที่ 3.9264 และให้ค่า MAPE อยู่ที่ 0.0275 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วยเทคนิค Lasso Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.7266 ให้ค่า RMSE อยู่ที่ 3.9264 และให้ค่า MAPE อยู่ที่ 0.0275 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Random Forest Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.8296 ให้ค่า RMSE อยู่ที่ 3.0999 และให้ค่า MAPE อยู่ที่ 0.0216 และเมื่อทดสอบด้วยชุดคุณลักษณะที่ถูกคัดเลือกด้วย Genetic Selection พบว่าแบบจำลอง SGD ให้ค่า R^2 อยู่ที่ 0.8246 ให้ค่า RMSE อยู่ที่ 3.1443 และให้ค่า MAPE อยู่ที่ 0.0220 ดังตาราง 56 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 7 วัน

ตาราง 56 ตารางแสดงผลการทดสอบแบบจำลองของ Stochastic Gradient Descent (SGD) แบบพยากรณ์ล่วงหน้า 7 วัน

ชุดคุณลักษณะ	R^2	RMSE	MAPE
All Feature	0.7266	3.9264	0.0275
Lasso Selection	0.7266	3.9264	0.0275
Random Forest Selection	0.8296	3.0999	0.0216
Genetic Selection	0.8246	3.1443	0.0220

ตาราง 57 ตารางแสดงค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R^2) ของแต่ละแบบจำลอง

จำนวนวันการพยากรณ์	คุณลักษณะ	Support Vector Regression	Decision Tree Regression	Random Forest Regression	Lasso Regression	Ridge Regression	Baysian Ridge Regression	Gradient Boosting Regression	Stochastic Gradient Decent
1	All Feature	0.9076	0.7721	0.8338	0.9612	0.9901	0.9933	0.8578	0.9813
	Lasso Selection	0.9878	0.8844	0.9277	0.9607	0.9627	0.9914	0.9216	0.9813
	Random Forest Selection	0.9475	0.8738	0.8782	0.9696	0.9612	0.9940	0.9123	0.9671
	Genetic Selection	0.9867	0.7439	0.8571	0.9696	0.9696	0.9932	0.8622	0.9801
3	All Feature	0.9065	0.5314	0.6013	0.9266	0.9220	0.9225	0.6560	0.9203
	Lasso Selection	0.9446	0.5306	0.7453	0.9336	0.9336	0.9310	0.6682	0.9203
	Random Forest Selection	0.8922	0.5770	0.5937	0.9266	0.9266	0.9483	0.7823	0.9313
	Genetic Selection	0.9077	0.6968	0.6907	0.9266	0.9266	0.9310	0.6410	0.9398
5	All Feature	0.8425	0.5926	0.5682	0.8894	0.8203	0.8169	0.5879	0.8612
	Lasso Selection	0.8708	0.6823	0.6514	0.8949	0.8494	0.8373	0.7017	0.8612
	Random Forest Selection	0.8724	0.5213	0.5755	0.8897	0.8894	0.8804	0.6373	0.8823
	Genetic Selection	0.7763	0.4602	0.6070	0.8897	0.8897	0.8279	0.6816	0.8815
7	All Feature	0.7349	0.6706	0.5815	0.8492	0.7282	0.7313	0.5029	0.7266
	Lasso Selection	0.2085	0.5295	0.6136	0.8032	0.8032	0.7402	0.4803	0.7266
	Random Forest Selection	0.7794	0.6931	0.6378	0.8454	0.8491	0.8103	0.5709	0.8296
	Genetic Selection	0.0148	0.3177	0.4733	0.8454	0.8454	0.6894	0.6010	0.8246

ตาราง 58 ตารางแสดงค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ของแต่ละแบบจำลอง

จำนวนวันการพยากรณ์	คุณลักษณะ	Support Vector Regression	Decision Tree Regression	Random Forest Regression	Lasso Regression	Ridge Regression	Baysian Ridge Regression	Gradient Boosting Regression	Stochastic Gradient Decent
1	All Feature	2.2826	3.5851	3.0607	1.4774	0.7459	0.6108	2.8313	1.0265
	Lasso Selection	0.8284	2.5529	2.0190	1.4871	1.4871	0.6928	2.1020	1.0265
	Random Forest Selection	1.7204	2.6671	2.6205	1.3090	1.4774	0.5795	2.2229	1.3619
	Genetic Selection	0.8628	3.8001	2.8379	1.3090	1.3090	0.6179	2.7867	1.0579
3	All Feature	2.2963	5.1405	4.7419	2.0344	2.0972	2.0902	4.4040	2.1194
	Lasso Selection	1.7660	5.1465	3.7899	1.9346	1.9346	1.9726	4.3253	2.1194
	Random Forest Selection	2.4649	4.8841	4.7866	2.0341	2.0343	1.7061	3.5037	1.9678
	Genetic Selection	2.2814	4.1348	4.1762	2.0341	2.0341	1.9719	4.4992	1.8421
5	All Feature	2.9802	4.7931	4.9344	2.4965	3.1834	3.2126	4.8205	2.7970
	Lasso Selection	2.6988	4.2327	4.4340	2.4336	2.4336	3.0290	4.1015	2.7970
	Random Forest Selection	2.6815	5.1954	4.8928	2.4940	2.4965	2.5970	4.5223	2.5754
	Genetic Selection	3.5516	5.5175	4.7074	2.4940	2.4940	3.1146	4.2370	2.5847
7	All Feature	8.8665	4.3099	4.8580	2.9170	3.9149	3.8921	5.2946	3.9264
	Lasso Selection	6.6811	5.1511	4.6679	3.3313	3.3313	3.8276	5.4135	3.9264
	Random Forest Selection	3.5269	4.1598	4.1594	2.9526	2.9172	3.2701	4.9188	3.0999
	Genetic Selection	7.4540	6.2031	5.4500	2.9526	2.9526	4.1850	4.7431	3.1443

ตาราง 59 ตารางแสดงค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) ของแต่ละแบบจำลอง

จำนวนวันการพยากรณ์	คุณลักษณะ	Support Vector Regression	Decision Tree Regression	Random Forest Regression	Lasso Regression	Ridge Regression	Baysian Ridge Regression	Gradient Boosting Regression	Stochastic Gradient Decent
1	All Feature	0.1797	0.0210	0.0175	0.0109	0.0048	0.0038	0.0158	0.0067
	Lasso Selection	0.0055	0.0148	0.0117	0.0110	0.0110	0.0042	0.0125	0.0067
	Random Forest Selection	0.0127	0.0167	0.0152	**0.0094**	**0.0109**	**0.0036**	0.0129	**0.0103**
	Genetic Selection	0.0058	0.0216	0.0156	0.0094	0.0094	0.0038	0.0154	0.0072
3	All Feature	0.0164	0.0324	0.0303	0.0142	0.0144	0.0142	0.03228	0.0147
	Lasso Selection	0.0119	0.0331	0.0246	0.0132	0.0132	0.0135	0.0320	0.0147
	Random Forest Selection	0.0180	0.0310	0.0308	**0.0141**	**0.0142**	**0.0115**	0.0240	**0.0138**
	Genetic Selection	0.0160	0.0271	0.0269	0.0141	0.0141	0.0134	0.0329	0.0124
5	All Feature	0.0204	0.3332	0.0332	0.0172	0.0220	0.0222	0.0334	0.0200
	Lasso Selection	**0.0191**	0.0295	0.0303	**0.0167**	**0.0166**	0.0213	0.0296	**0.0200**
	Random Forest Selection	0.0186	0.0357	0.0329	0.0172	0.0172	0.0181	0.0301	0.0183
	Genetic Selection	0.0257	0.0373	0.0317	0.0172	0.0172	0.0217	0.0306	0.0178
7	All Feature	0.0270	0.0313	0.0338	0.0205	0.0269	0.0268	0.0372	0.0275
	Lasso Selection	0.0528	0.0360	0.0327	0.0233	0.0233	0.0265	0.0393	0.0275
	Random Forest Selection	0.0249	0.0306	0.0318	**0.0204**	**0.0201**	**0.0227**	0.0355	**0.0216**
	Genetic Selection	0.0556	0.0492	0.0372	0.0204	0.0204	0.0290	0.0339	0.0220

4.5. การรวมโมเดล

กระบวนการรวมโมเดลจะใช้วิธีการคัดเลือกอันดับของโมเดลจากค่า MAPE โดยอันดับการคัดเลือกโมเดลเป็นดังนี้

การพยากรณ์ข้อมูลล่วงหน้า 1 วัน ได้เลือกใช้โมเดลที่ผ่านกระบวนการคัดเลือกคุณลักษณะด้วยวิธี Random Forest Feature Selection โดยนำมาประยุกต์ใช้ร่วมกับอัลกอริทึมการเรียนรู้ของเครื่อง ได้แก่ Bayesian Ridge Regression, Stochastic Gradient Descent, Lasso Regression และ Ridge Regression ตามลำดับ ซึ่งเป็นเครื่องมือการเรียนรู้ของเครื่องที่ดีที่สุด 4 อันดับที่คัดเลือกจากค่า MAPE ของการทดสอบโมเดลในเบื้องต้น และใช้เทคนิคการรวมโมเดลแบบ Voting และ Stacking เพื่อค้นหาโมเดลที่ดีที่สุดในการพยากรณ์ล่วงหน้าแบบ 1 วัน

ตาราง 60 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 1 วัน

จำนวนวัน	การวัดผล	เทคนิคการรวมโมเดล	จำนวนโมเดล			
			1	2	3	4
1	R^2	Voting	0.9940	0.9870	0.9827	0.9786
		Stacking	0.9939	0.9939	**0.9941**	0.9939
	RMSE	Voting	0.5795	0.8558	0.9864	1.0977
		Stacking	0.5751	0.5818	**0.5850**	0.5851
	MAPE	Voting	0.0036	0.0058	0.0071	0.0080
		Stacking	0.0036	0.0036	**0.0035**	0.0036

จากตาราง 60 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 1 วัน พบว่าการใช้เทคนิคการรวมโมเดลแบบ Stacking ให้ผลการประเมินที่ดีกว่า Voting ในทุกเครื่องมือการวัดประสิทธิภาพ โดยให้ค่า R^2 อยู่ที่ 0.9941 ให้ค่า RMSE อยู่ที่ 0.5850 และให้ค่า MAPE อยู่ที่ 0.0035 ซึ่งแสดงให้เห็นว่าการใช้เทคนิคการรวมโมเดลสำหรับการทำนายล่วงหน้าแบบ 1 วัน ใช้เครื่องมือแบบ Stacking ได้ผลดีกว่าการใช้เครื่องมือแบบ Voting และเครื่องมือการรวมโมเดลแบบ Voting ยังมีแนวโน้มที่เพิ่มขึ้นเรื่อย ๆ ซึ่งอาจเกิดจากความซับซ้อนของโมเดลหรือการรวมผลที่ไม่สามารถเพิ่มคุณภาพของการพยากรณ์ได้

การพยากรณ์ข้อมูลล่วงหน้า 3 วัน ได้เลือกใช้โมเดลที่ผ่านกระบวนการคัดเลือกคุณลักษณะด้วยวิธี Random Forest Feature Selection โดยนำมาประยุกต์ใช้ร่วมกับอัลกอริทึมการเรียนรู้ของเครื่อง ได้แก่ Bayesian Ridge Regression, Stochastic Gradient Descent, Lasso Regression และ Ridge Regression ตามลำดับ ซึ่งเป็นเครื่องมือการเรียนรู้ของเครื่องที่ดีที่สุด 4 อันดับที่คัดเลือกจากค่า MAPE ของการทดสอบโมเดลในเบื้องต้น และใช้เทคนิคการรวมโมเดลแบบ Voting และ Stacking เพื่อค้นหาโมเดลที่ดีที่สุดในการพยากรณ์ล่วงหน้าแบบ 3 วัน

ตาราง 61 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 3 วัน

จำนวนวัน	การวัดผล	เทคนิคการรวมโมเดล	จำนวนโมเดล			
			1	2	3	4
3	R^2	Voting	0.9483	0.9469	0.9454	0.9427
		Stacking	0.9456	0.9460	**0.9514**	0.9514
	RMSE	Voting	1.7061	1.7295	1.7539	1.7973
		Stacking	1.7500	1.7449	**1.6543**	1.6543
	MAPE	Voting	0.0115	0.0119	0.0122	0.0125
		Stacking	0.0120	0.0119	**0.0111**	0.0111

จากตาราง 61 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 3 วัน พบว่าการใช้เทคนิคการรวมโมเดลแบบ Stacking ให้ผลการประเมินที่ดีกว่า Voting ในทุกเครื่องมือการวัดประสิทธิภาพ โดยให้ค่า R^2 อยู่ที่ 0.9514 ให้ค่า RMSE อยู่ที่ 1.6543 และให้ค่า MAPE อยู่ที่ 0.0111 ซึ่งแสดงให้เห็นว่าการใช้เทคนิคการรวมโมเดลสำหรับการทำนายล่วงหน้าแบบ 3 วัน ใช้เครื่องมือแบบ Stacking ได้ผลดีกว่าการใช้เครื่องมือแบบ Voting และเครื่องมือการรวมโมเดลแบบ Voting ยังมีแนวโน้มที่เพิ่มขึ้นเรื่อย ๆ ซึ่งอาจเกิดจากความซับซ้อนของโมเดลหรือการรวมผลที่ไม่สามารถเพิ่มคุณภาพของการพยากรณ์ได้ และการรวมโมเดลจำนวน 3 และ 4 โมเดล ให้ผลการประเมินประสิทธิภาพของโมเดลได้เท่ากัน ซึ่งเป็นผลการประเมินที่ดีที่สุดด้วย แสดงให้เห็นว่า สำหรับการรวมโมเดลแบบ Stacking การเพิ่มเข้ามาของการเรียนรู้ของเครื่องแบบ Ridge Regression ไม่ได้ส่งผลให้การทำงานเปลี่ยนแปลงไป

การพยากรณ์ข้อมูลล่วงหน้า 5 วัน ได้เลือกใช้โมเดลที่ผ่านกระบวนการคัดเลือกคุณลักษณะด้วยวิธี Lasso Selection โดยนำมาประยุกต์ใช้ร่วมกับอัลกอริทึมการเรียนรู้ของเครื่อง ได้แก่ Ridge Regression, Lasso Regression, Support Vector Machine และ Stochastic Gradient Descent ตามลำดับ ซึ่งเป็นเครื่องมือการเรียนรู้ของเครื่องที่ดีที่สุด 4 อันดับที่คัดเลือกจากค่า MAPE ของการทดสอบโมเดลในเบื้องต้น และใช้เทคนิคการรวมโมเดลแบบ Voting และ Stacking เพื่อค้นหาโมเดลที่ดีที่สุดในการพยากรณ์ล่วงหน้าแบบ 5 วัน

ตาราง 62 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 5 วัน

จำนวนวัน	การวัดผล	เทคนิคการรวมโมเดล	จำนวนโมเดล			
			1	2	3	4
5	R^2	Voting	0.8949	0.8949	**0.9042**	0.8995
		Stacking	0.8946	0.8866	0.8710	0.8656
	RMSE	Voting	2.4336	2.4336	**2.3234**	2.3806
		Stacking	2.4379	2.5288	2.6964	2.7528
	MAPE	Voting	0.0166	0.0166	**0.0161**	0.0167
		Stacking	0.0168	0.0176	0.0191	0.0195

จากตาราง 62 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 5 วัน พบว่าการใช้เทคนิคการรวมโมเดลแบบ Voting ให้ผลการประเมินที่ดีกว่า Stacking ในทุกเครื่องมือการวัดประสิทธิภาพ โดยให้ค่า R^2 อยู่ที่ 0.9042 ให้ค่า RMSE อยู่ที่ 2.3234 และให้ค่า MAPE อยู่ที่ 0.0161 ซึ่งแสดงให้เห็นว่าการใช้เทคนิคการรวมโมเดลสำหรับการทำนายล่วงหน้าแบบ 5 วัน ใช้เครื่องมือแบบ Voting ได้ผลดีกว่าการใช้เครื่องมือแบบ Stacking

การพยากรณ์ข้อมูลล่วงหน้า 7 วัน ได้เลือกใช้โมเดลที่ผ่านกระบวนการคัดเลือกคุณลักษณะด้วยวิธี Random Forest Feature Selection โดยนำมาประยุกต์ใช้ร่วมกับอัลกอริทึมการเรียนรู้ของเครื่อง ได้แก่ Ridge Regression, Lasso Regression, Stochastic Gradient Descent และ Bayesian Ridge Regression ตามลำดับ ซึ่งเป็นเครื่องมือการเรียนรู้ของเครื่องที่ดีที่สุด 4 อันดับที่คัดเลือกจากค่า MAPE ของการทดสอบโมเดลในเบื้องต้น และใช้เทคนิคการรวมโมเดลแบบ Voting และ Stacking เพื่อค้นหาโมเดลที่ดีที่สุดในการพยากรณ์ล่วงหน้าแบบ 7 วัน

ตาราง 63 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 7 วัน

จำนวนวัน	การวัดผล	เทคนิคการรวมโมเดล	จำนวนโมเดล			
			1	2	3	4
7	R^2	Voting	0.8491	0.8491	0.8437	**0.8586**
		Stacking	0.8467	0.8368	0.8423	0.8424
	RMSE	Voting	2.9170	2.9170	2.9688	**2.8236**
		Stacking	2.9394	3.0332	2.9818	2.9806
	MAPE	Voting	0.0201	0.0201	0.0205	**0.0196**
		Stacking	0.0200	0.0207	0.0203	0.0205

จากตาราง 63 ตารางแสดงผลการทดสอบแบบจำลอง Ensemble แบบพยากรณ์ล่วงหน้า 7 วัน พบว่าการใช้เทคนิคการรวมโมเดลแบบ Voting ให้ผลการประเมินที่ดีกว่า Stacking ในทุกเครื่องมือการวัดประสิทธิภาพ โดยให้ค่า R^2 อยู่ที่ 0.8586 ให้ค่า RMSE อยู่ที่ 2.8236 และให้ค่า MAPE อยู่ที่ 0.0196 ซึ่งแสดงให้เห็นว่าการใช้เทคนิคการรวมโมเดลสำหรับการทำนายล่วงหน้าแบบ 5 วัน ใช้เครื่องมือแบบ Voting ได้ผลดีกว่าการใช้เครื่องมือแบบ Stacking

บทที่ 5
สรุปผลการวิจัย

5.1. สรุปผลการวิจัย

งานวิจัยนี้มุ่งเน้นการศึกษาเพื่อการทำนายราคากองทุนทองคำ และหาค่าความแม่นยำระหว่างราคากองทุนทองคำจากการทำนายจากการเรียนรู้ของเครื่อง และราคาจริง เพื่อช่วยให้นักลงทุน และผู้ที่เกี่ยวข้องเข้าใจแนวโน้มของราคากองทุนทองคำในอนาคตได้ดีขึ้น ช่วยในการตัดสินใจลงทุน และช่วยในการตัดสอนใจเพื่อปรับกลยุทธ์การลงทุนได้ดีขึ้น ซึ่งงานวิจัยนี้จะศึกษาวิเคราะห์ เปรียบเทียบแบบจำลองการเรียนรู้ของเครื่อง ประกอบไปด้วย ซัพพอร์ตเวกเตอร์รีเกสชัน, การถดถอยแบบต้นไม้ตัดสินใจ, การถดถอยของป่าแบบสุ่ม, การถดถอยลาสโซ, การถดถอยแบบริดจ์, การถดถอยของริดจ์แบบเบส์, ถดถอยแบบกราเดียนท์บูสติง และสโตแคสติกแกรเดียนท์เดสเซนท์

จากผลการศึกษาพบว่าโมเดลการพยากรณ์ล่วงหน้า 1 วัน โมเดลที่ใช้การคัดเลือกคุณลักษณะโดยใช้ Random Forest Selection มีประสิทธิภาพสูงสุด โดยการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด ได้แก่ Baysian Ridge Regression, Stochastic Gradient Decent, Lasso Regression และ Lasso Regression โดยมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) เป็น 0.0036, 0.0094, 0.0103 และ 0.0109 ตามลำดับ ซึ่งเป็นการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด 4 อันดับแรก โดยเมื่อนำโมเดลมาที่มีประสิทธิภาพสูงสุด 4 อันดับแรกเข้าสู่กระบวนการใช้เทคนิคการรวมโมเดล พบว่า การใช้โมเดลที่รวม 3 โมเดลโดยใช้เทคนิคแบบ Stacking ให้ MAPE ที่ต่ำที่สุด คือ 0.0035

การพยากรณ์ล่วงหน้า 3 วัน โมเดลที่ใช้การคัดเลือกคุณลักษณะโดยใช้ Random Forest Selection มีประสิทธิภาพสูงสุด โดยการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด ได้แก่ Baysian Ridge Regression, Stochastic Gradient Decent, Lasso Regression และ Lasso Regression โดยมีค่า MAPE เป็น 0.0115, 0.0138, 0.0141 และ 0.0142 ตามลำดับ ซึ่งเป็นการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด 4 อันดับแรก โดยเมื่อนำโมเดลมาที่มีประสิทธิภาพสูงสุด 4 อันดับแรกเข้าสู่กระบวนการใช้เทคนิคการรวมโมเดล พบว่า การใช้โมเดลที่รวม 3 และ 4 โมเดลโดยใช้เทคนิคแบบ Stacking ให้ MAPE ที่ต่ำที่สุด คือ 0.0111

การพยากรณ์ล่วงหน้า 5 วัน โมเดลที่ใช้การคัดเลือกคุณลักษณะโดยใช้ Lasso Selection มีประสิทธิภาพสูงสุด โดยการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด ได้แก่ Ridge Regression, Lasso Regression, Support Vector Regression และ Stochastic Gradient

Decent โดยมีค่า MAPE เป็น 0.0166, 0.0167, 0.0191 และ 0.200 ตามลำดับ ซึ่งเป็นการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด 4 อันดับแรก โดยเมื่อนำโมเดลมาที่มีประสิทธิภาพสูงสุด 4 อันดับแรกเข้าสู่กระบวนการใช้เทคนิคการรวมโมเดล พบว่า การใช้โมเดลที่รวม 3 โมเดล โดยใช้เทคนิคแบบ Voting ให้ MAPE ที่ต่ำที่สุด คือ 0.0161

การพยากรณ์ล่วงหน้า 7 วัน โมเดลที่ใช้การคัดเลือกคุณลักษณะโดยใช้ Random Forest Selection มีประสิทธิภาพสูงสุด โดยการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด ได้แก่ Ridge Regression, Lasso Regression, Stochastic Gradient Decent และ Baysian Ridge Regression โดยมีค่า MAPE เป็น 0.0201, 0.0204, 0.0216 และ 0.0227 ตามลำดับ ซึ่งเป็นการเรียนรู้ของเครื่องที่มีประสิทธิภาพสูงสุด 4 อันดับแรก เมื่อเข้าสู่กระบวนการใช้เทคนิคการรวมโมเดล พบว่า การใช้โมเดลที่รวม 4 โมเดล โดยใช้เทคนิคแบบ Voting ให้ RMSE ที่ต่ำที่สุด คือ 0.0196

สรุปได้ว่า การศึกษานี้สามารถสร้างแบบจำลองโดยประยุกต์ใช้เทคนิคการรวมโมเดลเพื่อพยากรณ์ราคาทองคำแบบล่วงหน้า 1, 3, 5 และ 7 วันสามารถทำนายราคาทองได้ค่อนข้างแม่นยำ ซึ่งการทำนายแบบล่วงหน้า 1 วันจะให้ความแม่นยำที่มาที่สุด และลดหลั่นกันไปตามรายะห่างของจำนวนวัน ซึ่งการพยากรณ์นี้สามารถนำไปประยุกต์ใช้ได้จริงในการพยากรณ์ราคากองทุนทองคำ เพื่อใช้ในคาดคะเนการลงทุน และความเสี่ยงในการลงทุนในกองทุนทองคำได้

5.2. อภิปรายผล

จากการศึกษาการทำนายราคากองทุนทองคำโดยใช้แบบจำลองการเรียนรู้ของเครื่อง พบว่าแบบจำลองมีประสิทธิภาพสูงในการพยากรณ์ระยะสั้น โดยเฉพาะการพยากรณ์ล่วงหน้า 1 วัน ซึ่งให้ผลลัพธ์ที่แม่นยำที่สุดเมื่อใช้การคัดเลือกคุณลักษณะด้วย Random Forest Selection ร่วมกับการรวมโมเดลแบบ Stacking ซึ่งสามารถลดค่า MAPE ลงได้ถึง 0.0035 อย่างมีนัยสำคัญ ในขณะที่การพยากรณ์ในระยะ 3, 5 และ 7 วัน มีแนวโน้มที่ความแม่นยำจะลดลง โดยแบบจำลองที่ใช้ในการพยากรณ์ล่วงหน้า 3 วันยังให้ความแม่นยำในระดับที่ดี ส่วนการพยากรณ์ระยะยาว เช่น 5 และ 7 วัน มีค่าความคลาดเคลื่อนเพิ่มขึ้นตามลำดับ ซึ่งสอดคล้องกับลักษณะของตลาดทองคำที่มีความผันผวนสูงและอ่อนไหวต่อปัจจัยภายนอก การเลือกใช้เทคนิคการคัดเลือกคุณลักษณะมีผลต่อประสิทธิภาพของแบบจำลองอย่างชัดเจน โดย Random Forest Selection เหมาะสมในกรณีส่วนใหญ่ ยกเว้นการพยากรณ์ล่วงหน้า 5 วันที่ Lasso Selection ให้ผลดีกว่า ขณะเดียวกันเทคนิคการรวมโมเดลทั้งแบบ Stacking และ Voting มีส่วนช่วยเพิ่มความแม่นยำของแบบจำลอง โดยเฉพาะ Stacking ที่เหมาะกับระยะสั้น และ Voting ที่เหมาะกับการพยากรณ์ในระยะยาว

อย่างไรก็ตาม แบบจำลองยังมีข้อจำกัดในด้านการพยากรณ์ระยะยาว ซึ่งอาจได้รับผลกระทบจากปัจจัยภายนอกที่ไม่รวมอยู่ในชุดข้อมูล ดังนั้นในอนาคตควรผนวกข้อมูลเศรษฐกิจ ข่าวสาร หรือเหตุการณ์สำคัญระดับมหภาคเข้ามา เพื่อเสริมความสามารถของแบบจำลองในการพยากรณ์อย่างครอบคลุมและแม่นยำยิ่งขึ้น

5.3. ข้อจำกัดและอุปสรรคของงานวิจัย

ข้อจำกัดของงานวิจัยฉบับนี้อยู่ในหลายประเด็นที่ควรพิจารณาในการนำผลลัพธ์ไปใช้ในเชิงปฏิบัติ ประการแรก งานวิจัยใช้ข้อมูลราคากองทุนทองคำในช่วงระยะเวลาที่จำกัด และแหล่งข้อมูลเฉพาะ ซึ่งอาจไม่ครอบคลุมปัจจัยภายนอกที่มีผลกระทบต่อราคาทองคำในโลกแห่งความเป็นจริง เช่น ปัจจัยเศรษฐกิจมหภาค เหตุการณ์ทางการเมือง หรือความผันผวนในตลาดโลก นอกจากนี้ แบบจำลองที่ใช้ยังเน้นการเรียนรู้จากข้อมูลเชิงตัวเลขย้อนหลัง โดยไม่ได้พิจารณาตัวแปรเพิ่มเติม เช่น ดอกเบี้ยนโยบาย หรือข่าวสารที่อาจสะท้อนแนวโน้มราคาทองคำได้อย่างมีประสิทธิภาพมากยิ่งขึ้น อีกทั้งการพัฒนาแบบจำลองยังมีข้อจำกัดด้านการเลือกและปรับพารามิเตอร์ ซึ่งมีผลต่อความแม่นยำในการพยากรณ์ โดยเฉพาะอย่างยิ่งเมื่อช่วงเวลาการพยากรณ์เพิ่มขึ้นจาก 1 วัน เป็น 3, 5 และ 7 วัน จะพบว่าค่าความคลาดเคลื่อนเพิ่มสูงขึ้นอย่างชัดเจน แสดงให้เห็นถึงข้อจำกัดของแบบจำลองในการทำนายระยะยาว นอกจากนี้ โมเดลที่พัฒนาขึ้นยังได้รับการทดสอบเฉพาะในบริบทของกองทุนทองคำเท่านั้น จึงไม่สามารถสรุปผลไปยังสินทรัพย์ประเภทอื่นได้โดยตรง โดยไม่ผ่านกระบวนการปรับแบบจำลองให้เหมาะสมกับบริบทใหม่ ด้วยข้อจำกัดดังกล่าว ผู้วิจัยแนะนำให้มีการศึกษาเพิ่มเติมที่ผสานข้อมูลจากหลากหลายแหล่ง และนำตัวแปรภายนอกมาร่วมในการสร้างแบบจำลอง เพื่อเพิ่มความแม่นยำและความสามารถในการใช้งานจริงของระบบพยากรณ์ราคากองทุนในอนาคต

5.4. ข้อเสนอแนะ

5.4.1. งานวิจัยในอนาคตควรพิจารณาใช้ข้อมูลราคากองทุนทองคำในช่วงเวลาที่ยาวนานขึ้น เพื่อเพิ่มความหลากหลายของข้อมูลและรองรับความผันผวนในหลายสภาวะเศรษฐกิจ ซึ่งจะช่วยให้แบบจำลองสามารถเรียนรู้รูปแบบของข้อมูลได้ดียิ่งขึ้น และสามารถปรับตัวให้เข้ากับสถานการณ์ที่แตกต่างกันในอนาคตได้อย่างมีประสิทธิภาพ การใช้ข้อมูลในช่วงเวลาที่จำกัดอาจทำให้แบบจำลองเกิดความเอนเอียงและไม่สามารถคาดการณ์ได้แม่นยำเมื่อเกิดสถานการณ์ที่ไม่เคยเกิดขึ้นมาก่อนในชุดข้อมูลฝึก

5.4.2. นอกจากข้อมูลราคาย้อนหลังของกองทุนทองคำแล้ว ควรมีการบูรณาการข้อมูลเชิงเศรษฐกิจอื่นๆ ที่มีความสัมพันธ์กับราคาทองคำ เช่น อัตราแลกเปลี่ยน ดอกเบี้ยเงินฝาก อัตราเงินเฟ้อ ราคาน้ำมัน หรือข้อมูลตลาดหุ้นระหว่างประเทศ รวมไปถึงข้อมูลจากข่าวสารและเหตุการณ์สำคัญระดับโลก การเพิ่มตัวแปรเหล่านี้จะช่วยให้แบบจำลองสามารถพิจารณาความสัมพันธ์เชิงสาเหตุที่ลึกซึ้งยิ่งขึ้น และสะท้อนสภาพแวดล้อมจริงได้ดีขึ้น ซึ่งจะนำไปสู่ความแม่นยำในการพยากรณ์ที่สูงขึ้น

5.4.3. การเลือกคุณลักษณะเป็นขั้นตอนสำคัญที่ส่งผลต่อประสิทธิภาพของโมเดลอย่างมาก งานวิจัยในอนาคตควรเปรียบเทียบวิธีการคัดเลือกคุณลักษณะที่หลากหลาย เช่น Recursive Feature Elimination (RFE), Principal Component Analysis (PCA) หรือ Mutual Information กับวิธีที่ใช้อยู่ในปัจจุบันอย่าง Random Forest และ Lasso ซึ่งอาจช่วยให้สามารถลดจำนวนตัวแปรที่ไม่จำเป็น และเพิ่มความเร็วในการประมวลผล โดยไม่ลดทอนความแม่นยำของแบบจำลอง

5.4.4. แม้ว่าการใช้เทคนิคการรวมโมเดล เช่น Stacking และ Voting จะให้ผลลัพธ์ที่ดี แต่ยังมีโอกาสในการพัฒนาเพิ่มเติมโดยการผสมผสานเทคนิคต่างๆ เข้าด้วยกันอย่างเป็นระบบ เช่น การนำ Voting มารวมกับ Boosting หรือใช้ Meta-Learner ที่ซับซ้อนขึ้น เพื่อหาวิธีการรวมโมเดลที่ให้ความแม่นยำและเสถียรภาพสูงที่สุดในแต่ละช่วงการพยากรณ์

บรรณานุกรม

Akhtaruzzaman, M., Boubaker, S., Lucey, B. M., & Sensoy, A. (2021). Is gold a hedge or a safe-haven asset in the COVID–19 crisis? *Economic Modelling*, 102, 105588.

Alkhazali, O. M., & Zoubi, T. A. (2020). Gold and portfolio diversification: A stochastic dominance analysis of the Dow Jones Islamic indices. *Pacific-Basin Finance Journal*, 60, 101264.

Appel, G. (2005). Advanced Moving Average Convergence-Divergence (MACD): The Ultimate Market Timing Indicator. *Technical Analysis (Power Tools for Active Investors); Prentice Hall: Hoboken, NJ, USA*, 165-200.

Cardoso, J. P. V. L. (2019). *Technical analysis on foreign exchange markets: MACD and RSI*

Ciner, C. (2001). On the long run relationship between gold and silver prices A note. *Global finance journal*, 12(2), 299-303.

Eryigit, M. (2017). Short-term and long-term relationships between gold prices and precious metal (palladium, silver and platinum) and energy (crude oil and gasoline) prices. *Economic research-Ekonomska istraživanja*, 30(1), 499-510.

Ghule, R., Gadhave, A., Dubey, M., Kharade, J., & Mumbai, N. (2022). Gold price prediction using machine learning. *International Journal of Environmental Engineering*, 6(2022).

Ghute, M., & Korde, M. (2023). Efficient Machine Learning Algorithm for Future Gold Price Prediction. 2023 International Conference on Inventive Computation Technologies (ICICT),

Gilmore, C. G., McManus, G. M., Sharma, R., & Tezel, A. (2009). The dynamics of gold prices, gold mining stock prices and stock market prices comovements. *Research in Applied Economics*, 1(1), 1-19.

Houcine, B., Zouheyr, G., Abdessalam, B., Youcef, H., & Hanane, A. (2020). The relationship between crude oil prices, EUR/USD exchange rate and gold prices. *International Journal of Energy Economics and Policy*, 10(5), 234-242.

Ismail, Z., Yahya, A., & Shabri, A. (2009). Forecasting gold prices using multiple linear regression method. *American Journal of Applied Sciences*, 6(8), 1509.

Kearney, A. A., & Lombra, R. E. (2009). Gold and platinum: Toward solving the price puzzle. *The Quarterly Review of Economics and Finance*, 49(3), 884-892.

Kuck, K. (2021). Gold and the S&P500: An analysis of the return and volatility relationship. *Available at SSRN* 3930548.

Kumar, B. S., & Raj, G. R. (2019). Gold vs gold exchange traded funds: An empirical study in India. *Economic Affairs*, 64(4), 703-710.

Le, T.-H., & Chang, Y. (2011). Oil and gold: correlation or causation?

Mamcarz, K. (2024). The Causal Relationship Between Stocks, Gold, Crude Oil, and Bond Returns in Poland. *Annales Universitatis Mariae Curie-Skłodowska, Sectio H Oeconomia*, 58(4), 127-147.

Manjula, K., & Karthikeyan, P. (2019). Gold price prediction using ensemble based machine learning techniques. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI),

Pati, A., Mishra, S., Sahoo, S., Kar, A., & Student, M. (2024). Forecasting Gold Prices: Leveraging Machine Learning Algorithms. *INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS*, 12, 2320-2882.

Shahani, R., & Bansal, A. (2021). Is gold an effective hedge and/or safe haven instrument against stocks, rupee–dollar rate and crude: an empirical investigation from India. *Decision*, 48(2), 129-151.

มีชัยเจริญยิ่ง, ช. (2024, 02). เจาะลึกราคาทองคำ หลัง FED ลดดอกเบี้ย.

ประวัติผู้เขียน

ชื่อ-สกุล ธนกฤต พลอยเกตุ

วัน เดือน ปี เกิด 12 กันยายน 2537