การประเมินเชิงเปรียบเทียบของ Random Forest, Gradient Boosting และ LSTM ในการพยากรณ์ราคาทองรูปพรรณของประเทศไทย

Comparative Evaluation of Random Forest, Gradient Boosting, and LSTM for Forecasting Gold Jewelry Prices in Thailand

พงศ์พัฒน์ ฉายศิริพันธ์[1*], ณรงค์ฤทธิ์ สุคนธสิงห์[2], อรรณพ กางกั้น[3] , ศรัญธร มั่งมี[4] และ นิวัฒน์ เตชะเกียรตินันท์[5]

Pongpat Chaisiripan[1*], Narongrit Sukonthasing[2] , Unnop Kangkan[3], Saranthon Maungmee[4] and Niwat Techakiattinun[5]

[1,2,3] อาจารย์สาขาวิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม

[1,2,3] Lecturer in Information Technology Faculty of Information Technology Siam University

[4,5] อาจารย์สาขาวิชาธุรกิจดิจิทัลและเทคโนโลยี คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยสยาม

[4,5] Lecturer in Digital Business and Technology Faculty of Information Technology Siam University

*** Corresponding author, E-mail**: pongpat.cha@siam.edu

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อประเมินและเปรียบเทียบสมรรถนะของแบบจำลอง Random Forest, Gradient Boosting และ Long Short-Term Memory (LSTM) สำหรับการพยากรณ์ราคาทองรูปพรรณขายออกในประเทศไทย โดยใช้ข้อมูลอนุกรมเวลารายวันช่วงปี พ.ศ. 2564-2569 ภายใต้กรอบการทดลองเดียวกัน ทั้งในด้านการเตรียมข้อมูล การแบ่งข้อมูลตามลำดับเวลา และการประเมินผลเชิงสถิติ ผลการประเมินบนชุดข้อมูลทดสอบพบว่า LSTM ให้ผลดีที่สุด โดยมีค่า MAE เท่ากับ 3,587.05 ค่า RMSE เท่ากับ 4,452.09 ค่า MAPE เท่ากับร้อยละ 5.94 และค่า R^2 เท่ากับ 0.3569 ขณะที่ Random Forest และ Gradient Boosting มีค่าคลาดเคลื่อนสูงและมีค่า R^2 ติดลบ สะท้อนข้อจำกัดในการติดตามแนวโน้มราคาที่เปลี่ยนแปลงต่อเนื่อง ผลการทดสอบ Diebold-Mariano แบบ pairwise ยืนยันว่า LSTM ให้ความแม่นยำเหนือกว่า Random Forest และ Gradient Boosting อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ในขณะที่ Random Forest และ Gradient Boosting ไม่แตกต่างกันอย่างมีนัยสำคัญ ผลการวิเคราะห์ความสำคัญของตัวแปรจากแบบจำลอง Random Forest ชี้ว่า ราคาต่ำสุด ราคาสูงสุด และราคาเปิด เป็นตัวแปรสำคัญต่อการพยากรณ์มากที่สุด ผลการศึกษานี้สะท้อนว่าความสอดคล้องระหว่างโครงสร้างของแบบจำลองกับธรรมชาติของข้อมูลอนุกรมเวลาเป็นปัจจัยสำคัญต่อคุณภาพการพยากรณ์ และชี้ให้เห็นศักยภาพของ LSTM สำหรับการพัฒนาเป็นเครื่องมือสนับสนุนการตัดสินใจด้านราคาทองรูปพรรณในระดับเบื้องต้น

คำสำคัญ: ราคาทองรูปพรรณ, ข้อมูลอนุกรมเวลา, Random Forest, Gradient Boosting, LSTM, การพยากรณ์ราคา

Abstract

This study evaluates and compares the performance of Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) for forecasting Thailand's gold jewelry selling price using daily time-series data from 2021 to 2026 under a unified experimental framework. Test-set results show that LSTM achieved the best performance with an MAE of 3,587.05, RMSE of 4,452.09, MAPE of 5.94%, and R^2 of 0.3569, while Random Forest and Gradient Boosting produced substantially larger errors and negative R^2 values. Pairwise Diebold-Mariano tests further confirmed that LSTM significantly outperformed both Random Forest and Gradient Boosting at the 0.05 level, whereas the difference between the two tree-based models was not statistically significant. Feature-importance analysis based on the Random Forest model also indicated that the daily low, high, and open prices were the most influential predictors. The findings highlight that alignment between model structure and the temporal nature of the data is a key determinant of forecasting quality and point to the potential of LSTM as a preliminary decision-support component for gold jewelry price analysis in Thailand.

Keywords: Gold jewelry price, Time-series forecasting, Random Forest, Gradient Boosting, LSTM

บทนำ

ทองรูปพรรณเป็นสินทรัพย์ที่มีความสำคัญทั้งในด้านการออม การลงทุน และการใช้ประโยชน์ในชีวิตประจำวันของประชาชนไทยราคาทองรูปพรรณในประเทศจึงมีนัยสำคัญ ทั้งเชิงเศรษฐกิจระดับครัวเรือนและเชิงพาณิชย์ อย่างไรก็ดี ราคาดังกล่าวมิได้เคลื่อนที่อย่างคงที่ หากแต่ผันผวนตามราคาทองคำในตลาดโลก อัตราแลกเปลี่ยนเงินบาทต่อดอลลาร์สหรัฐ ต้นทุนภายในประเทศ และสภาวะความไม่แน่นอนของเศรษฐกิจโลก ทำให้ผู้ซื้อ ผู้ค้า และผู้ลงทุนต้องตัดสินใจภายใต้ข้อมูลที่เปลี่ยนแปลงอย่างรวดเร็ว (สมาคมค้าทองคำ, 2568; World Gold Council, 2025)

แม้สมาคมค้าทองคำจะประกาศราคาทองรูปพรรณรายวันอย่างต่อเนื่อง แต่การพยากรณ์แนวโน้มระยะสั้นยังคงเป็นโจทย์ที่ท้าทาย เนื่องจากข้อมูลมีทั้งแนวโน้ม ความผันผวน การแกว่งตัวตามเหตุการณ์ และความสัมพันธ์ตามลำดับเวลา งานก่อนหน้าบางส่วนให้ความสนใจกับผลลัพธ์เชิงความแม่นยำ แต่ยังอธิบายไม่เพียงพอว่าวิธีที่ใช้สอดคล้องกับธรรมชาติของข้อมูลมากน้อยเพียงใด โดยเฉพาะกรณีที่นำแบบจำลองเชิงตารางมาใช้กับข้อมูลอนุกรมเวลา (Mosavi, Ozturk, & Chau, 2018; Sezer, Gudelek, & Ozbayoglu, 2020)

เมื่อพิจารณาข้อมูลราคาทองรูปพรรณขายออกในช่วงเวลาที่ศึกษา พบว่าราคามีช่วงการเคลื่อนไหวตั้งแต่ 25,000 ถึง 68,000 บาท โดยมีค่าเฉลี่ย 37,048.63 บาท และส่วนเบี่ยงเบนมาตรฐาน 10,127.56 บาท สะท้อนให้เห็นว่าตลาดมีทั้งแนวโน้มระยะยาวและความผันผวนในระดับที่มีนัยสำคัญ ซึ่งทำให้การพยากรณ์จากข้อมูลในอดีตเป็นโจทย์ที่ต้องอาศัยแบบจำลองที่สามารถรับมือกับ temporal dependency ได้อย่างเหมาะสม

ในเชิงวิชาการ ปัญหาสำคัญจึงไม่ได้อยู่ที่การตอบเพียงว่าแบบจำลองใดให้ผลดีกว่าเท่านั้น แต่รวมถึงการอธิบายด้วยว่าเหตุใดแบบจำลองนั้น จึงเหมาะสมหรือไม่เหมาะสมกับภารกิจการพยากรณ์ข้อมูลอนุกรมเวลา โดยเฉพาะเมื่อตลาดทองรูปพรรณมีทั้งแนวโน้มระยะสั้น การเกิดค่าราคาสูงใหม่ และแรงกระทบจากปัจจัยภายนอกที่เปลี่ยนแปลงอยู่เสมอ (Hooker & Mentch, 2019; Bontempi, Taieb, & Le Borgne, 2013)

เมื่อทบทวนวรรณกรรมที่เกี่ยวข้องพบว่า งานวิจัยส่วนใหญ่ยังมุ่งเน้นราคาทองคำสปอตในตลาดโลกมากกว่าราคาทองรูปพรรณในประเทศไทย ขณะเดียวกันหลายการศึกษามักใช้ชุดข้อมูล วิธีเตรียมข้อมูล หรือวิธีประเมินผลที่ต่างกัน ทำให้ยากต่อการเปรียบเทียบผลอย่างเป็นธรรม ช่องว่างดังกล่าวทำให้ยังจำเป็นต้องมีการประเมินแบบจำลองภายใต้กรอบการทดลองเดียวกัน ในบริบทของราคาทองรูปพรรณไทย (Demšar, 2006; Wang & Lin, 2024; Gong, 2024)

ดังนั้น การศึกษานี้จึงมุ่งประเมินสมรรถนะของ Random Forest, Gradient Boosting และ LSTM ภายใต้ข้อมูลและเงื่อนไขการทดลองเดียวกัน โดยมีเป้าหมายเพื่อวิเคราะห์ทั้งระดับความแม่นยำ ความเหมาะสมเชิงโครงสร้าง และข้อจำกัดของแต่ละแบบจำลอง มากกว่าการสรุปเชิงกว้างว่าแบบจำลองประเภทหนึ่งดีกว่าอีกประเภทหนึ่งเสมอไป

วัตถุประสงค์ของการวิจัย

1. เพื่อพัฒนาแบบจำลอง Random Forest, Gradient Boosting และ LSTM สำหรับการพยากรณ์ราคาทองรูปพรรณรายวันของประเทศไทย โดยใช้ข้อมูลราคาในอดีตและตัวแปรเชิงเทคนิค
2. เพื่อเปรียบเทียบสมรรถนะของแบบจำลองทั้งหมดภายใต้กรอบการทดลองเดียวกัน โดยควบคุมชุดข้อมูล วิธีการเตรียมข้อมูล วิธีแบ่งข้อมูลตามลำดับเวลา และตัวชี้วัดประสิทธิภาพ
3. เพื่ออภิปรายความเหมาะสม ข้อจำกัด และศักยภาพในการประยุกต์ใช้ของแบบจำลองแต่ละแบบสำหรับงานพยากรณ์ข้อมูลอนุกรมเวลาในบริบทราคาทองรูปพรรณ

ทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง

1. แบบจำลองอนุกรมเวลาเชิงสถิติแบบดั้งเดิม

งานพยากรณ์ข้อมูลอนุกรมเวลาในทางเศรษฐศาสตร์ และการเงินมักเริ่มต้นจากแบบจำลองเชิงสถิติ เช่น ARIMA ซึ่งพัฒนาภายใต้ Box-Jenkins methodology โดยอาศัยแนวคิด Autoregressive, Integrated และ Moving Average เพื่ออธิบายความสัมพันธ์ของข้อมูลย้อนหลัง แบบจำลองดังกล่าวมีประโยชน์ในฐานะกรอบอ้างอิงสำคัญ สำหรับการอธิบายโครงสร้างเชิงเส้นของข้อมูล แต่มีข้อจำกัดเมื่อข้อมูลมีความไม่เชิงเส้นหรือเกิดการเปลี่ยนแปลงเชิงโครงสร้างบ่อยครั้ง (Box, Jenkins, Reinsel, & Ljung, 2015)

2. แบบจำลองเชิงต้นไม้สำหรับการพยากรณ์ข้อมูล

Random Forest และ Gradient Boosting เป็นแบบจำลองที่ได้รับความนิยมในงานพยากรณ์และงานการวิเคราะห์การถดถอย เนื่องจากสามารถเรียนรู้ความสัมพันธ์ไม่เชิงเส้นและปฏิสัมพันธ์ของตัวแปรได้ โดยไม่ต้องกำหนดรูปแบบสมการล่วงหน้า Random Forest อาศัยแนวคิด bootstrap aggregation เพื่อลดความแปรปรวนของการพยากรณ์ ขณะที่ Gradient Boosting ใช้การเรียนรู้แบบลำดับขั้น เพื่อลดความผิดพลาดของแบบจำลองเดิมอย่างต่อเนื่อง อย่างไรก็ตามเมื่อประยุกต์ใช้กับข้อมูล

อนุกรมเวลาในรูปแบบตัวแปร lag แบบจำลองกลุ่มนี้ อาจเผชิญข้อจำกัดในการติดตามแนวโน้มที่เปลี่ยนแปลงต่อเนื่องหรือการคาดการณ์ค่าที่อยู่นอกช่วงข้อมูลฝึก (Breiman, 2001; Friedman, 2001; Hooker & Mentch, 2019)

3. แบบจำลอง LSTM สำหรับข้อมูลอนุกรมเวลา

LSTM เป็นโครงข่ายประสาทเทียมที่ได้รับการออกแบบมา เพื่อจัดการข้อมูลเชิงลำดับผ่านกลไกหน่วยความจำและประตูควบคุม ที่ช่วยให้แบบจำลองสามารถรักษาบริบทจากข้อมูลในอดีตได้ดีกว่าโครงข่ายประสาทเทียมแบบทั่วไป จึงได้รับความสนใจอย่างแพร่หลายในงานพยากรณ์ข้อมูลการเงินและข้อมูลราคาสินทรัพย์ (Hochreiter & Schmidhuber, 1997; Fischer & Krauss, 2018) งานเฉพาะด้านทองคำก็รายงานแนวโน้มสอดคล้องกัน เช่น Livieris, Pintelas, & Pintelas (2020) พบว่า โครงสร้าง CNN-LSTM ปรับตัวเข้ากับข้อมูลราคาทองคำได้ดี ขณะที่ Wang & Lin (2024) เสนอกรอบ deep learning เชิงกำหนดและเชิงความน่าจะเป็นสำหรับการพยากรณ์ราคาทอง และ Gong (2024) เปรียบเทียบ LSTM กับ linear regression ในบริบทการคาดการณ์การขึ้นลงของราคาทอง

4. ช่องว่างงานวิจัย

แม้งานพยากรณ์ราคาทองคำจะมีอยู่จำนวนมาก แต่การศึกษาในบริบทของราคาทองรูปพรรณของประเทศไทยยังมีจำกัด อีกทั้งงานวิจัยก่อนหน้าหลายชิ้นใช้ข้อมูล วิธีการเตรียมข้อมูลและการแบ่งชุดข้อมูลที่แตกต่างกัน ทำให้การเปรียบเทียบผลระหว่างแบบจำลองข้ามงานวิจัยเป็นไปได้ยาก งานวิจัยนี้จึงมุ่งเติมช่องว่างดังกล่าวโดยประเมิน Random Forest, Gradient Boosting และ LSTM ภายใต้กรอบการทดลองเดียวกัน เพื่อพิจารณาว่าแบบจำลองใดมีความเหมาะสมกับข้อมูลราคาทองรูปพรรณรายวันของประเทศไทยมากที่สุด

ตารางที่ 1 สรุปงานวิจัยที่เกี่ยวข้องและตำแหน่งของงานวิจัยนี้

ผู้วิจัย/ปี	บริบทข้อมูล	วิธีที่ศึกษา	สาระสำคัญ	จุดต่างจากงานนี้
Fischer & Krauss (2018)	ข้อมูลตลาดการเงิน	LSTM	LSTM มีศักยภาพต่อข้อมูลลำดับเวลา	ไม่ได้ศึกษาราคาทองรูปพรรณไทย
Livieris, Pintelas, & Pintelas (2020)	ราคาทองคำ	LSTM, CNN-LSTM	แบบจำลองลึกปรับตัวต่อข้อมูลทองได้ดี	ไม่ได้เปรียบเทียบ RF/GB ภายใต้กรอบเดียวกัน
Wang & Lin (2024)	ราคาทองคำและตัวแปรภายนอก	Quantile Regression deep learning	เน้นกรอบพยากรณ์เชิงกำหนด/เชิงความน่าจะเป็น	บริบทข้อมูลและโจทย์ต่างจากราคาทองรูปพรรณไทย
Gong (2024)	ราคาทองคำรายวัน	LSTM, Linear Regression	เปรียบเทียบ LSTM กับโมเดลง่าย	ไม่ได้ประเมิน tree-based models
งานวิจัยนี้	ราคาทองรูปพรรณขายออกของไทย	Random Forest, Gradient Boosting, LSTM	เปรียบเทียบโมเดลต่างโครงสร้างภายใต้กรอบเดียวกัน พร้อม DM test	มุ่งตอบความเหมาะสมเชิงโครงสร้างกับข้อมูลไทย

แนวคิด ทฤษฎี กรอบแนวคิด

ราคาทองรูปพรรณในประเทศไทยเป็นผลลัพธ์ของกลไกตลาดที่เชื่อมโยง ทั้งในระดับจุลภาคและมหภาค โดยมีความสัมพันธ์กับราคาทองคำสปอตในตลาดโลก อัตราแลกเปลี่ยนเงินบาทต่อดอลลาร์สหรัฐ และสภาวะความไม่แน่นอนทางเศรษฐกิจระหว่างประเทศ (World Gold Council, 2025; สมาคมค้าทองคำ, 2568) ความสัมพันธ์ดังกล่าวไม่ได้มีลักษณะเชิงเส้นอย่างสมบูรณ์ เนื่องจากยังได้รับผลกระทบจากแรงเก็งกำไร ความคาดหวังของตลาดและเหตุการณ์ภูมิรัฐศาสตร์

ทฤษฎีการเรียนรู้ของเครื่องตั้งอยู่บนหลักการประมาณฟังก์ชันเชื่อมโยงตัวแปรอิสระกับตัวแปรตามผ่านกระบวนการลดค่าความสูญเสียเชิงประจักษ์ ความท้าทายสำคัญอยู่ที่การจัดสมดุลระหว่างอคติและความแปรปรวน (Geman, Bienenstock, & Doursat, 1992; Hastie, Tibshirani, & Friedman, 2009). สำหรับข้อมูลอนุกรมเวลา ความสอดคล้องระหว่างโครงสร้างแบบจำลองกับ temporal dependency มีผลต่อสมรรถนะของการพยากรณ์ อย่างมีนัยสำคัญ

กรอบแนวคิดของการศึกษานี้ ตั้งอยู่บนข้อพิจารณาว่าราคาทองรูปพรรณเป็นผลลัพธ์ของความสัมพันธ์เชิงซับซ้อนระหว่างพฤติกรรมราคาในอดีต และบริบทเศรษฐกิจที่ส่งผลต่อการเคลื่อนไหวของราคาในแต่ละช่วงเวลาภายใต้เงื่อนไขดังกล่าว แบบจำลองที่มีโครงสร้างต่างกันย่อมมีความสามารถแตกต่างกันในการเรียนรู้ความสัมพันธ์เชิงเวลาและการตอบสนองต่อแนวโน้มของข้อมูล

ดังนั้น การตัดสินความเหมาะสมของแบบจำลอง จึงควรอาศัยการเปรียบเทียบเชิงประจักษ์ภายใต้กรอบการทดลองที่ควบคุมเงื่อนไขอย่างเป็นระบบ เพื่อให้ได้ข้อสรุปที่สมดุลทั้งด้านความแม่นยำ การอธิบายข้อจำกัด และศักยภาพในการประยุกต์ใช้จริง (Demšar, 2006; Bontempi, Taieb, & Le Borgne, 2013)

วิธีดำเนินการวิจัย

1. ระเบียบวิธีวิจัยและกรอบการทดลองเชิงคำนวณ งานวิจัยนี้ใช้ระเบียบวิธีวิจัยเชิงปริมาณผ่านการทดลองเชิงคำนวณ เพื่อประเมินและเปรียบเทียบประสิทธิภาพของแบบจำลองพยากรณ์ราคาทองรูปพรรณรายวัน การออกแบบการทดลองยึดหลัก comparative experimental design ซึ่งมุ่งควบคุมเงื่อนไขของการทดลองให้เหมือนกันในทุกแบบจำลอง เพื่อให้ความแตกต่างของผลลัพธ์สะท้อนศักยภาพของอัลกอริทึมโดยแท้จริง มิใช่ผลจากความแตกต่างของกระบวนการเตรียมข้อมูล โครงสร้างคุณลักษณะ หรือขั้นตอนการประเมินผล

2. แหล่งข้อมูลและลักษณะข้อมูล ข้อมูลที่ใช้ในการศึกษาเป็นข้อมูลทุติยภูมิจากสมาคมค้าทองคำแห่งประเทศไทย ครอบคลุมราคาทองรูปพรรณขายออกรายวัน ในช่วงวันที่ 1 มกราคม พ.ศ. 2564 ถึง 1 มกราคม พ.ศ. 2569 รวมทั้งสิ้น 1,826 ระเบียนข้อมูล จากการสำรวจข้อมูลเบื้องต้น พบว่า ราคาขายออกมีค่าต่ำสุด 25,000 บาท ค่าสูงสุด 68,000 บาท ค่าเฉลี่ย 37,048.63 บาท และส่วนเบี่ยงเบนมาตรฐาน 10,127.56 บาท ข้อมูลมีลักษณะเป็นอนุกรมเวลา และตรวจพบค่าที่ขาดหาย คิดเป็นร้อยละ 7.2 ของข้อมูลทั้งหมด

3. การเตรียมข้อมูลและการสร้างชุดข้อมูล ค่าที่ขาดหายได้รับการจัดการด้วยวิธี linear interpolation และ forward filling ตามลักษณะของข้อมูลเชิงเวลา ขณะที่ค่าผิดปกติได้รับการตรวจสอบด้วยเกณฑ์ Z-score และ Interquartile Range เพื่อป้องกันอิทธิพลของ outliers ต่อกระบวนการเรียนรู้ ตัวแปรเชิงต่อเนื่องถูกปรับขนาดด้วยวิธี Min-Max normalization โดย fit scaler เฉพาะบนชุดฝึกเท่านั้น จากนั้นจึงใช้ transformation เดียวกันกับชุดตรวจสอบและชุดทดสอบ ก่อนการ

คำนวณตัวชี้วัดประสิทธิภาพได้ทำ inverse transformation กับค่าพยากรณ์เพื่อให้ตัวชี้วัดสะท้อนหน่วยราคาจริง

4. การสร้างคุณลักษณะและพารามิเตอร์ของแบบจำลอง งานวิจัยนี้สร้างตัวแปรนำเข้า 15 ตัวจากข้อมูลราคาในอดีต โดยครอบคลุมตัวแปรเชิง lag ตัวแปรแนวโน้ม และตัวแปรความผันผวน พร้อมใช้ข้อมูลย้อนหลัง 14 วันเป็นอินพุตสำหรับ one-step-ahead forecasting. Random Forest กำหนด n_estimators = 200, max_depth = 10, min_samples_split = 5, random_state = 42 และ n_jobs = -1. Gradient Boosting กำหนด n_estimators = 100, learning_rate = 0.1 และ random_state = 42

แบบจำลอง LSTM ใช้โครงสร้างสองชั้น โดยกำหนด LSTM 64 หน่วยในชั้นแรก (return_sequences =True) ตามด้วย Dropout 0.2, LSTM 32 หน่วยในชั้นที่สอง, Dropout 0.2, Dense 16 หน่วยพร้อม activation แบบ ReLU และ Dense 1 หน่วยในชั้นเอาต์พุต ใช้ Adam optimizer, loss แบบ MSE, batch size = 32, epochs สูงสุด = 100 และ Early Stopping ด้วย patience = 10 เพื่อเลือกแบบจำลองที่ให้ผลดีที่สุดจากชุดตรวจสอบ

5. การแบ่งชุดข้อมูล งานวิจัยนี้ใช้การแบ่งข้อมูลตามลำดับเวลาแบบ holdout split เพื่อสะท้อนสถานการณ์การพยากรณ์จริงและป้องกันข้อมูลรั่วไหล โดยกำหนดชุดฝึก ชุดตรวจสอบ และชุดทดสอบตามลำดับเวลาอย่างชัดเจน ได้แก่ ชุดฝึก (70%) ครอบคลุมปี พ.ศ. 2564-2566 ชุดตรวจสอบ (15%) ครอบคลุมปี พ.ศ. 2567 และชุดทดสอบ (15%) ครอบคลุมปี พ.ศ. 2568-2569 ไม่มีการนำข้อมูลจากชุดทดสอบไปใช้ในการปรับแต่งพารามิเตอร์ของแบบจำลอง

6. การประเมินประสิทธิภาพและกรอบการเปรียบเทียบเชิงสถิติ งานวิจัยนี้ใช้ตัวชี้วัด MAE, RMSE, MAPE และ R^2 ในการประเมินสมรรถนะของแบบจำลอง และใช้การทดสอบ Diebold-Mariano เพื่อเปรียบเทียบความแม่นยำเป็นรายคู่ โดยพิจารณาความแตกต่างของ squared error ในแต่ละช่วงเวลา สถิติทดสอบคำนวณจากค่าเฉลี่ยของ loss differential และปรับส่วนเบี่ยงเบนมาตรฐานด้วย Newey-West estimator เพื่อรองรับ Autocorrelation และ heteroskedasticity การทดสอบใช้สมมติฐานว่าง (Null Hypothesis) ว่าความแม่นยำของแบบจำลองทั้งสองไม่แตกต่างกัน และใช้การทดสอบแบบสองด้านที่ระดับนัยสำคัญ 0.05

ตารางที่ 2 รายการตัวแปรนำเข้าที่ใช้ในการพัฒนาแบบจำลอง

กลุ่มตัวแปร	ตัวอย่างตัวแปร	ความหมาย	ใช้กับแบบจำลอง
ราคาในวันเดียวกัน	open, high, low	ข้อมูลราคาเปิด สูงสุด และต่ำสุดในวันปัจจุบัน	ทุกแบบจำลอง
ตัวแปรหน่วงเวลา	sell_lag1 - sell_lag7	ราคาขายออกย้อนหลัง 1-7 วัน	ทุกแบบจำลอง
ตัวแปรแนวโน้ม	sell_ma5, sell_ma20	ค่าเฉลี่ยเคลื่อนที่เพื่อสะท้อนแนวโน้มระยะสั้นและกลาง	ทุกแบบจำลอง
ตัวแปรความผันผวน	roll_std, daily_return	ส่วนเบี่ยงเบนมาตรฐานเคลื่อนที่และอัตราการเปลี่ยนแปลงรายวัน	ทุกแบบจำลอง

ผลการวิจัย

การศึกษานี้มีวัตถุประสงค์เพื่อประเมินสมรรถนะของแบบจำลอง Random Forest, Gradient Boosting และ LSTM ในการพยากรณ์ราคาทองรูปพรรณขายออก โดยกำหนดเงื่อนไขการทดลองให้เหมือนกันทุกแบบจำลอง ทั้งในด้านชุดข้อมูล การเตรียมข้อมูล วิธีแบ่งข้อมูลตามลำดับเวลา และตัวชี้วัดประสิทธิภาพ เพื่อให้ผลการเปรียบเทียบสะท้อนความแตกต่างของแบบจำลองอย่างเป็นธรรม

1. ผลการเปรียบเทียบประสิทธิภาพของแบบจำลอง

จากตารางที่ 3 จะเห็นได้ว่าการประเมินบนชุดข้อมูลทดสอบ พบว่าแบบจำลอง LSTM ให้ผลดีที่สุดเมื่อพิจารณาจากตัวชี้วัดหลักทุกตัวขณะที่ Random Forest และ Gradient Boosting มีค่าความคลาดเคลื่อนสูงใกล้เคียงกันและมีค่า R^2 ติดลบ แสดงให้เห็นว่าแบบจำลองทั้งสองไม่สามารถติดตามโครงสร้างของข้อมูลในช่วงทดสอบได้ดี

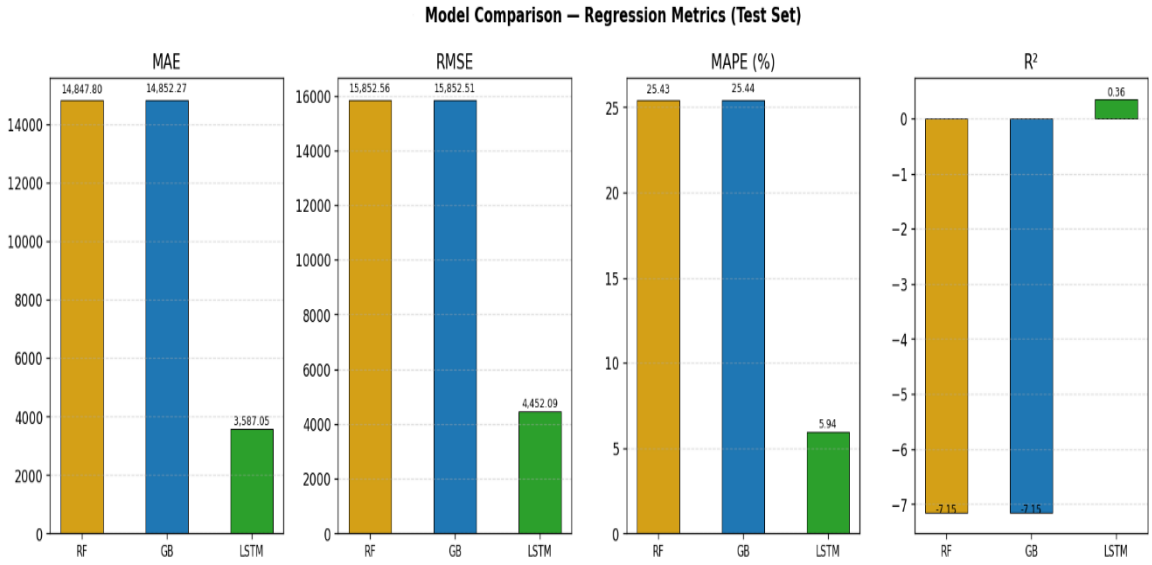
ตารางที่ 3 ตารางเปรียบเทียบประสิทธิภาพของแบบจำลอง

Model	MAE	RMSE	MAPE (%)	R^2
Random Forest Regressor	14,847.80	15,852.56	25.4331	-7.1539
Gradient Boosting Regressor	14,852.27	15,852.51	25.4431	-7.1539
LSTM	3,587.05	4,452.09	5.9395	0.3569

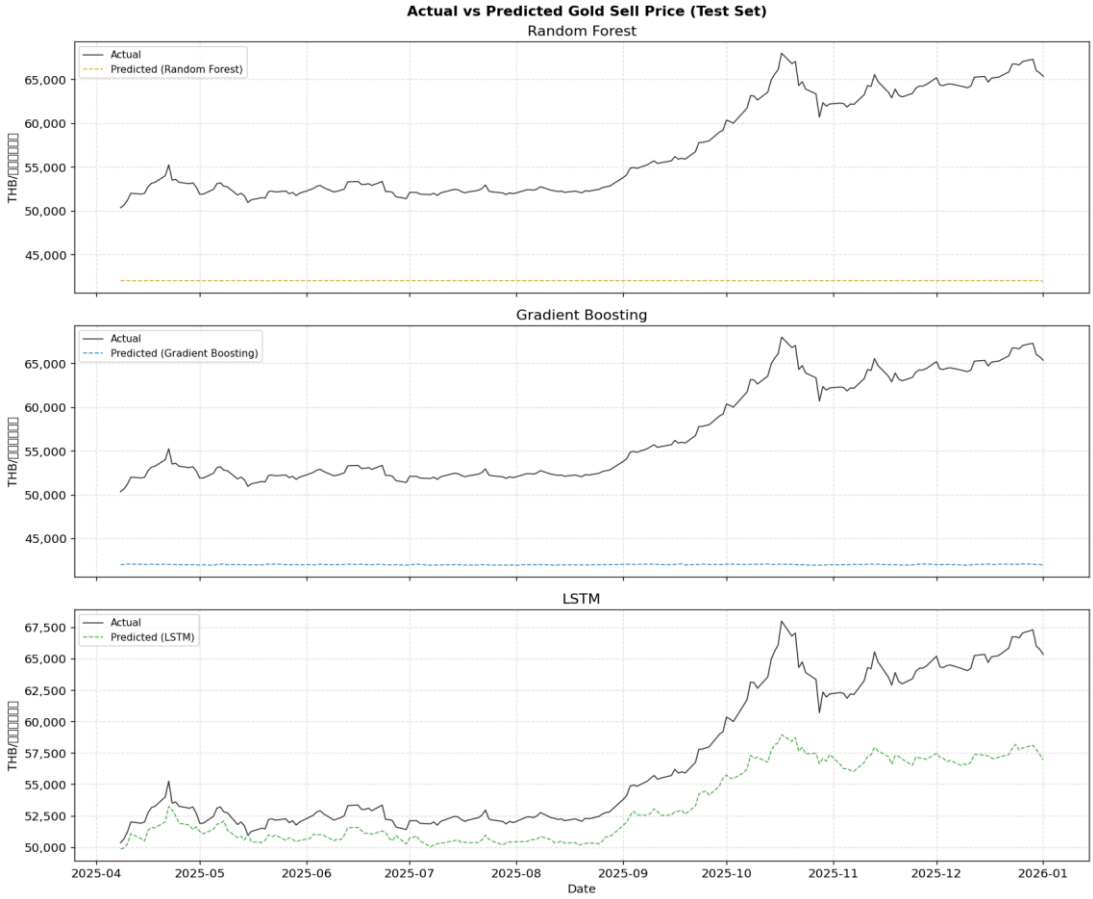
ตารางที่ 4 ผลการทดสอบ Diebold-Mariano แบบ pairwise

คู่แบบจำลอง	**DM statistic**	**p-value**	**ข้อสรุป**
Random Forest vs Gradient Boosting	0.0176	0.9860	ไม่แตกต่างกันอย่างมีนัยสำคัญ
Random Forest vs LSTM	20.0083	< 0.001	LSTM แม่นยำกว่า อย่างมีนัยสำคัญ
Gradient Boosting vs LSTM	20.0537	< 0.001	LSTM แม่นยำกว่า อย่างมีนัยสำคัญ

จากตารางที่ 4 จะเห็นได้ว่า LSTM มีความแม่นยำเหนือกว่า Random Forest และ Gradient Boosting อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 ขณะที่ Random Forest และ Gradient Boosting ไม่แตกต่างกันอย่างมีนัยสำคัญ ซึ่งสอดคล้องกับผลจากตัวชี้วัด MAE, RMSE และ MAPE



ภาพที่ 1 การเปรียบเทียบค่าชี้วัดบนชุดทดสอบ

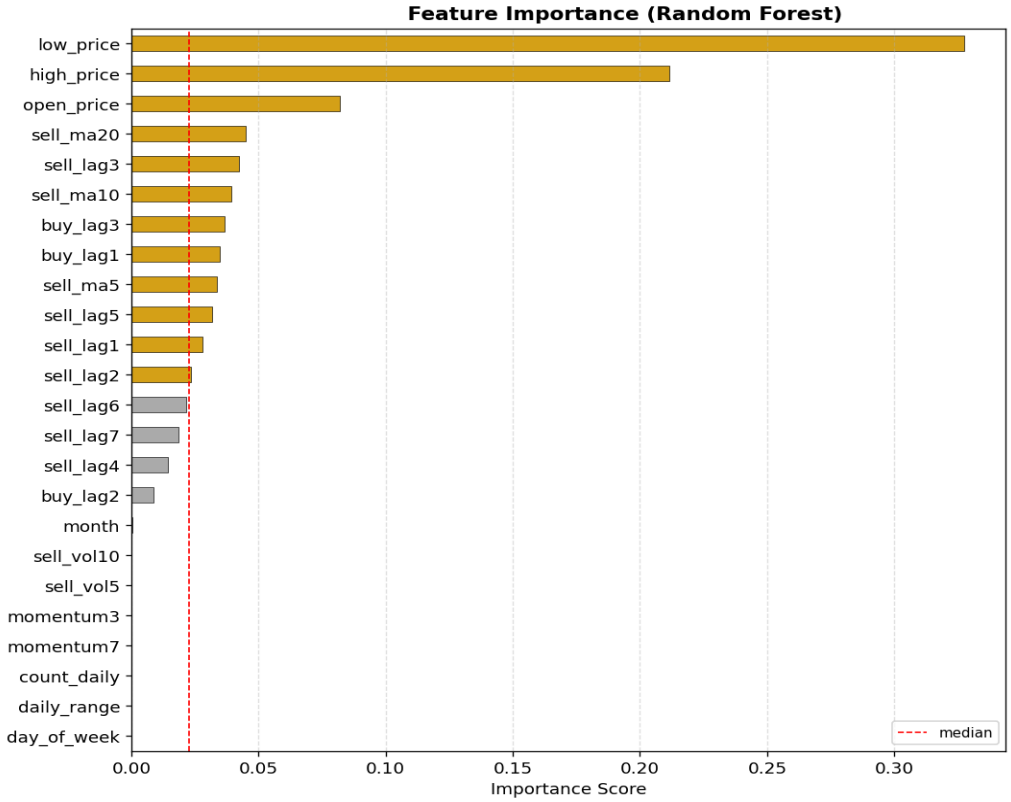


ภาพที่ 2 Actual vs Predicted Gold Sell Price

เมื่อพิจารณาแนวโน้มของค่าพยากรณ์ดังแสดงในภาพที่ 2 พบว่า LSTM สามารถเกาะติดทิศทางราคาจริงได้ใกล้เคียงกว่า ขณะที่แบบจำลอง Tree-based ให้ค่าพยากรณ์ค่อนข้างแบนและตอบสนองต่อการเปลี่ยนแปลงเชิงแนวโน้มได้จำกัด โดยเฉพาะในช่วงที่ข้อมูลจริงมีการปรับตัวขึ้นต่อเนื่อง

2. ผลการวิเคราะห์ความสำคัญของตัวแปร

การวิเคราะห์ความสำคัญของตัวแปรในส่วนนี้อ้างอิงจากแบบจำลอง Random Forest ซึ่งสามารถคำนวณค่า feature importance ได้โดยตรง ผลลัพธ์ชี้ให้เห็นว่าตัวแปรเชิงราคาภายในวันยังคงเป็นแกนหลักของการพยากรณ์



ภาพที่ 3 Feature Importance (Random Forest)

ผลการวิเคราะห์ความสำคัญของตัวแปร พบว่า ราคาต่ำสุด (Low Price) มีค่าความสำคัญ 0.3275 ราคาสูงสุด (High Price) มีค่าความสำคัญ 0.2115 และราคาเปิด (Open Price) มีค่าความสำคัญ 0.0818 ซึ่งเป็นสามอันดับแรกของตัวแปรทั้งหมด นอกจากนี้ ตัวแปรย้อนหลังและตัวแปรเชิงแนวโน้ม เช่น sell_ma20 และ sell_lag3 ก็มีส่วนช่วยในการพยากรณ์เช่นกัน

สรุปผลการวิจัย

การศึกษานี้ประเมินแบบจำลอง Random Forest, Gradient Boosting และ LSTM สำหรับการพยากรณ์ราคาทองรูปพรรณภายใต้กรอบการทดลองเดียวกัน ผลการประเมินแสดงให้เห็นว่า LSTM ให้ผลดีที่สุดทั้งในเชิงค่าความคลาดเคลื่อนและผลการทดสอบ Diebold-Mariano ขณะที่ Random Forest และ Gradient Boosting ให้ผลใกล้เคียงกันและไม่แตกต่างกัน อย่างมีนัยสำคัญ ผลดังกล่าวสะท้อนว่าความสอดคล้องระหว่างโครงสร้างของแบบจำลองกับธรรมชาติ ของข้อมูลอนุกรมเวลาเป็นปัจจัยสำคัญต่อคุณภาพการพยากรณ์ อย่างไรก็ตาม ข้อค้นพบนี้ควรถูกตีความภายใต้ขอบเขตของข้อมูล ตัวแปร และช่วงเวลาที่ใช้ศึกษาในงานวิจัยนี้

อภิปรายผล

1. ความเหมาะสมของ LSTM ต่อข้อมูลอนุกรมเวลา

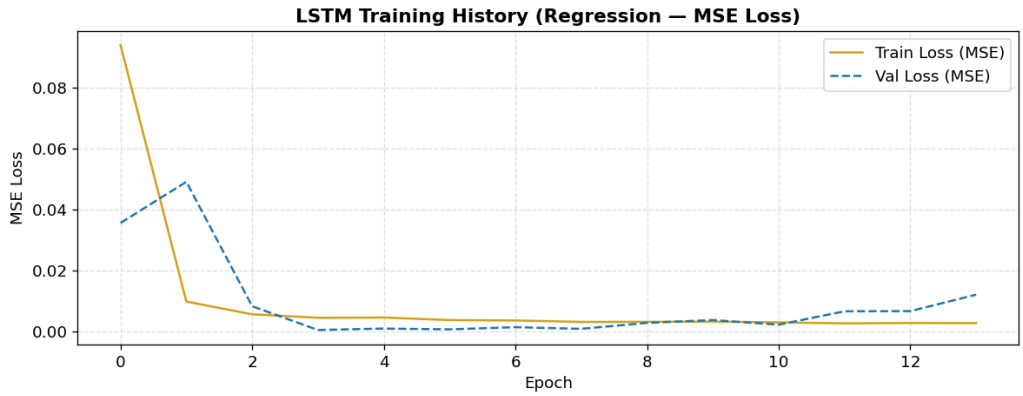
การที่ LSTM ให้ผลดีกว่า Random Forest และ Gradient Boosting สอดคล้องกับคุณสมบัติของสถาปัตยกรรมที่ออกแบบมา เพื่อเรียนรู้ความสัมพันธ์ตามลำดับเวลาและรักษาบริบทจากข้อมูลในอดีตผ่านกลไกหน่วยความจำ การทดสอบ Diebold-Mariano ที่พบความแตกต่างอย่างมีนัยสำคัญระหว่าง LSTM กับแบบจำลองอีกสองแบบช่วยยืนยันว่า ความเหนือกว่านี้ไม่ได้เกิดจากความแตกต่างของค่าความคลาดเคลื่อนโดยบังเอิญเพียงอย่างเดียว (Hochreiter & Schmidhuber, 1997; Fischer & Krauss, 2018) อย่างไรก็ตาม การที่ LSTM ให้ค่า R^2 เท่ากับ 0.3569 สะท้อนว่าความสามารถในการพยากรณ์ยังอยู่ในระดับปานกลางแม้จะดีกว่าแบบจำลองอื่นอย่างชัดเจน แต่ยังไม่ควรตีความว่าเพียงพอสำหรับการใช้งานเชิงตัดสินใจโดยลำพัง

2. ข้อจำกัดของแบบจำลอง Tree-based

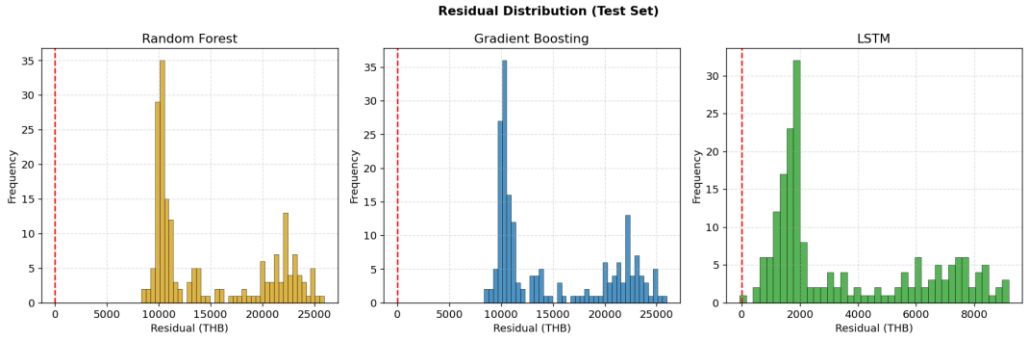
ผลการทดลองที่ Random Forest และ Gradient Boosting มีค่า R^2 ติดลบและให้แนวเส้นพยากรณ์ค่อนข้างราบ สะท้อนข้อจำกัดของแบบจำลองตระกูลต้นไม้ เมื่อถูกนำมาใช้กับข้อมูลอนุกรมเวลาในรูปแบบตัวแปร lag ภายใต้กรอบการทดลองนี้ แบบจำลองทั้งสองสามารถเรียนรู้รูปแบบจากข้อมูลฝึกได้ในระดับหนึ่ง แต่มีข้อจำกัดในการติดตามแนวโน้มของข้อมูลที่เปลี่ยนแปลงต่อเนื่องและการคาดการณ์ค่าที่เคลื่อนออกนอกช่วงข้อมูลเดิม (Hooker & Mentch, 2019) ทั้งนี้ ผลดังกล่าวมิได้หมายความว่า Random Forest และ Gradient Boosting ไม่สามารถใช้กับปัญหาอนุกรมเวลาได้ในทุกกรณี หากแต่ชี้ให้เห็นว่าความเหมาะสมของแบบจำลองขึ้นอยู่กับธรรมชาติของข้อมูลและวิธีการออกแบบตัวแปรนำเข้าร่วมด้วย

3. นัยเชิงข้อมูลและข้อจำกัดของงานวิจัย

การที่ราคาสูงสุด ราคาต่ำสุด และราคาเปิดมีความสำคัญสูงสุด สะท้อนว่ากรอบการเคลื่อนไหวของราคาในแต่ละวันมีความสัมพันธ์ต่อราคาขายออกวันถัดไปอย่างมีนัยสำคัญ อย่างไรก็ตาม งานวิจัยนี้ยังใช้ตัวแปรจากข้อมูลราคาเป็นหลัก จึงยังไม่ครอบคลุมอิทธิพลของตัวแปรเศรษฐศาสตร์มหภาค เช่น อัตราแลกเปลี่ยน ราคาทองคำตลาดโลก หรืออัตราดอกเบี้ย ซึ่งอาจช่วยยกระดับความแม่นยำได้ในอนาคต ในเชิงปฏิบัติ ผลการศึกษานี้ จึงควรถูกตีความว่าเป็นหลักฐานเชิงเปรียบเทียบของแบบจำลองภายใต้ชุดข้อมูลและเงื่อนไขที่กำหนด มากกว่าจะเป็นข้อยืนยันว่าแบบจำลองใดสามารถใช้แทนการวิเคราะห์ปัจจัยด้านอื่นได้ทั้งหมด



ภาพที่ 4 LSTM Training History



ภาพที่ 5 Residual Distribution

ข้อเสนอแนะ

1. ด้านการนำไปใช้ประโยชน์ ควรนำแบบจำลองที่ให้ผลดีที่สุดไปพัฒนาเป็นเครื่องมือสนับสนุนการตัดสินใจเบื้องต้น โดยออกแบบให้นำเสนอค่าพยากรณ์ร่วมกับช่วงความไม่แน่นอนและข้อมูลประกอบอื่นเพื่อให้ผู้ใช้งานสามารถประเมินความเสี่ยงได้รอบด้านมากขึ้น

2. ด้านการวิจัยในอนาคต ควรเพิ่มตัวแปรเศรษฐศาสตร์มหภาค เช่น อัตราแลกเปลี่ยนเงินบาท ราคาทองคำในตลาดโลก อัตราดอกเบี้ย และดัชนีเศรษฐกิจที่เกี่ยวข้อง ตลอดจนทดลองเปรียบเทียบกับแบบจำลองที่ออกแบบมาสำหรับข้อมูลอนุกรมเวลาโดยตรงเพิ่มเติม เช่น SARIMA, SARIMAX, Prophet และ Transformer พร้อมทั้งทำ robustness check ภายใต้หลายช่วงเวลาและหลายรูปแบบการแบ่งข้อมูล

เอกสารอ้างอิง

สมาคมค้าทองคำ. (2568). ข้อมูลราคาทองคำย้อนหลัง. ดึงข้อมูลเมื่อ 15 กุมภาพันธ์ 2569, จาก https://www.goldtraders.or.th

Bontempi, G., Taieb, S. B., & Le Borgne, Y.-A. (2013). Machine learning strategies for time series forecasting. Lecture Notes in Business Information Processing, 138, 62-77.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research, 7, 1-30.

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189-1232.

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. Neural Computation, 4(1), 1-58.

Gong, W. (2024). Research on gold price forecasting based on LSTM and linear regression. SHS Web of Conferences, 181, 02005. https://doi.org/10.1051/shsconf/202418102005

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Springer.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780.

Hooker, G., & Mentch, L. (2019). Please stop permuting features: An explanation and alternatives. arXiv preprint arXiv:1905.03151.

Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting, 37(4), 1748-1764.

Livieris, I. E., Pintelas, E., & Pintelas, P. (2020). A CNN-LSTM model for gold price time-series forecasting. Neural Computing and Applications, 32(23), 17351-17360.

Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. Water, 10(11), 1536. https://doi.org/10.3390/w10111536

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005-2019. Applied Soft Computing, 90, 106181.

Taylor, S. J., & Letham, B. (2018). Forecasting at scale. The American Statistician, 72(1), 37-45.

Wang, Y., & Lin, T. (2024). A novel deterministic probabilistic forecasting framework for gold price with a new pandemic index based on quantile regression deep learning and multi-objective optimization. Mathematics, 12(1), 29. https://doi.org/10.3390/math12010029

World Gold Council. (2025). Gold demand trends. Retrieved from https://www.gold.org