

การเปรียบเทียบประสิทธิภาพของค่าพยากรณ์ราคาทองคำ จากวิธีเฉลี่ยตัวแบบบนต้นไม้การถดถอยกับป่าสุ่ม

ศุภาพิชญ์ สังขกฤษณ์¹ และ สุปราณี ลิขสวัสดิ์²

สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ปทุมธานี 12120

Received: 03 April 2025; Revised: 28 May 2025; Accepted: 04 June 2025

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการพยากรณ์ราคาทองคำ โดยใช้วิธีการเฉลี่ยตัวแบบบนต้นไม้การถดถอย 2 วิธี ได้แก่ Multi-Model Tree Averaging และ K-Fold Tree Averaging ซึ่งมีการถ่วงน้ำหนักค่าพยากรณ์แบบน้ำหนักที่เท่ากันและน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะ รวมถึงเปรียบเทียบกับวิธีป่าสุ่ม เพื่อประเมินความแม่นยำและความเสถียรของค่าพยากรณ์ ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลทุติยภูมิรายเดือนเกี่ยวกับตัวแปรทางเศรษฐกิจ เช่น อัตราเงินเฟ้อและอัตราดอกเบี้ยของประเทศไทย สหรัฐอเมริกา จีน และอินเดีย ระหว่างเดือนมกราคม พ.ศ. 2557 ถึงกรกฎาคม พ.ศ. 2567 ผลการวิจัยพบว่า วิธีป่าสุ่มให้ผลลัพธ์ที่เหมาะสมที่สุดในแง่ของค่า R^2 , RMSE และ MAPE ทั้งในชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบ โดยแสดงถึงความแม่นยำและความเสถียรสูงสุด ขณะที่ Multi-Model Tree Averaging และ K-Fold Tree Averaging มีประสิทธิภาพรองลงมา โดยการถ่วงน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะให้ผลลัพธ์ที่ดีกว่าแบบน้ำหนักที่เท่ากันในทุกกรณี งานวิจัยนี้แสดงให้เห็นถึงศักยภาพของวิธีการพยากรณ์ที่หลากหลายและเสนอแนะแนวทางการเลือกวิธีการที่เหมาะสมสำหรับการพยากรณ์ราคาทองคำและสินทรัพย์ที่มีความผันผวนสูง

คำสำคัญ: การพยากรณ์ราคาทองคำ, multi-model tree averaging, K-fold tree averaging, ป่าสุ่ม, การเฉลี่ยตัวแบบ

*Corresponding author: E-mail: supapitch.san@dome.tu.ac.th

¹นักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์ มหาวิทยาลัยธรรมศาสตร์

²รองศาสตราจารย์ ดร. คณะวิทยาศาสตร์ มหาวิทยาลัยธรรมศาสตร์

A comparative analysis of gold price forecasting efficiency using model averaging over regression trees and random forest

Supapitch Sangakulchorn^{*1} and Supranee Lisawadi²

^{1,2}Department of Mathematics and Statistics, Faculty of Science and Technology,
Thammasat University, Pathum Thani 12120

Received: 03 April 2025; Revised: 28 May 2025; Accepted: 04 June 2025

Abstract

This research aims to compare the forecasting performance of gold prices using two tree-averaging methods: Multi-Model Tree Averaging and K-Fold Tree Averaging. Both methods employ weighting schemes, namely Equal Weight and AIC Weight. The study also compares these methods with Random Forest to evaluate forecasting accuracy and stability. The data used are monthly secondary data on economic variables, such as inflation rates and interest rates in Thailand, the United States, China, and India, spanning from January 2014 to July 2024. The results show that Random Forest achieves the most suitable performance in terms of R^2 , RMSE, and MAPE in both training and testing datasets, demonstrating superior accuracy and stability. Multi-Model Tree Averaging and K-Fold Tree Averaging rank slightly lower in performance, with AIC Weight consistently outperforming Equal Weight in all scenarios. This study highlights the potential of diverse forecasting methods and provides recommendations for selecting appropriate techniques for predicting gold prices and other highly volatile assets.

Keywords: gold price forecasting, multi-model tree averaging, K-fold tree averaging, random forest, model averaging

*Corresponding author: E-mail: supapitch.san@dome.tu.ac.th

¹ Student in Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University

² Associate Professor in Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University

1. บทนำ

ในช่วงหลายปีที่ผ่านมา ทองคำยังคงรักษาสถานะของตนในฐานะสินทรัพย์ที่ปลอดภัย โดยเฉพาะในช่วงเวลาที่เศรษฐกิจโลกเผชิญกับความไม่แน่นอน อัตราเงินเฟ้อสูง หรือความผันผวนในตลาดทุน โดยจากข้อมูลของ World Gold Council ณ เดือนกรกฎาคม ปี 2567 ราคาทองคำโลกอยู่ที่ประมาณ 1,960–2,000 ดอลลาร์สหรัฐต่อออนซ์ ซึ่งนับว่าเป็นจุดสูงสุดใหม่เป็นประวัติการณ์ สาเหตุสำคัญมาจากความกังวลเกี่ยวกับเศรษฐกิจจีน ค่าเงินดอลลาร์สหรัฐที่อ่อนค่า และแนวโน้มการปรับลดอัตราดอกเบี้ยของธนาคารกลางสหรัฐ (FED) อีกทั้งในประเทศไทย ขณะที่ราคาทองคำแท่งในประเทศไทยเคลื่อนไหวอยู่ในช่วง 31,000–32,000 บาทต่อบาททองคำ ซึ่งยังคงเป็นระดับราคาที่สูงกว่าค่าเฉลี่ยในระยะยาว

การพยากรณ์ราคาทองคำมีความสำคัญต่อการตัดสินใจของ นักลงทุน นักเก็งกำไร และผู้บริโภคทั่วไป เนื่องจากทองคำเป็นสินทรัพย์ ที่มีมูลค่าสูงและราคามีความผันผวน การพยากรณ์ดังกล่าวช่วยเสริมประสิทธิภาพในการวางแผนการเงิน การบริหารจัดการ และการกำหนด นโยบายที่เกี่ยวข้องกับการใช้ทองคำในกระบวนการผลิต โดยในปี 1959 Arthur Samuel [1] มีการนำเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งเป็นส่วนหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence: AI) มาใช้ในการพยากรณ์ราคาทองคำ เทคนิคนี้มุ่งเน้น การค้นหารูปแบบจากข้อมูลโดยไม่ต้องพึ่งพาการป้อนคำสั่งที่ชัดเจน

หนึ่งในวิธีการที่ได้รับความนิยมคือต้นไม้การถดถอย ซึ่งช่วยแบ่งข้อมูลออกเป็นกลุ่มย่อยเพื่อเพิ่มความแม่นยำในการพยากรณ์ อย่างไรก็ตาม วิธีนี้อาจประสบปัญหาการเข้ากันเกินไป (Overfitting) วิธีการหนึ่งที่ช่วยแก้ปัญหานี้คือการเฉลี่ยตัวแบบ (Model Averaging) โดยเป็นการรวบรวมตัวแบบต่าง ๆ เข้าด้วยกัน และให้น้ำหนักกับค่าทำนายแต่ละตัวแบบ ซึ่งช่วยเพิ่มความแม่นยำของค่าทำนายรวม

ในการศึกษาเรื่องนี้ มีการใช้เทคนิคการเฉลี่ยตัวแบบผ่านสองแนวทาง ได้แก่ Multi-Model Tree Averaging (MMTA) และ K-Fold Tree Averaging

(KFTA) เทคนิค MMTA ใช้การรวมค่าพยากรณ์จากต้นไม้การถดถอยหลายตัวแบบเข้าด้วยกัน โดยการกำหนดน้ำหนักเพื่อเพิ่มความแม่นยำ ในขณะที่ KFTA ใช้กระบวนการ K-Fold Cross-Validation เพื่อประเมินและเฉลี่ยค่าพยากรณ์จากตัวแบบในความลึกต่าง ๆ ซึ่งช่วยลดความลำเอียงและข้อผิดพลาดในการพยากรณ์

นอกจากนี้วิธีป่าสุ่มเป็นอีกหนึ่งเทคนิคที่ถูกพัฒนาขึ้นเพื่อแก้ไขข้อจำกัดของต้นไม้การถดถอย โดยใช้การสุ่มตัวอย่างข้อมูล (Bagging) และการสร้างต้นไม้การถดถอยหลายต้น เทคนิคนี้เหมาะกับข้อมูลที่มีความซับซ้อนสูง เช่น ข้อมูลที่มีตัวแปรจำนวนมากหรือมีความไม่เป็นเชิงเส้น ทำให้วิธีป่าสุ่มได้รับความนิยมนำไปแพร่หลายในหลากหลายสาขา เช่น การพยากรณ์ราคาสินทรัพย์ และการจำแนกประเภทข้อมูล

งานวิจัยนี้จึงมุ่งเน้นการพัฒนาและประยุกต์ใช้เทคนิคดังกล่าวในการพยากรณ์ราคาทองคำ เพื่อลดปัญหาการเข้ากันเกินไป และเพิ่มความแม่นยำในการพยากรณ์ ทั้งนี้ยังช่วยสนับสนุนการวางแผนการลงทุนและการตัดสินใจที่เกี่ยวข้องกับทองคำอย่างมีประสิทธิภาพ

2. การทบทวนวรรณกรรม

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ต้นไม้การถดถอย (Regression Tree)

ต้นไม้การถดถอยคือหนึ่งในรูปแบบของต้นไม้ตัดสินใจ (Decision Tree) ที่ใช้ในการทำงานเชิงการถดถอย โดยใช้ Residual Sum of Squares (RSS) เป็นฟังก์ชันวัตถุประสงค์ (Objective Function) เพื่อค้นหาจุดแบ่ง (Split Point) ที่เหมาะสมที่สุด โดยการลดค่า RSS ให้น้อยที่สุด

2.1.2 การเฉลี่ยตัวแบบ (Model Averaging)

แนวคิดของการเฉลี่ยตัวแบบถูกเสนอครั้งแรกโดย Francis Galton ในปี 1886 [2] เพื่อแก้ปัญหากลุ่มตัวอย่างที่ดีที่สุดเพียงตัวเดียว โดยการเฉลี่ยผลลัพธ์จากหลายตัวแบบที่มีประสิทธิภาพใกล้เคียงกัน ช่วยเพิ่มความแม่นยำและความแข็งแกร่งของผลลัพธ์ (Robustness) โดยมีสมการทั่วไปดังนี้

$$\hat{y}^* = \sum_{i=1}^M w_i \hat{y}_i \quad (1)$$

โดยที่ $\hat{y}^*$ แทน ค่าพยากรณ์รวม
 w_i แทน ค่าน้ำหนักที่ให้แต่ละตัวแบบ
 (Model Weight) เมื่อ $i = 1, 2, \dots, M$
 $\hat{y}_i$ แทน ค่าพยากรณ์ที่มาจากแต่ละตัวแบบ
 M แทน จำนวนตัวแบบทั้งหมดที่นำมา
 เฉลี่ย

โดยมีเงื่อนไข 2 ประการ คือ $\sum_{i=1}^M w_i = 1$ และ $0 \leq w_i \leq 1$

2.1.2.1 น้ำหนักที่เท่ากัน (Equal Weight)

การเฉลี่ยตัวแบบโดยกำหนดให้ค่า
 น้ำหนักของแต่ละตัวแบบมีค่าที่เท่ากัน การกำหนด
 น้ำหนักในลักษณะนี้ แม้ว่าเป็นการนำเสนอในรูปแบบที่
 ง่ายและไม่ซับซ้อน แต่สามารถให้ผลลัพธ์ที่ดีเมื่อนำมาใช้
 กับการเฉลี่ยตัวแบบภายใต้การทดสอบแบบไขว้ (Cross
 Validation) ตามที่ระบุไว้ในงานของ Marmion et al.
 (2009) [3] และ Rapacciuolo et al. (2012) [4] โดยมี
 รูปแบบทั่วไป ดังนี้

$$w_{eq} = \frac{1}{M} \quad (2)$$

โดยที่ w_{eq} แทน ค่าน้ำหนักที่ให้แต่ละตัวแบบ
 เมื่อใช้
 ค่าน้ำหนักที่เท่ากัน
 M แทน จำนวนตัวแบบทั้งหมดที่นำมา
 เฉลี่ย

2.1.2.2 น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศ ของอะกะอิเกะ (AIC Weight)

Kenneth P. Burnham and
 David R. Anderson (2022) [5] นำเสนอการเฉลี่ยตัว
 แบบโดยพิจารณาจากเกณฑ์สารสนเทศของอะกะอิเกะ
 เป็นตัวเลือกพื้นฐานสำหรับวิธีการเฉลี่ยตัวแบบโดยมี
 รูปแบบทั่วไป ดังนี้

$$w_{AIC_i} = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_{a=1}^M \exp\left(-\frac{1}{2}\Delta_a\right)} \quad (3)$$

$$\text{เมื่อ } \Delta_i = AIC_i - AIC_{min} \geq 0 \quad (4)$$

$$AIC_i = \log\left(\frac{SSE}{n}\right) + \frac{2T}{n} \quad (5)$$

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

โดยที่ AIC_{min} แทน ค่า AIC ของตัวแบบที่ให้ค่า AIC
 ต่ำที่สุด

AIC_i แทน ค่า AIC ของแต่ละตัวแบบ

T แทนจำนวนพารามิเตอร์ซึ่งมีค่าเท่ากับ
 จำนวนโนดใบ

M แทน จำนวนตัวแบบทั้งหมดที่นำมาเฉลี่ย

n แทน ขนาดตัวอย่าง

y_i แทน ค่าสังเกตของตัวแปรตามที่ i

$\hat{y}_i$ แทน ค่าพยากรณ์ที่ i

2.1.3 ป่าสุ่ม (Random Forest)

วิธีป่าสุ่มถูกนำเสนอครั้งแรกในปี 1995 โดย Tin K.
 Ho [6] และพัฒนาต่อในปี 2001 โดย Leo Breiman [7] เป็น
 เทคนิคการเรียนรู้ของเครื่องที่ต่อยอดจากต้นไม้ตัดสินใจและ
 ต้นไม้การถดถอย มีความสามารถในการจัดการข้อมูลที่ไม่เป็น
 เชิงเส้น ตัวแปรอิสระจำนวนมากและข้อมูลที่มีความซับซ้อน
 โดยไม่จำเป็นต้องมีการจัดการข้อมูลล่วงหน้า (Data
 Preprocessing)

กระบวนการทำงานของวิธีป่าสุ่มเริ่มจากการสร้าง
 ต้นไม้การถดถอยหลายต้นจากตัวอย่างข้อมูลสุ่ม (Bootstrap
 Sample) พร้อมกับการสุ่มตัวแปรอิสระบางส่วนในแต่ละ
 ต้นไม้ เพื่อลดปัญหาความสัมพันธ์กันของข้อมูล เทคนิคนี้ช่วย
 เพิ่มความแม่นยำของการพยากรณ์และลดปัญหาการเข้ากัน
 เกินไป ผลลัพธ์จากต้นไม้แต่ละต้นจะถูกเฉลี่ยร่วมกันเพื่อสร้าง
 การพยากรณ์ที่เสถียรและเชื่อถือได้ นอกจากนี้ ข้อมูลที่ไม่ได้
 ถูกสุ่มมา (Out-of-Bag Dataset) ยังใช้สำหรับประเมินความ
 แม่นยำของตัวแบบได้อีกด้วย เทคนิควิธีป่าสุ่มจึงเหมาะ
 สำหรับการประยุกต์ใช้ในงานวิจัยที่ต้องการจัดการข้อมูลที่
 ซับซ้อนและมีตัวแปรอิสระจำนวนมาก

2.2 งานวิจัยที่เกี่ยวข้อง

ในปี 2550 ปรกรณ์ จารุตระกูลชัย [8] ได้ศึกษาเรื่อง
 การทำนายปริมาณสารส้มที่ใช้ในกระบวนการผลิตน้ำประปา
 ด้วยวิธีป่าของต้นไม้ตัดสินใจและโปรแกรมเชิงพันธุกรรม โดย
 การศึกษานี้มีวัตถุประสงค์เพื่อหาวิธีที่เหมาะสมในการทำนาย

ปริมาณสารส้มในกระบวนการผลิตน้ำประปา โดยใช้ซอฟต์แวร์เวกเตอร์แมชชีน, ต้นไม้ตัดสินใจ, ป่าของต้นไม้ตัดสินใจ และเปรียบเทียบกับวิธีการแบบนิรอลเน็ตเวิร์ก ผลการศึกษาพบว่า วิธีป่าของต้นไม้ตัดสินใจสามารถทำนายได้แม่นยำที่สุดเมื่อใช้ตัวแปรอินพุตที่ปรับจากโปรแกรมเชิงพันธุกรรม

ต่อมาในปี 2021 Arisa Sirichartchai [9] ได้ศึกษาเกี่ยวกับการพัฒนาตัวแบบเพื่อทำนายระยะเวลาการพักอยู่ของผู้ป่วยในโรงพยาบาล (Length of Stay: LOS) โดยเปรียบเทียบตัวแบบต้นไม้การถดถอย, โคโรข่ายประสาทเทียม, วิธีป่าสุ่ม และ Voting Ensemble ผลการศึกษาพบว่าตัวแบบวิธีป่าสุ่มมีประสิทธิภาพดีที่สุดในการทำพยากรณ์ LOS ด้วยเวลาประมวลผลที่สั้น และความแม่นยำที่ดี

จากนั้นในปี 2564 ชนินทร์ แก้ววิบูลย์พันธุ์ [10] ได้ทำการศึกษาเพื่อเพิ่มประสิทธิภาพของต้นไม้การถดถอย โดยการสร้างต้นไม้การถดถอยหลายต้นย่อยและทำการเฉลี่ยตัวแบบผ่านวิธีการต่างๆ รวมถึงการใช้เกณฑ์สารสนเทศของอะกะอิเกะ การศึกษาพบว่า วิธีการเฉลี่ยตัวแบบบนต้นไม้การถดถอยสามารถเพิ่มความแม่นยำให้กับค่าพยากรณ์ได้

ในปี 1995 Ron Kohavi [11] ได้ทำการเปรียบเทียบประสิทธิภาพของวิธีการประเมินตัวแบบสองวิธี 10-fold cross-validation และ bootstrap ผลการศึกษาแสดงว่า 10-fold cross-validation แบบ stratified ให้ผลลัพธ์ที่แม่นยำกว่าและช่วยลดปัญหา การเข้ากันเกินไป เป็นตัวอย่างในการทำ K-fold Cross Validation

ในปี 2564 ศิวกร บรรลือทรัพย์ [12] ได้ศึกษาเรื่องการจำแนกผู้ป่วยโควิด-19 ด้วยเทคนิควิธีป่าสุ่ม โดยใช้ข้อมูลจากผู้ป่วย 1,608,923 รายและทำการเปรียบเทียบประสิทธิภาพของการจำแนกผู้ป่วยระหว่างวิธีป่าสุ่ม, นิรอลเน็ตเวิร์ค และ Naive Bayes ผลการศึกษาแสดงว่า วิธีป่าสุ่มมีความถูกต้องสูงสุดที่ 93.33% และสามารถใช้ในการพัฒนาระบบจำแนกผู้ป่วยโควิด-19 ได้

ในปี 2564 สุภัตสร สมนเจตนา และ จาริ ทองคำ [13] ได้ศึกษาเพื่อเปรียบเทียบประสิทธิภาพของเทคนิค

ต่างๆ ในการสร้างแบบจำลองจำแนกความคิดเห็นของผู้ปกครองต่อการใช้สมาร์ทโฟนของบุตร โดยใช้เทคนิคต่างๆ เช่น C4.5, Naive Bayes และ วิธีป่าสุ่ม พบว่า วิธีป่าสุ่มให้ผลลัพธ์ที่ดีที่สุดในการจำแนกความคิดเห็น

ในปี 2023 Purwa Hasan Putra et al. [14] ได้ทำการศึกษาเกี่ยวกับผลกระทบของการนำพลังงานทดแทน เช่น รถยนต์ไฟฟ้า มาใช้ในการสนับสนุนของภาครัฐ ซึ่งส่งผลกระทบต่อความต้องการตลาดรถยนต์มือสอง การกำหนดราคาที่เหมาะสมสำหรับการซื้อขายรถยนต์มือสองจึงกลายเป็นปัญหาสำคัญสำหรับผู้บริโภคในยุคที่มีการพัฒนางานวิจัยดังกล่าวมุ่งเน้นการใช้เทคนิคแมชชีนเลิร์นนิงในการทำนายราคาของรถยนต์มือสอง โดยใช้วิธีป่าสุ่มและต้นไม้การถดถอย เพื่อวิเคราะห์ความแม่นยำในการทำนายราคา ผลการวิจัยพบว่าวิธีป่าสุ่มมีความแม่นยำสูงกว่าการใช้ต้นไม้การถดถอย โดยมีความแม่นยำเท่ากับ 72.13% สำหรับวิธีป่าสุ่ม และ 67.21% สำหรับต้นไม้การถดถอย ซึ่งบ่งชี้ว่าวิธีป่าสุ่มเป็นเครื่องมือที่มีประสิทธิภาพสูงกว่าในการคาดการณ์ราคาของรถยนต์มือสอง งานวิจัยนี้ได้สรุปว่าการใช้วิธีป่าสุ่มมีความสามารถในการคาดการณ์ที่ดีกว่าเมื่อเทียบกับต้นไม้การถดถอย ซึ่งจะใช้ 2 วิธีนี้ในการทำในงานวิจัยเล่มนี้

ในปี 2558 ณภัส อัญญ์ พิบูลพาณิชย์การ [15] การศึกษาค้นคว้ามุ่งศึกษาการวิเคราะห์ปัจจัยที่มีผลต่อราคาทองคำในประเทศไทย โดยการศึกษาใช้วิธีการวิเคราะห์สมการถดถอยพหุคูณ (Multiple Regression Analysis) เพื่อสำรวจความสัมพันธ์ระหว่างปัจจัยต่างๆ ที่ส่งผลกระทบต่อราคาทองคำแท่งในประเทศไทย ผลการศึกษาพบว่า ปัจจัยที่มีผลต่อราคาทองคำ ได้แก่ ราคาทองคำตลาดโลก Spot Gold, อัตราแลกเปลี่ยนสกุลเงินบาทต่อดอลลาร์สหรัฐฯ, ดัชนีราคาผู้บริโภคในประเทศไทย, อัตราดอกเบี้ย, และนโยบายธนาคารแห่งประเทศไทย

ในปี 1989 Thomas D. Kaufmann and Richard A. Winters [16] ศึกษาการพยากรณ์ราคาทองคำในปี 1989 โดยใช้การวิเคราะห์การถดถอย (Regression Analysis) เพื่อพยากรณ์ราคาทองคำจากการพิจารณาตัวแปรหลายตัว ได้แก่ การเปลี่ยนแปลงในอัตราเงินเฟ้อ ดัชนีการแลกเปลี่ยนเงินดอลลาร์สหรัฐฯ และการผลิตทองคำรายปี งานวิจัยพบว่าดัชนี

ของสหรัฐฯและการแลกเปลี่ยนเงินดอลลาร์สหรัฐฯ มีความสัมพันธ์ที่แข็งแกร่งกับราคาทองคำ

ข้อมูลในปี 2022 จาก World Gold Council [17] ระบุว่าความต้องการทองคำในอุตสาหกรรมเครื่องประดับ มีสัดส่วนถึง 44% ของความต้องการทองคำทั่วโลก โดยประเทศที่บริโภคทองคำในอุตสาหกรรมเครื่องประดับมากที่สุด ได้แก่ อินเดีย จีน และสหรัฐอเมริกา นอกจากนี้ทองคำยังมีความต้องการในด้านเทคโนโลยีและอุตสาหกรรม ซึ่งคิดเป็นอีก 7.5% ของความต้องการทั่วโลก ซึ่งให้เห็นว่าประเทศเหล่านี้มีความสัมพันธ์สูงกับราคาทองคำโลก

ในปี 2013 Samveg A. Patel [18] ศึกษาบทบาทของทองคำในการป้องกันความเสี่ยงจากเงินเฟ้อและอัตราแลกเปลี่ยนในอินเดีย โดยการใช้ข้อมูลรายเดือนตั้งแต่ปี 1991 ถึง 2012 ผลการวิจัยพบว่า ราคาทองคำมีความสัมพันธ์เชิงลบกับอัตราเงินเฟ้อและมีการใช้งานทองคำเป็นเครื่องป้องกันความเสี่ยงจากอัตราเงินเฟ้อในระยะยาว

ในปี 2018 Narinder Pal Singh และ Navneet Joshi [19] ศึกษาความสัมพันธ์ระยะยาวระหว่างราคาทองคำและดัชนีราคาผู้บริโภค (CPI) ในอินเดียระหว่างปี 2011-2017 โดยใช้วิธีการวิเคราะห์ของ Johansen's Cointegration พบว่า ทองคำและ CPI มีความสัมพันธ์ในระยะยาวที่สมดุล แต่ไม่พบความสัมพันธ์เชิงสาเหตุระหว่างทั้งสองตัวแปร

ในปี 2022 Kamaran Qader Yaqub [20] วิเคราะห์ผลกระทบของอัตราเงินเฟ้อ, อัตราดอกเบี้ย, อัตราการว่างงาน และผลิตภัณฑ์มวลรวมภายในประเทศ (GDP) ต่อราคาทองคำโดยใช้ตัวแบบ Generalized Method of Moments (GMM) พบว่าอัตราเงินเฟ้อส่งผลกระทบต่อราคาทองคำ ขณะที่อัตราดอกเบี้ยส่งผลกระทบต่อราคาทองคำ

3. ระเบียบวิธีวิจัย

3.1 ข้อมูลที่ใช้ในการศึกษา

ในการศึกษาครั้งนี้ผู้วิจัยมุ่งเป้าหมายในการเปรียบเทียบประสิทธิภาพของค่าพยากรณ์ราคาทองคำ โดยการรวบรวมข้อมูลทุติยภูมิ (Secondary Data) แบบอนุกรมเวลา (Time Series Data) รายเดือน ซึ่งเริ่มเก็บข้อมูลตั้งแต่เดือนมกราคม พ.ศ. 2557 ถึงเดือนกรกฎาคม พ.ศ. 2567 จากแหล่งข้อมูลออนไลน์ที่เชื่อถือได้ ได้แก่ เว็บไซต์ Trading Economics, Investing.com และ WORLD GOLD COUNCIL ข้อมูลที่รวบรวมมานั้นเกี่ยวข้องกับตัวแปรทางเศรษฐกิจต่างๆ ทั้งในประเทศไทยและต่างประเทศ โดยแบ่งเป็นข้อมูลหลักดังนี้

1. อัตราเงินเฟ้อในประเทศไทย (Thai Inflation Rate)
2. อัตราดอกเบี้ยในประเทศไทย (Thai Interest Rate)
3. อัตราเงินเฟ้อในประเทศอินเดีย (India Inflation Rate)
4. อัตราดอกเบี้ยในประเทศอินเดีย (India Interest Rate)
5. อัตราเงินเฟ้อในประเทศจีน (China Inflation Rate)
6. อัตราดอกเบี้ยในประเทศจีน (China Interest Rate)
7. ดัชนีราคาผู้บริโภคในประเทศสหรัฐอเมริกา (Consumer Price Index: CPI)
8. ดัชนีราคาจากรายจ่ายเพื่อการบริโภคส่วนบุคคลในประเทศสหรัฐอเมริกา (Personal Consumption Expenditure: PCE)
9. ดัชนีราคาผู้ผลิตในประเทศสหรัฐอเมริกา (Producer Price Index: PPI)
10. อัตราดอกเบี้ยในประเทศสหรัฐอเมริกา (United States Interest Rate)

ข้อมูลเหล่านี้จะถูกนำมาใช้ในการศึกษาความสัมพันธ์และการคาดการณ์ราคาทองคำ ซึ่งคาดหวังว่าจะมีบทบาทสำคัญในการวิเคราะห์และทำนายแนวโน้มของราคาทองคำในอนาคตโดยพิจารณาจากปัจจัยทางเศรษฐกิจต่างๆ ที่เกี่ยวข้อง

3.2 เครื่องมือที่ใช้ในการศึกษา

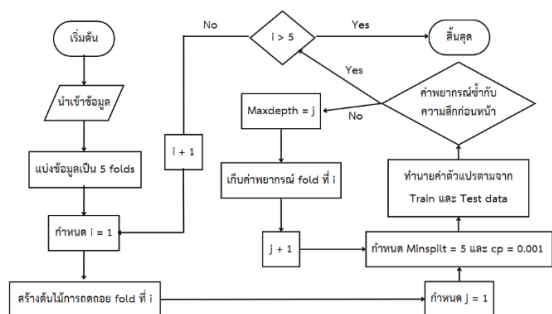
ในการศึกษาครั้งนี้ผู้วิจัยใช้เครื่องมือ RStudio ในการประมวลผลข้อมูลและวิเคราะห์ผลการพยากรณ์ราคาทองคำ ผู้วิจัยนำวิธีการเฉลี่ยตัวแบบบนต้นไม้การถดถอย (MMTA และ KFTA) และวิธีป่าสุ่มมาใช้ในการพยากรณ์ เพื่อเปรียบเทียบประสิทธิภาพของแต่ละวิธี

3.3 การทดสอบแบบไขว้ (Cross-Validation)

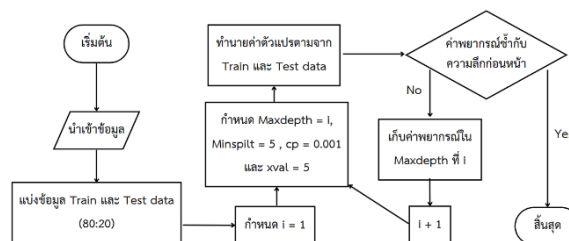
การทดสอบแบบไขว้เป็นที่ใช้ประเมินและเปรียบเทียบประสิทธิภาพของตัวแบบ โดยแบ่งชุดข้อมูลออกเป็น K ส่วนเท่า ๆ กัน ซึ่ง K ในการวิจัยนี้ถูกกำหนดเป็น 5 หมายถึง การแบ่งข้อมูลเป็น 5 ส่วน โดยแต่ละส่วนจะถูกใช้เป็นข้อมูลทดสอบ ในขณะที่ส่วนที่เหลือใช้เป็นข้อมูลฝึกสอน ตัวอย่างเช่น ในรอบที่ 1 ใช้ชุดข้อมูลที่ 1 เป็นชุดทดสอบ และชุดที่ 2-5 เป็นข้อมูลฝึกสอน และดำเนินการวนซ้ำจนครบ 5 รอบ ผลลัพธ์จากการฝึกสอนแต่ละรอบจะนำมาคำนวณค่าเฉลี่ยเพื่อให้ได้ค่าประสิทธิภาพโดยรวม

3.4 ขั้นตอนการดำเนินงาน

1. นำเข้าข้อมูลที่ต้องการศึกษา
2. สร้างค่าพยากรณ์ต้นไม้การถดถอยโดยวิธี MMTA
3. สร้างค่าพยากรณ์ต้นไม้การถดถอยโดยวิธี KFTA
4. สร้างค่าพยากรณ์โดยวิธีป่าสุ่ม
5. พิจารณาเกณฑ์ในการเปรียบเทียบประสิทธิภาพของค่าพยากรณ์จากทุกวิธีที่ทำการศึกษา โดยใช้สัมประสิทธิ์ การกำหนด, รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย, และค่าสัมบูรณ์เปอร์เซ็นต์ความคลาดเคลื่อน
6. เปรียบเทียบประสิทธิภาพของวิธี MMTA กับวิธีป่าสุ่ม
7. เปรียบเทียบประสิทธิภาพของวิธี KFTA กับวิธีป่าสุ่ม
8. สรุปขั้นตอนการดำเนินงาน



รูปที่ 1 ขั้นตอนการทำวิธี Multi-Model Tree Averaging



รูปที่ 2 ขั้นตอนการทำวิธี K-Fold Tree Averaging

สำหรับวิธี MMTA จะเริ่มต้นจากการแบ่งชุดข้อมูลออกเป็นชุดฝึกสอนและชุดทดสอบ แบบคงที่ เช่น 80:20 จากนั้นทำการสร้างโมเดลต้นไม้การถดถอย โดยใช้ maxdepth ที่แตกต่างกันหลายค่า เพื่อศึกษาผลของระดับความซับซ้อนของโมเดลต่อประสิทธิภาพในการพยากรณ์ ทั้งนี้มีการกำหนดพารามิเตอร์ร่วมในการสร้างต้นไม้เพิ่มเติม ได้แก่ minsplit = 5 ซึ่งหมายถึงจำนวนข้อมูลขั้นต่ำในโหนดที่จำเป็นต้องมีเพื่อให้เกิดการแบ่ง ช่วยลดการแตกโหนดโดยไม่จำเป็น cp = 0.001 คือ complexity parameter ใช้ควบคุมการเติบโตของต้นไม้ หากค่าลดลงต่ำกว่าค่านี้จะไม่แตกโหนดเพิ่ม xval = 5 คือจำนวนรอบของ cross-validation ภายในต้นไม้เพื่อประเมินโครงสร้างโมเดลย่อย

โมเดลที่ได้จากแต่ละค่า maxdepth จะมีลักษณะและผลพยากรณ์ที่ต่างกัน จากนั้นจึงนำค่าพยากรณ์ทั้งหมดมาทำการเฉลี่ย เพื่อให้ได้ค่าพยากรณ์รวมที่มีเสถียรภาพและแม่นยำมากขึ้น

ในทางตรงกันข้าม วิธี KFTA ใช้วิธีการแบ่งข้อมูลแบบ K-Fold Cross-Validation โดยข้อมูลจะถูกแบ่งเป็น 5 ส่วนเท่า ๆ กัน ($K = 5$) และทำการฝึกสอนโมเดลต้นไม้การถดถอยจากชุดข้อมูลฝึกในแต่ละ fold โดยกำหนดพารามิเตอร์หลักให้คงที่ ได้แก่ maxdepth ความลึกของต้นไม้, minsplit และ cp เช่นเดียวกับที่ใช้ใน MMTA โมเดลที่ได้จากแต่ละ fold จะมีค่าพยากรณ์แตกต่างกันไปตามชุดข้อมูลที่ฝึกสอน และค่าพยากรณ์ทั้งหมดจาก 5 fold จะถูกนำมาเฉลี่ยเป็นค่าพยากรณ์สุดท้าย

4. ผลการวิจัยและอภิปรายผล

ผลการวิจัยแบ่งออกเป็น 3 ส่วน ได้แก่

1. ผลการศึกษาต้นไม้การถดถอยย่อยที่นำมาเฉลี่ย
2. ผลลัพธ์จากวิธีการเฉลี่ยตัวแบบบนต้นไม้การถดถอย (MMTA และ KFTA) กับวิธีป่าสุ่ม
3. การเปรียบเทียบประสิทธิภาพของค่าพยากรณ์จากวิธีการเฉลี่ยตัวแบบบนต้นไม้การถดถอย (MMTA และ KFTA) กับวิธีป่าสุ่ม

ผลการศึกษาแสดงในรูปแบบของตารางเปรียบเทียบ ซึ่งได้กำหนดสัญลักษณ์แทนความหมายต่าง ๆ ดังนี้

n	หมายถึง ขนาดตัวอย่าง
Fully grown tree	หมายถึง ต้นไม้การถดถอยที่มีขนาดใหญ่ที่สุด
Depth	หมายถึง ความลึกของต้นไม้การถดถอย
maxdepth	หมายถึง ความลึกสูงสุดที่ต้นไม้สามารถเติบโตได้
minsplit	หมายถึง จำนวนข้อมูลขั้นต่ำที่ต้องมีในโหนดก่อนที่จะแบ่ง
cp	หมายถึง ค่าพารามิเตอร์ที่ใช้ควบคุมความซับซ้อนของโมเดล ยิ่งค่าสูง ยิ่งลดการแตกแขนงของต้นไม้

Multi-Model Tree Averaging (MMTA)

หมายถึง วิธีการพยากรณ์ที่เฉลี่ยค่าจากหลายตัวแบบต้นไม้การถดถอย โดยใช้การถ่วงน้ำหนักตัวแบบตามความแม่นยำของการพยากรณ์

K-Fold Tree Averaging (KFTA)

หมายถึง วิธีการพยากรณ์ที่เฉลี่ยค่าจากหลายตัวแบบต้นไม้การถดถอย โดยใช้ K-Fold Cross-Validation เพื่อประเมินความแม่นยำของการพยากรณ์ในแต่ละความลึก

4.1 ต้นไม้การถดถอยย่อยที่นำมาเฉลี่ย

ทำการศึกษาต้นไม้การถดถอยย่อยหลายต้น ซึ่งจะถูกนำมาเฉลี่ยและทำการคัดเลือกตัวแปรอิสระของต้นไม้แต่ละต้นย่อย MMTA ภายใต้การทดสอบแบบไขว้ โดยกำหนดค่า K=5 และ KFTA ภายใต้การแบ่งชุดข้อมูลแบบไขว้ โดยกำหนดค่า K=5

4.1.1 ผลลัพธ์จากต้นไม้การถดถอยสำหรับวิธี Multi-Model Tree Averaging (MMTA)

ผลการวิเคราะห์ต้นไม้การถดถอยด้วยวิธี MMTA แสดงให้เห็นถึงความสามารถในการพยากรณ์ราคาทองคำได้อย่างมีประสิทธิภาพ โดยใช้ค่าพยากรณ์จากตัวแบบหลายตัวที่ถ่วงน้ำหนักตามประสิทธิภาพของแต่ละตัวแบบ ซึ่งวัดจากตัวชี้วัดหลัก ได้แก่ R^2 , RMSE และ MAPE

ตารางที่ 4.1 ต้นไม้การถดถอยย่อย สำหรับวิธี Multi-Model Tree Averaging

ต้นไม้การถดถอยของ	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R^2	RMSE	MAPE	R^2	RMSE	MAPE
1. Fully Grown Tree	0.94	2.51	4.15%	0.86	3.55	5.12%
2. Depth 1	0.81	4.32	7.64%	0.86	3.76	6.09%
3. Depth 2	0.90	3.20	5.62%	0.80	4.33	7.03%
4. Depth 3	0.93	2.54	4.50%	0.87	3.42	5.40%
5. Depth 4	0.94	2.32	3.89%	0.87	3.36	5.13%
6. Depth 5	0.95	2.26	3.69%	0.88	3.31	4.88%
7. Depth 6	0.95	2.26	3.69%	0.88	3.31	4.88%

หมายเหตุ: ตัวเลขหนา หมายถึง ต้นไม้การถดถอยย่อยที่นำมาเฉลี่ย

จากตารางที่ 4.1 พบว่าตัวแบบที่มี Depth 5 ให้ค่า R^2 สูงที่สุด ซึ่งสะท้อนถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้อย่างมีประสิทธิภาพ ค่า RMSE และ MAPE ของตัวแบบเหล่านี้ก็อยู่ในระดับต่ำที่สุด โดยเฉพาะในชุดข้อมูลทดสอบซึ่งแสดงให้เห็นว่าตัวแบบสามารถพยากรณ์ข้อมูลใหม่ได้อย่างแม่นยำและเสถียร

ในทางกลับกัน ตัวแบบ Fully Grown Tree ซึ่งไม่มีการจำกัดความลึก แม้ว่าจะให้ค่า R^2 สูงในชุดข้อมูลฝึกสอน แต่เมื่อทำการทดสอบในชุดข้อมูลทดสอบพบว่าค่า RMSE และ MAPE สูงขึ้นอย่างมีนัยสำคัญ ซึ่งบ่งชี้ถึงปัญหาการเข้ากันเกินไปที่ทำให้ตัวแบบสูญเสียความสามารถในการพยากรณ์ข้อมูลใหม่ได้อย่างแม่นยำ

การใช้วิธี MMTA ช่วยลดปัญหานี้ โดยการรวมค่าพยากรณ์จากตัวแบบหลายตัว ส่งผลให้ค่าพยากรณ์มีความเสถียรและเชื่อถือได้มากยิ่งขึ้น

ตารางที่ 4.2 ตัวแปรอิสระที่เกี่ยวข้องในการสร้างต้นไม้การถดถอยย่อย สำหรับวิธี Multi-Model Tree

ต้นไม้การถดถอยของ MMTA	ตัวแปรอิสระ
1. Fully grown tree	China Inflation, China Interest, Thai Inflation, US-CPI, US-PCE
2. Depth 1	China Interest
3. Depth 2	China Inflation, China Interest, Thai Inflation
4. Depth 3	China Inflation, China Interest, Thai Inflation, US-CPI, US-PCE
5. Depth 4	China Inflation, China Interest, Thai Inflation, US-CPI, US-PCE
6. Depth 5	China Inflation, China Interest, Thai Inflation, US-CPI, US-PCE
7. Depth 6	China Inflation, China Interest, Thai Inflation, US-CPI, US-PCE

หมายเหตุ: ตัวเลขหนา หมายถึง ต้นไม้การถดถอยย่อยที่นำมาเฉลี่ย

4.1.2 ผลลัพธ์จากต้นไม้การถดถอย สำหรับวิธี K-Fold Tree Averaging (KFTA)

ผลการวิเคราะห์ต้นไม้การถดถอยด้วยวิธี KFTA แสดงให้เห็นถึงประสิทธิภาพในการพยากรณ์ข้อมูลราคาทองคำ โดยใช้กระบวนการ K-Fold Cross-Validation เพื่อเฉลี่ยค่าพยากรณ์จากตัวแบบหลายตัว ซึ่งแต่ละตัวแบบถูกสร้างขึ้นจากชุดข้อมูลที่แบ่งเป็น K ส่วนอย่างเป็นระบบ

ตารางที่ 4.3 ต้นไม้การถดถอยย่อยเฉลี่ย สำหรับวิธี K-Fold Tree Averaging

ต้นไม้การถดถอยของ KFTA	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R^2	RMSE	MAPE	R^2	RMSE	MAPE
1. Fully Grown Tree	0.94	2.43	4.00%	0.88	3.51	5.77%
2. Depth 1	0.82	4.21	7.30%	0.79	4.50	7.70%
3. Depth 2	0.89	3.26	5.52%	0.82	4.26	7.40%
4. Depth 3	0.94	2.38	4.01%	0.88	3.44	6.13%
5. Depth 4	0.97	1.82	3.11%	0.90	3.16	5.25%
6. Depth 5	0.98	1.38	2.40%	0.90	3.17	5.01%
7. Depth 6	0.98	1.27	2.20%	0.89	3.28	5.11%
8. Depth 7	0.98	1.23	2.10%	0.89	3.29	5.14%

หมายเหตุ: ตัวเลขหนา หมายถึง ต้นไม้การถดถอยย่อยที่นำมาเฉลี่ย

จากผลลัพธ์ที่แสดงในตารางที่ 4.3 เมื่อใช้วิธี KFTA ในการเฉลี่ยต้นไม้การถดถอยย่อยจากทุก Fold พบว่าตัวแบบที่ Depth 5 ให้

ค่า R^2 สูงที่สุด ซึ่งสะท้อนถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ดี นอกจากนี้ ตัวแบบนี้ยังให้ค่า RMSE และ MAPE ที่ต่ำในชุดข้อมูลทดสอบแสดงให้เห็นว่าตัวแบบสามารถพยากรณ์ข้อมูลใหม่ได้อย่างแม่นยำและเสถียร

ในทางกลับกัน ตัวแบบ Fully Grown Tree ที่ไม่มีการจำกัดความลึก แม้ว่าจะให้ค่า R^2 สูงในชุดข้อมูลฝึกสอน แต่กลับมีค่า RMSE และ MAPE ที่สูงขึ้นในชุดข้อมูลทดสอบ ซึ่งสะท้อนถึงปัญหา การเข้ากันเกินไป ที่ทำให้ตัวแบบสูญเสียความสามารถในการพยากรณ์ข้อมูลใหม่

การใช้วิธี KFTA ช่วยลดปัญหานี้ได้โดยการรวมค่าพยากรณ์จากหลายตัวแบบที่ผ่านการประเมินความเสถียรในแต่ละรอบของ K-Fold ส่งผลให้ค่าพยากรณ์มีความเสถียรและแม่นยำมากยิ่งขึ้น

ตารางที่ 4.4 ตัวแปรอิสระที่เกี่ยวข้องในการสร้างต้นไม้การถดถอยย่อย สำหรับวิธี K-Fold Tree Averaging

ต้นไม้การถดถอยของ KFTA	ตัวแปรอิสระ				
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
1. Fully grown tree	CN-INF, IN-INF, TH-INF, US-CPI, US-PCE	CN-INF, CN-INT, TH-INF, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PCE
2. Depth 1	US-CPI	US-CPI	CN-INT	CN-INT	CN-INT
3. Depth 2	CN-INF, TH-INF, US-CPI	TH-INF, US-CPI, US-PPI	CN-INF, CN-INT, TH-INF	CN-INF, CN-INT, US-CPI	CN-INF, CN-INT, TH-INF
4. Depth 3	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PCE	CN-INT, IN-INF, TH-INF, US-CPI, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE
5. Depth 4	CN-INF, CN-INT, IN-INF, TH-INF, TH-INT, US-CPI, US-PCE	CN-INT, IN-INF, TH-INF, US-CPI, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE
6. Depth 5	CN-INF, CN-INT, IN-INF, TH-INF, TH-INT, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-PCE

ต้นไม้การถดถอยของ KFTA	ตัวแปรอิสระ				
	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
	IN-INF, IN-INF, TH-INF, TH-INT, US-CPI, US-PCE	TH-INF, US-CPI, US-PPI	TH-INF, US-CPI, US-INT, US-PCE, US-PPI	TH-INF, TH-INT, US-CPI, US-INT, US-PCE	TH-INT, US-CPI, US-INT, US-PCE
7. Depth 6	CN-INF, CN-INT, IN-INF, IN-INF, TH-INF, TH-INT, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-INT, US-PCE, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-INT, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-INT, US-PCE
8. Depth 7	CN-INF, CN-INT, IN-INF, TH-INF, TH-INT, US-CPI, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-INT, US-PCE, US-PPI	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-INT, US-PCE	CN-INF, CN-INT, IN-INF, TH-INF, US-CPI, US-INT, US-PCE

หมายเหตุ: ตัวเลขหนา หมายถึง ต้นไม้การถดถอยย่อยที่นำมาเฉลี่ย

คำอธิบายอักษรย่อของตัวแปรในตาราง

- อัตราเงินเฟ้อของไทย: Thai Inflation Rate (TH-INF)
- อัตราดอกเบี้ยของไทย: Thai Interest Rate (TH-INT)
- อัตราเงินเฟ้อของอินเดีย: India Inflation Rate (IN-INF)
- อัตราดอกเบี้ยของอินเดีย: India Interest Rate (IN-INT)

5. อัตราเงินเฟ้อของจีน: China Inflation Rate (CN-INF)
6. อัตราดอกเบี้ยของจีน: China Interest Rate (CN-INT)
7. ดัชนีราคาผู้บริโภคของสหรัฐฯ: Consumer Price Index (US-CPI)
8. การใช้จ่ายเพื่อการบริโภคส่วนบุคคลของสหรัฐฯ: Personal Consumption Expenditure (US-PCE)
9. ดัชนีราคาผู้ผลิตของสหรัฐฯ: Producer Price Index (US-PPI)
10. อัตราดอกเบี้ยของสหรัฐฯ: United States Interest Rate (US-INT)

4.2 การเปรียบเทียบประสิทธิภาพค่าพยากรณ์ของ

วิธีการเฉลี่ยตัวแบบบนต้นไม้การถดถอยทั้ง 2 วิธี กับวิธีป่าสุ่ม

การเปรียบเทียบประสิทธิภาพระหว่าง MMTA และ KFTA และวิธีป่าสุ่ม แสดงให้เห็นถึงแนวทางที่คล้ายคลึงกันในการใช้ตัวแบบหลายตัวเพื่อเพิ่มความแม่นยำในการพยากรณ์ โดยทั้งสามวิธีมุ่งเน้นการลดความคลาดเคลื่อนและเพิ่มความเสถียรของผลลัพธ์ อย่างไรก็ตาม วิธีการแต่ละแบบมีลักษณะเฉพาะและข้อดีที่แตกต่างกัน

MMTA ใช้หลักการถ่วงน้ำหนักตัวแบบโดยอิงจากเกณฑ์เกณฑ์น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะ ซึ่งช่วยให้ตัวแบบที่มีประสิทธิภาพสูงสุดมีผลต่อค่าพยากรณ์รวมมากที่สุด ในขณะที่ KFTA ใช้กระบวนการ K-Fold Cross-Validation แบ่งข้อมูลออกเป็นหลายส่วน และเฉลี่ยค่าพยากรณ์จากตัวแบบหลายตัวในแต่ละส่วน

ในทางกลับกัน วิธีป่าสุ่มใช้การสร้างตัวแบบต้นไม้การถดถอยจำนวนมากผ่านกระบวนการสุ่มตัวอย่าง และเฉลี่ยผลลัพธ์จากตัวแบบทั้งหมดเพื่อเพิ่มความเสถียรและลดความผันผวนของค่าพยากรณ์

ข้อดีของ MMTA คือการลดความลำเอียงจากการเลือกตัวแบบและเพิ่มความแม่นยำผ่านการถ่วงน้ำหนักตามประสิทธิภาพของตัวแบบ ขณะที่ KFTA มี

ความสามารถในการประเมินความเสถียรของตัวแบบได้ดีผ่าน

4.2.2 ผลลัพธ์จากวิธี KFTA

อย่างไรก็ตาม MMTA และ KFTA มีข้อจำกัดด้านความซับซ้อนของกระบวนการคำนวณ โดยเฉพาะเมื่อใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะในการถ่วงน้ำหนักตัวแบบ ขณะที่วิธีป่าสุ่มอาจมีข้อเสียในแง่ของการตีความผลลัพธ์เนื่องจากเป็นตัวแบบแบบ black box ที่ไม่สามารถอธิบายรายละเอียดของการตัดสินใจได้อย่างชัดเจน

สรุป ทั้งสามวิธีมีข้อดีและข้อเสียที่แตกต่างกัน การเลือกใช้งานจึงขึ้นอยู่กับลักษณะของข้อมูลและเป้าหมายของการพยากรณ์ โดยวิธีป่าสุ่มมักเหมาะสมสำหรับข้อมูลที่ซับซ้อนและไม่เป็นเชิงเส้น ขณะที่ MMTA และ KFTA เหมาะสำหรับการวิเคราะห์ที่ต้องการการควบคุมและประเมินตัวแบบในหลายมิติอย่างละเอียด

4.2.1 ผลลัพธ์จากวิธี Multi-Model Tree

Averaging

ตารางที่ 4.5 การเปรียบเทียบประสิทธิภาพของวิธีการหาค่าพยากรณ์ สำหรับวิธี Multi-Model Tree Averaging

วิธีของ MMTA	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R ²	RMSE	MAPE	R ²	RMSE	MAPE
Equal Weight	0.94	2.52	4.36%	0.87	3.32	5.06%
AIC Weight	0.94	2.44	4.16%	0.88	3.30	5.03%

การวิเคราะห์ผลลัพธ์ในตารางที่ 4.5 แสดงให้เห็นว่า การใช้เกณฑ์น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะ ในวิธี MMTA ให้ผลลัพธ์ที่ดีกว่าการเฉลี่ยแบบเท่ากันในหลายมิติ โดยพิจารณาจากตัวชี้วัดหลัก ได้แก่ R², RMSE, และ MAPE ในชุดข้อมูลฝึกสอนตัวแบบที่ใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะแสดงค่า R² ที่สูงกว่าน้ำหนักที่เท่ากันโดยมีค่า 0.94 สะท้อนถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ดียิ่งขึ้น นอกจากนี้ ค่า RMSE และ MAPE ของน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะยังต่ำกว่าที่ 2.44 และ 4.16% ตามลำดับ แสดงถึงความแม่นยำที่เพิ่มขึ้นในชุดข้อมูลฝึกสอน

ในชุดข้อมูลทดสอบค่า R² ของน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะอยู่ที่ 0.88 ซึ่งสูงกว่าน้ำหนักที่

เท่ากันที่มีค่า 0.87 เล็กน้อย บ่งบอกถึงความสามารถในการพยากรณ์ข้อมูลใหม่ที่ดีกว่า ขณะที่ค่า RMSE และ MAPE ของน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะอยู่ที่ 3.30 และ 5.03% ซึ่งลดลงเล็กน้อยเมื่อเทียบกับน้ำหนักที่เท่ากันที่มีค่า 3.32 และ 5.06% ตามลำดับ การลดลงของข้อผิดพลาดเหล่านี้แสดงให้เห็นถึงความแม่นยำในการพยากรณ์ข้อมูลใหม่ที่เพิ่มขึ้น

โดยสรุป การใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะในวิธี MMTA สามารถเพิ่มประสิทธิภาพของการพยากรณ์ได้อย่างชัดเจน ทั้งในด้านความสามารถในการอธิบายข้อมูลและการลดข้อผิดพลาด ทำให้วิธีนี้มีความเหมาะสมสำหรับการพยากรณ์ราคาทองคำที่มีความผันผวนสูง

4.2.2 ผลลัพธ์จากวิธี K-Fold Tree Averaging

ตารางที่ 4.6 การเปรียบเทียบประสิทธิภาพของวิธีหาค่าพยากรณ์ สำหรับวิธี K-Fold Tree Averaging

วิธีของ KFTA	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R ²	RMSE	MAPE	R ²	RMSE	MAPE
Equal Weight fold: 1	0.97	1.72	2.84%	0.91	3.08	4.65%
AIC Weight fold: 1	0.98	1.29	2.11%	0.91	2.99	4.60%
Equal Weight fold: 2	0.97	1.72	2.84%	0.91	3.08	4.65%
AIC Weight fold: 2	0.98	1.29	2.11%	0.91	2.99	4.60%
Equal Weight fold: 3	0.97	1.89	3.11%	0.94	2.22	4.07%
AIC Weight fold: 3	0.98	1.33	2.15%	0.93	2.60	4.95%
Equal Weight fold: 4	0.97	1.64	2.86%	0.83	4.46	7.38%
AIC Weight fold: 4	0.99	1.16	2.12%	0.80	4.80	6.82%
Equal Weight fold: 5	0.97	1.59	2.59%	0.90	3.12	5.34%
AIC Weight fold: 5	0.99	1.13	1.90%	0.93	2.74	4.69%

การวิเคราะห์ผลลัพธ์ในตารางที่ 4.6 แสดงให้เห็นว่าการใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะในวิธี KFTA ให้ผลลัพธ์ที่ดีกว่าการเฉลี่ยแบบเท่ากันในบางกรณี โดยพิจารณาจากตัวชี้วัดหลัก ได้แก่ R², RMSE, และ

MAPE ในชุดข้อมูลฝึกสอนตัวแบบที่ใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะใน fold ที่ 5 แสดงค่า R² ที่สูงถึง 0.99 ซึ่งสะท้อนถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ดียิ่งขึ้นเมื่อเทียบกับน้ำหนักที่เท่ากันที่มีค่า R² ต่ำกว่าเล็กน้อย นอกจากนี้ค่า RMSE และ MAPE ของน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะยังต่ำกว่าที่ 1.13 และใกล้ 1.90% ตามลำดับ แสดงถึงความแม่นยำที่เพิ่มขึ้นในชุดข้อมูลฝึกสอน

ในชุดข้อมูลทดสอบค่า R² ของน้ำหนักที่เท่ากันอยู่ที่ 0.94 ซึ่งสูงกว่าน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะที่มีค่า 0.93 บ่งบอกถึงความสามารถในการพยากรณ์ข้อมูลใหม่ที่ดีกว่า ขณะที่ค่า RMSE และ MAPE ของน้ำหนักที่เท่ากันอยู่ที่ต่ำกว่า 2.22 และ 4.07% ใน fold ที่ 3 ซึ่งดีกว่าค่าที่ได้จากน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะอย่างชัดเจน การลดลงของข้อผิดพลาดเหล่านี้แสดงให้เห็นถึงความแม่นยำที่เพิ่มขึ้นในการพยากรณ์ข้อมูลใหม่

โดยสรุป การใช้น้ำหนักที่เท่ากันในวิธี KFTA สามารถเพิ่มประสิทธิภาพของการพยากรณ์ได้อย่างชัดเจน ทั้งในด้านความสามารถในการอธิบายข้อมูลและการลดข้อผิดพลาด ทำให้วิธีนี้มีความเหมาะสมสำหรับการพยากรณ์ราคาทองคำที่มีความผันผวนสูง

ตารางที่ 4.7 การเปรียบเทียบประสิทธิภาพของวิธีหาค่าพยากรณ์เฉลี่ย สำหรับวิธี K-Fold Tree Averaging

วิธี	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R ²	RMSE	MAPE	R ²	RMSE	MAPE
Equal Weight	0.97	1.71	2.85%	0.90	3.19	5.22%
AIC Weight	0.98	1.24	2.08%	0.90	3.22	5.13%

4.2.3 ผลลัพธ์จากวิธีป่าสุ่ม

ตารางที่ 4.8 การเปรียบเทียบประสิทธิภาพของวิธีการหาค่าพยากรณ์ สำหรับวิธีป่าสุ่ม

วิธี	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R ²	RMSE	MAPE	R ²	RMSE	MAPE
ป่าสุ่ม	0.99	0.87	1.39%	0.95	2.34	4.13%

จากการวิเคราะห์ผลลัพธ์ในตารางที่ 4.8 สำหรับวิธีป่าสุ่มตัวแบบสามารถพยากรณ์ข้อมูลราคาทองคำได้อย่างมี

ประสิทธิภาพ โดยใน ชุดข้อมูลฝึกสอนตัวแบบแสดงค่า R^2 สูงถึง 0.99 ซึ่งบ่งชี้ว่าตัวแบบสามารถอธิบายความแปรปรวนของข้อมูลได้เกือบทั้งหมด อีกทั้งค่า RMSE และ MAPE อยู่ในระดับต่ำเพียง 0.87 และ 1.39% ตามลำดับ ซึ่งสะท้อนถึงความแม่นยำสูงและข้อผิดพลาดต่ำในชุดข้อมูลฝึกสอน

สำหรับชุดข้อมูลทดสอบค่า R^2 เท่ากับ 0.95 ซึ่งยังคงแสดงถึงความสามารถในการพยากรณ์ที่ดี แม้ว่าจะลดลงเล็กน้อยเมื่อเทียบกับ ชุดข้อมูลฝึกสอน ส่วนค่า RMSE อยู่ที่ 2.34 และ MAPE เท่ากับ 4.13% แม้ว่าค่าทั้งสองนี้จะสูงกว่าที่ได้ในชุดข้อมูลฝึกสอน แต่ยังคงอยู่ในเกณฑ์ที่ยอมรับได้สำหรับการพยากรณ์ข้อมูลใหม่

สรุป วิธีป่าสุ่มมีประสิทธิภาพสูงในการพยากรณ์ราคาทองคำ โดยสามารถรักษาความแม่นยำและลดปัญหาการเข้ากันเกินไปได้ดี เหมาะสำหรับข้อมูลที่มีความผันผวนสูง

4.3 การเปรียบเทียบประสิทธิภาพของค่าพยากรณ์

4.3.1 เปรียบเทียบประสิทธิภาพของวิธีการหาค่าพยากรณ์ สำหรับวิธี Multi-Model Tree Average กับวิธีป่าสุ่ม

จากการเปรียบเทียบประสิทธิภาพของวิธี MMTA และวิธีป่าสุ่มในการพยากรณ์ราคาทองคำ พบว่า วิธีป่าสุ่มมีความแม่นยำสูงกว่าในด้านค่า R^2 , RMSE และ MAPE ตามผลลัพธ์ที่แสดงในตารางที่ 4.5 และ 4.8

ในชุดข้อมูลฝึกสอนวิธีป่าสุ่มแสดงค่า R^2 สูงถึง 0.99 และ RMSE ต่ำเพียง 0.87 ซึ่งสะท้อนถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลและลดข้อผิดพลาดได้อย่างมีประสิทธิภาพ ในขณะที่ MMTA แม้จะมีค่า R^2 ที่ใกล้เคียงกันในกรณีของน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะ (0.94) แต่ค่า RMSE และ MAPE ยังสูงกว่า โดย RMSE เท่ากับ 2.44 และ MAPE อยู่ที่ 2.08%

สำหรับชุดข้อมูลทดสอบค่า R^2 ของวิธีป่าสุ่มอยู่ที่ 0.95 ซึ่งสูงกว่า MMTA ที่มีค่าสูงสุดในกรณีน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะ 0.90 นอกจากนี้ ค่า RMSE และ MAPE ของวิธีป่าสุ่มในชุดข้อมูลทดสอบ

ยังอยู่ในระดับที่ต่ำกว่า โดยมีค่า RMSE เท่ากับ 2.34 และ MAPE อยู่ที่ 4.13% ซึ่งดีกว่าค่า RMSE และ MAPE ของ MMTA ในทั้งกรณีน้ำหนักที่เท่ากันและน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะที่มีค่าสูงกว่าในชุดข้อมูลทดสอบ

วิธีป่าสุ่มแสดงให้เห็นถึงความสามารถในการพยากรณ์ที่เหนือกว่าในทุกด้าน ทั้งในชุดข้อมูลฝึกสอนและทดสอบ โดยสามารถพยากรณ์ข้อมูลใหม่ได้อย่างแม่นยำและมีความเสถียรสูงกว่า MMTA ทำให้เป็นตัวเลือกที่เหมาะสมสำหรับการพยากรณ์ราคาทองคำที่มีความผันผวนสูง

4.3.2 เปรียบเทียบประสิทธิภาพของวิธีการหาค่าพยากรณ์ สำหรับวิธี K-Fold Tree Average กับวิธีป่าสุ่ม

จากการเปรียบเทียบประสิทธิภาพของวิธี K-Fold Tree Average (KFTA) และวิธีป่าสุ่มในการพยากรณ์ราคาทองคำ พบว่าวิธีป่าสุ่มมีความแม่นยำสูงกว่าในด้านค่า R^2 , RMSE และ MAPE ตามผลลัพธ์ที่แสดงในตารางที่ 4.7 และ 4.8

ในชุดข้อมูลฝึกสอน วิธีป่าสุ่มแสดงค่า R^2 สูงถึง 0.99 และ RMSE ต่ำเพียง 0.87 ซึ่งสะท้อนถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลและลดข้อผิดพลาดได้อย่างมีประสิทธิภาพ ในขณะที่ KFTA แม้จะมีค่า R^2 ที่ใกล้เคียงกันในกรณีของน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะ (0.98) แต่ค่า RMSE และ MAPE ยังสูงกว่า โดย RMSE เท่ากับ 1.24 และ MAPE อยู่ที่ 2.08%

สำหรับชุดข้อมูลทดสอบค่า R^2 ของวิธีป่าสุ่มอยู่ที่ 0.95 ซึ่งสูงกว่า KFTA ที่มีค่าสูงสุดในกรณีน้ำหนักที่เท่ากันที่ 0.90 นอกจากนี้ ค่า RMSE และ MAPE ของวิธีป่าสุ่มในชุดข้อมูลทดสอบ ยังอยู่ในระดับที่ต่ำกว่า โดยมีค่า RMSE เท่ากับ 2.34 และ MAPE อยู่ที่ 4.13% ซึ่งดีกว่าค่า RMSE และ MAPE ของ KFTA ในทั้งกรณีน้ำหนักที่เท่ากันและน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะที่มีค่าสูงกว่าในชุดข้อมูลทดสอบ

สรุป วิธีป่าสุ่มแสดงให้เห็นถึงความสามารถในการพยากรณ์ที่เหนือกว่าในทุกด้าน ทั้งในชุดข้อมูลฝึกสอนและทดสอบ โดยสามารถพยากรณ์ข้อมูลใหม่ได้อย่างแม่นยำและมีความเสถียรสูงกว่า KFTA ทำให้เป็นตัวเลือกที่เหมาะสมสำหรับการพยากรณ์ราคาทองคำที่มีความผันผวนสูง

4.3.3 การวิเคราะห์ผลลัพธ์ของวิธี Multi-Model Tree Averaging และ K-Fold Tree Averaging เปรียบเทียบกับวิธีป่าสุ่ม

การวิเคราะห์ผลลัพธ์จากการใช้ MMTA และ KFTA เปรียบเทียบกับวิธีป่าสุ่ม ในการพยากรณ์ราคาทองคำพบว่าวิธีป่าสุ่มมีประสิทธิภาพดีกว่าในหลายด้าน ได้แก่ R^2 , RMSE, และ MAPE

ตารางที่ 4.9 การเปรียบเทียบประสิทธิภาพของวิธีการหาค่าพยากรณ์ของวิธี Multi-Model Tree Averaging และ K-Fold Tree Averaging เปรียบเทียบกับวิธีป่าสุ่ม

วิธี	ชุดข้อมูลฝึกสอน			ชุดข้อมูลทดสอบ		
	R^2	RMSE	MAPE	R^2	RMSE	MAPE
Equal Weight MMTA	0.94	2.52	4.36%	0.87	3.32	5.06%
AIC Weight MMTA	0.94	2.44	4.16%	0.88	3.30	5.03%
Equal Weight KFTA	0.97	1.71	2.85%	0.90	3.19	5.22%
AIC Weight KFTA	0.98	1.24	2.08%	0.90	3.22	5.13%
ป่าสุ่ม	0.99	0.87	1.39%	0.95	2.34	4.13%

ในชุดข้อมูลฝึกสอนวิธีป่าสุ่มแสดงค่า R^2 สูงถึง 0.99 ซึ่งบ่งชี้ถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้อย่างมีประสิทธิภาพ และมีค่า RMSE ต่ำเพียง 0.87 ซึ่งสะท้อนถึงการลดข้อผิดพลาดในการพยากรณ์ได้ดีมาก ในขณะที่น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะของวิธี MMTA แสดงค่า R^2 เท่ากับ 0.94 และ RMSE เท่ากับ 2.44 ซึ่งแตกต่างจากวิธีป่าสุ่ม ส่วนน้ำหนักที่เท่ากันของวิธี KFTA แสดงค่า R^2 เท่ากับ 0.97 และ RMSE เท่ากับ 1.71 ซึ่งยังคงต่ำกว่าวิธีป่าสุ่มในด้านประสิทธิภาพการพยากรณ์

สำหรับชุดข้อมูลทดสอบ วิธีป่าสุ่มยังคงแสดงผลลัพธ์ที่ดีกว่า โดย R^2 ของวิธีป่าสุ่มอยู่ที่ 0.95 ซึ่งสูงกว่าน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะของวิธี MMTA (0.88) และ น้ำหนักที่เท่ากันของวิธี KFTA (0.90) นอกจากนี้ RMSE ของวิธีป่าสุ่มในชุดข้อมูลทดสอบอยู่ที่ 2.34 และ MAPE อยู่ที่ 4.13% ซึ่งต่ำกว่าน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะของวิธี MMTA และ

น้ำหนักที่เท่ากันของวิธี KFTA ในทั้งสองกรณีที่มีค่า RMSE และ MAPE สูงกว่า

สรุป วิธีป่าสุ่มแสดงให้เห็นถึงความสามารถในการพยากรณ์ที่เหนือกว่าในทุกด้านทั้งในชุดข้อมูลฝึกสอนและทดสอบ โดยสามารถพยากรณ์ข้อมูลใหม่ได้อย่างแม่นยำและมีความเสถียรสูงกว่า MMTA และ KFTA ทั้งในกรณีที่ใช้น้ำหนักที่เท่ากันและน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะของวิธี ทำให้วิธีป่าสุ่มเป็นตัวเลือกที่เหมาะสมสำหรับการพยากรณ์ราคาทองคำที่มีความผันผวนสูง

อย่างไรก็ตาม การใช้วิธีป่าสุ่มก็มีประเด็นที่ต้องพิจารณาเกี่ยวกับปัญหาของการเข้ากันเกินไปซึ่งอาจเกิดขึ้นได้ โดยเฉพาะใน วิธีป่าสุ่มสร้างจากการสุ่มสร้างต้นไม้หลายๆ ต้น ซึ่งอาจมีการพยากรณ์ที่แม่นยำเกินไปในชุดข้อมูลฝึกสอน แต่ในบางกรณีก็อาจทำให้ประสิทธิภาพลดลงเมื่อทดสอบกับข้อมูลใหม่

อย่างไรก็ตาม วิธีป่าสุ่มสามารถลดการเข้ากันเกินไปได้ โดยการสร้างหลายๆ ต้นไม้และการสุ่มเลือกตัวแปรในการแบ่งแต่ละต้นไม้ ซึ่งช่วยให้ตัวแบบมีความยืดหยุ่นและลดความเสี่ยงจากการที่ต้นไม้บางต้นอาจจะการเข้ากันเกินไป ในขณะที่ MMTA และ KFTA แม้จะใช้การเฉลี่ยของหลายๆ ตัวแบบเพื่อลดความซับซ้อนของการพยากรณ์ แต่ในบางกรณีก็ยังมีโอกาสเกิดการเข้ากันเกินไป หากมีการเลือก features หรือ hyperparameters ที่ไม่เหมาะสม

5. สรุปผลการศึกษาและข้อเสนอแนะ

5.1 สรุปผลการศึกษา

จากการวิเคราะห์และเปรียบเทียบวิธีการพยากรณ์พบว่า MMTA และ KFTA ทั้งในแบบน้ำหนักที่เท่ากันและน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะมีประสิทธิภาพที่ดีในชุดข้อมูลฝึกสอน โดยเฉพาะในกรณีที่ใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกะอิเกะซึ่งสามารถลดความคลาดเคลื่อนได้ดี แต่ยังคงมีข้อจำกัดในชุดข้อมูลทดสอบ เนื่องจากค่าพยากรณ์ยังมีข้อผิดพลาดเล็กน้อยเมื่อเทียบกับวิธีป่าสุ่ม วิธีป่าสุ่มแสดงผลลัพธ์ที่โดดเด่นที่สุดในทุกตัวชี้วัด โดยให้ค่า R^2 สูงสุดใน ทั้งชุดข้อมูลฝึกสอนและทดสอบ สะท้อนถึงความสามารถในการพยากรณ์ที่แม่นยำและมี

เสถียรภาพมากกว่า แม้ว่าจะมีข้อจำกัดด้านการตีความ เนื่องจากเป็น black box model

การใช้ KFTA ช่วยเพิ่มความแม่นยำและลดความคลาดเคลื่อนในชุดข้อมูลทดสอบโดยเฉพาะเมื่อใช้น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกอะเกเพื่อปรับน้ำหนักตัวแบบ

จากการวิเคราะห์ค่าตัวชี้วัด RMSE, MAE, และ MAPE พบว่าวิธีที่ใช้การเฉลี่ยตัวแบบสามารถลดปัญหาการเข้ากันเกินไปได้ดีในกรณีที่มีการปรับปรุงการเลือกตัวแบบ

5.2 ข้อเสนอแนะ

5.2.1 ข้อเสนอแนะเกี่ยวกับการวิจัย

1. ควรทำการศึกษาเพิ่มเติมเกี่ยวกับการปรับปรุงการเลือกตัวแบบ โดยเฉพาะการพัฒนาเกณฑ์การถ่วงน้ำหนักที่เหมาะสมสำหรับข้อมูลที่มีลักษณะต่างกัน เช่น น้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของเบย์มาเปรียบเทียบกับน้ำหนักที่ขึ้นกับเกณฑ์สารสนเทศของอะกอะเกเพื่อดูผลกระทบต่อประสิทธิภาพของการพยากรณ์

2. ควรศึกษาวิธีการพยากรณ์ในบริบทอื่น ๆ ที่ไม่ใช่ราคาทองคำ เช่น การพยากรณ์ในตลาดหุ้นหรือสินค้าโภคภัณฑ์อื่น ๆ เพื่อวิเคราะห์ความยืดหยุ่นและประสิทธิภาพของแต่ละวิธี

5.2.2 ข้อเสนอแนะเกี่ยวกับการนำไปประยุกต์ใช้

1. การนำวิธีการเฉลี่ยตัวแบบไปใช้กับข้อมูลจริง ควรพิจารณาลักษณะของข้อมูลและเลือกตัวแบบที่เหมาะสมที่สุด โดยไม่ควรยึดติดกับวิธีใดวิธีหนึ่ง เพื่อให้ได้ผลลัพธ์ที่แม่นยำและเชื่อถือได้

2. ในกรณีที่ข้อมูลมีความซับซ้อนและมีจำนวนตัวแปรมาก การใช้วิธีป่าสุ่ม อาจเป็นทางเลือกที่ดี เนื่องจากสามารถจัดการกับข้อมูลที่มีความสัมพันธ์ซับซ้อนได้อย่างมีประสิทธิภาพ

6. กิตติกรรมประกาศ

งานวิจัยฉบับนี้สามารถสำเร็จลงได้ด้วย การสนับสนุนและคำแนะนำจากหลายฝ่าย ซึ่งมีส่วนสำคัญในการทำให้การวิจัยดำเนินไปอย่างราบรื่นและสมบูรณ์

ขอแสดงความขอบคุณอย่างจริงใจต่ออาจารย์ที่ปรึกษา รองศาสตราจารย์ ดร.สุปราณี ลิขิตวิทย์, นางสาวพัชรชาพร เกษพัฒน์วระกุล นักศึกษาระดับปริญญาโทบัณฑิต คณะวิทยาศาสตร์และเทคโนโลยี สาขาวิชาสถิติประยุกต์ มหาวิทยาลัยธรรมศาสตร์, นางสาวชุตินันท์ อินทร และ นางสาวศัทธา เต่าร้าง นักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์และเทคโนโลยี สาขาวิชาสถิติประยุกต์ มหาวิทยาลัยธรรมศาสตร์ผู้ได้ให้คำแนะนำและข้อคิดเห็นที่มีคุณค่า พร้อมทั้งช่วยแนะนำแนวทางในการดำเนินการวิจัยด้วยความเอาใจใส่และเป็นประโยชน์อย่างยิ่งในการพัฒนาผลงานวิจัยนี้

7. เอกสารอ้างอิง

- [1] A. L. Samuel, "Title of the article," IBM Journal of Research and Development, vol. 3, no. 3, pp. 210–229, Jul. 1959, doi: 10.1147/rd.33.0210.
- [2] F. Galton, "Regression towards mediocrity in hereditary stature," The Journal of the Anthropological Institute of Great Britain and Ireland, vol. 15, pp. 246–263, 1886, doi: 10.2307/2841583.
- [3] M. Marmion, M. Parviainen, M. Luoto, R. K. Heikkinen, and W. Thuiller, "Evaluation of consensus methods in predictive species distribution modelling," Diversity and Distributions, vol. 15, pp. 59–69, 2009, doi: 10.1111/j.1472-4642.2008.00491.x.
- [4] G. Rapacciuolo, D. B. Roy, S. Gillings, R. Fox, K. Walker, and A. Purvis, "Climatic associations of British species distributions show good transferability in time but low predictive accuracy for range change," PLoS One, vol. 7,

- no. 7, p. e40212, Jul. 2012, doi: 10.1371/journal.pone.0040212.
- [5] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. New York, NY, USA: Springer, 2002, doi: 10.1007/b97636.
- [6] T. K. Ho, "Random Decision Forests," in *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, Montreal, QC, Canada, 14-16 Aug. 1995, pp. 278-282, doi: 10.1109/ICDAR.1995.598994.
- [7] L. Breiman, "Random Forests," in *Ensemble Machine Learning: Methods and Applications*, C. Zhang and Y. Ma, Eds. Springer, 2011, pp. 157-176, doi: 10.1007/978-1-4419-9326-7_5.
- [8] ปกรณ์ จารุตระกูลชัย, "การทำนายปริมาณสารส้มที่ใช้ในกระบวนการผลิตน้ำประปาด้วยวิธีป่าของต้นไม้ตัดสินใจและโปรแกรมเชิงพันธุกรรม," *วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์, ปทุมธานี*, 2550.
- [9] A. Sirichartchai, "Forecasting patient length of stay in hospital using machine learning algorithms," *University Tunku Abdul Rahman*, 2021.
- [10] ชนินทร์ แก้ววิบูลย์พันธุ์, "การเฉลี่ยตัวแบบบนต้นไม้การถดถอยสำหรับการพยากรณ์," *วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยธรรมศาสตร์, ปทุมธานี*, 2565.
- [11] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Stanford University*, 1995.
- [12] ศิวกร บรรลือทรัพย์, "ระบบการจำแนกผู้ป่วยโควิด-19 ด้วยเทคนิคป่าสุ่ม," *วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยธุรกิจบัณฑิตย์, กรุงเทพมหานคร*, 2564.
- [13] สุภัตสร สมเจตนา, & จาริ ทองคำ, "การเปรียบเทียบประสิทธิภาพเทคนิคเหมืองข้อมูลในการสร้างแบบจำลองเพื่อจำแนกความคิดเห็นของผู้ปกครองต่อการใช้สมาร์ตโฟนของบุตร," *วารสารวิทยาศาสตร์และเทคโนโลยี*, vol. 23, no. 1, pp. 21-30, 2564.
- [14] P. H. Putra, B. Azanuddin, B. Purba, and Y. A. Dalimunthe, "Random forest and decision tree algorithms for car price prediction," *Jurnal Matematika dan Ilmu Pengetahuan Alam*, vol. 3, no. 2, pp. 81-89, 2023.
- [15] ณาสรณ์ ญูพิบูลพาณิชย์การ, "การวิเคราะห์ปัจจัยที่มีผลต่อราคาทองคำแท่งในประเทศไทย," *วิทยานิพนธ์วิทยาศาสตร์มหาบัณฑิต, มหาวิทยาลัยกรุงเทพ, ปทุมธานี*, 2558.
- [16] T. D. Kaufmann and R. A. Winters, "The price of gold: A simple model," *Resources Policy*, vol. 15, no. 4, pp. 309-313, 1989, doi: 10.1016/0301-4207(89)90004-4.
- [17] World Gold Council, "Gold demand trends: Full year 2022," *World Gold Council*. [Online]. Available: <https://www.gold.org/goldhub/research/gold-demand-trends>. [Accessed: 30 May 2024]
- [18] S. A. Samveg, "Gold as a strategic prophecy against inflation and exchange rate," *Institute of Management and Computer Studies*, 2013.
- [19] N. P. Singh and N. Joshi, "Investigating gold investment as an inflationary hedge," *Jagan Institute of Management Studies*, 2018.
- [20] K. Q. Yaqub, "The impact of the United States macroeconomics on the price of gold," in *7th International Conference on Business, Management & Economics*, Oxford, United Kingdom, Dec. 16-18, 2022.