

แบบจำลองทำนายราคาข้าวโพดโดยใช้การเรียนรู้เชิงลึก Corn Price Prediction Model Using Deep Learning

สิรินทรา สามชัยรัตน์¹ พัทธล ศรีอินทร์¹ สุกฤษฎ์ ปัญญาสิทธิ์¹ ปกรณ์ แวส่วงวงศ์² และ กฤตคม ศรีจิรานนท์^{1*}
Sirintra Samchairat¹ Phattadon Sri-in¹ Sukit Panyasit¹ Pakorn Waewsawangwong²
and Krittakom Srijiranon^{1*}

¹หน่วยวิจัยด้านนวัตกรรมข้อมูลและปัญญาประดิษฐ์แห่งมหาวิทยาลัยธรรมศาสตร์ สาขาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ จ.ปทุมธานี ประเทศไทย

²สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ จ.ปทุมธานี ประเทศไทย
¹Thammasat University Research Unit in Data Innovation and Artificial Intelligence, Department of
Computer Science, Faculty of Science and Technology, Thammasat University, Pathum Thani, Thailand

²Department of Computer Science, Faculty of Science and Technology, Thammasat University,
Pathum Thani, Thailand

วันที่ส่งบทความ : 30 มีนาคม 2567 วันที่แก้ไขบทความ : 28 พฤษภาคม 2568 วันที่ตอบรับบทความ : 30 พฤษภาคม 2568
Received: 30 March 2024, Revised: 28 May 2025, Accepted: 30 May 2025

บทคัดย่อ

ข้าวโพดเลี้ยงสัตว์เป็นสินค้าการเกษตรที่มีความต้องการสูงในหลายอุตสาหกรรม อย่างไรก็ตาม ราคาของข้าวโพดเลี้ยงสัตว์มีการเปลี่ยนแปลงอย่างไม่แน่นอนในแต่ละปี ส่งผลให้เกษตรกรประสบความยากลำบากในการวางแผนเพาะปลูก การคาดการณ์ราคาข้าวโพดล่วงหน้าจึงมีความสำคัญเพื่อให้เกษตรกรสามารถวางแผนได้อย่างมีประสิทธิภาพ งานวิจัยนี้นำเสนอการพัฒนาแบบจำลองทำนายราคาข้าวโพด โดยใช้ทั้งข้อมูลราคาข้าวโพดในอดีตและคุณลักษณะสำคัญอื่น ๆ งานวิจัยได้เปรียบเทียบประสิทธิภาพของแบบจำลอง 4 ประเภท ได้แก่ ARIMA, ARIMAX, LSTM และ GRU โดยใช้ข้อมูลระหว่างปี พ.ศ. 2558 ถึง พ.ศ. 2564 เพื่อทำนายราคาข้าวโพดในปี พ.ศ. 2563 และ พ.ศ. 2564 ผลการเปรียบเทียบระหว่างแบบจำลองที่ใช้ข้อมูลราคาข้าวโพดเพียงอย่างเดียวในการสร้าง และแบบจำลองที่ใช้ข้อมูลราคาข้าวโพดร่วมกับคุณลักษณะอื่นพบว่า แบบจำลอง GRU ที่ใช้ข้อมูลราคาข้าวโพดย้อนหลังร่วมกับคุณลักษณะอื่นได้แก่ ปริมาณส่งออกข้าวโพดทั้งหมด ราคาข้าวโพดส่งออก ราคาซื้อขายล่วงหน้าตลาดชิคาโก และมูลค่าส่งออกข้าวโพดทั้งหมด มีประสิทธิภาพดีที่สุด โดยมีค่า RMSE เท่ากับ 0.0780 และค่า MAE เท่ากับ 0.0662 การคาดการณ์ราคาที่แม่นยำมีความสำคัญอย่างยิ่งสำหรับเกษตรกรและผู้ที่เกี่ยวข้องในอุตสาหกรรมข้าวโพดเลี้ยงสัตว์ เนื่องจากช่วยให้การวางแผนและการบริหารจัดการทรัพยากรสามารถทำได้มีประสิทธิภาพมากขึ้น

คำสำคัญ : ราคาข้าวโพด จีอาร์ยู การเรียนรู้เชิงลึก แบบจำลองการทำนาย

*ที่อยู่ติดต่อ E-mail address: krittakom_s@sci.tu.ac.th

<https://doi.org/10.55003/scikmitl.2025.262715>

Abstract

Animal feed corn is a highly demanded agricultural commodity across various industries. However, its prices fluctuate unpredictably each year, causing difficulties for farmers in planning their crops. Therefore, accurate price forecasting is crucial for helping farmers plan effectively. This research presents the development of a corn price prediction model using both historical price data and other important features. The study compares the performance of four models: ARIMA, ARIMAX, LSTM, and GRU, using data from 2015 to 2021 to predict corn prices for 2020 and 2021. The comparison between models that use only historical price data and those that incorporate additional features shows that the GRU model, utilizing both historical price data and features such as total corn exports, export prices, Chicago Board of Trade futures prices, and total export value, performs the best. The GRU model achieves an RMSE of 0.0780 and an MAE of 0.0662, demonstrating the highest accuracy in the test datasets. Accurate price forecasting is crucial for farmers and stakeholders in the corn industry, as it enables more efficient planning and resource management.

Keywords: Corn price, GRU, Deep learning, Prediction model

1. บทนำ

ข้าวโพดเลี้ยงสัตว์เป็นพืชเศรษฐกิจที่สามารถนำไปแปรรูปเป็นอาหารมนุษย์และสัตว์ได้ ทั้งนี้ทั่วโลกมีความต้องการใช้ข้าวโพดเลี้ยงสัตว์จำนวนหลายล้านตัน ในปี พ.ศ. 2561 ประเทศไทยมีการนำเข้าข้าวโพดเลี้ยงสัตว์ปริมาณรวม 154,000 ตัน มูลค่ารวม 900.93 ล้านบาท และมีการส่งออกปริมาณรวม 83,000 ตัน มูลค่ารวมกว่า 685 ล้านบาท ซึ่งแสดงให้เห็นว่าประเทศไทยมีปริมาณการนำเข้าข้าวโพดเลี้ยงสัตว์มากกว่าการส่งออก เนื่องจากอุตสาหกรรมต่าง ๆ มีความต้องการใช้ข้าวโพดเลี้ยงสัตว์เป็นวัตถุดิบหลักในการแปรรูป โดยเฉพาะการแปรรูปเป็นอาหารสัตว์ซึ่งข้าวโพดเลี้ยงสัตว์นั้นมีความต้องการใช้มากขึ้นเรื่อย ๆ จนถึงปัจจุบัน สำนักงานเศรษฐกิจการเกษตรรายงานว่าในช่วงปี พ.ศ. 2560 - 2561 พื้นที่ปลูกมากที่สุดอยู่ในภาคเหนือร้อยละ 67.70 รองลงมาได้แก่ ภาคตะวันออกเฉียงเหนือร้อยละ 21.46 และในภาคกลางและภาคตะวันตกร้อยละ 10.87 (Office of Agricultural Research and Development Region 5, Department of Agriculture, Ministry of Agriculture and Cooperatives, 2019)

ข้าวโพดเลี้ยงสัตว์เป็นพืชที่มีตลาดรองรับ เพราะผลผลิตยังไม่เพียงพอต่อความต้องการใช้ภายในประเทศ (Office of Agricultural Economics Region 1, Chiang Mai, 2023) ปัญหาของเกษตรกรที่พบเจอในการปลูกข้าวโพดเลี้ยงสัตว์คือ ก่อนทำการเกษตร เกษตรกรไม่ทราบแนวโน้มราคาขายของข้าวโพดเลี้ยงสัตว์ที่จะเปลี่ยนแปลงในอนาคต ทำให้การวางแผนในการทำการเกษตรไม่มีประสิทธิภาพ การทำนายแนวโน้มการเปลี่ยนแปลงราคาของข้าวโพดเลี้ยงสัตว์จะช่วยให้เกษตรกรได้วางแผนในการทำ

การเกษตรที่เหมาะสม โดยการทำนายดังกล่าวจำเป็นต้องมีการนำวิทยาการข้อมูล (Data science) เข้ามาใช้ในการจัดการและทำนายด้วยวิธีที่มีประสิทธิภาพ สำหรับการทำนายราคาข้าวโพดเลี้ยงสัตว์นั้นเป็นกระบวนการที่มีความซับซ้อนเนื่องจากเป็นข้อมูลอนุกรมเวลา (Time series) และยังมีปัจจัยต่าง ๆ ทั้งภายในและภายนอก เช่น ตลาดสินค้าเกษตรกรรมมีความผันผวนสูงและยังมีผลต่อการเปลี่ยนแปลงของราคาข้าวโพดเลี้ยงสัตว์อีกด้วย ดังนั้นแบบจำลองที่ได้รับการปรับปรุงอย่างมีประสิทธิภาพจะช่วยให้การทำนายราคาข้าวโพดเลี้ยงสัตว์มีความแม่นยำมากขึ้น

งานวิจัยในช่วง พ.ศ. 2561 - 2566 ที่เกี่ยวข้องกับการทำนายราคาผลิตภัณฑ์ทางการเกษตรมีหลายงานวิจัย ในงานวิจัยนี้จึงเลือกงานวิจัยที่มีความเกี่ยวข้องมาเป็นแนวทางในการทำวิจัยจำนวน 9 งานวิจัย งานวิจัยแรกเป็นของ Ge & Wu (2019) ได้วิเคราะห์แนวโน้มการเปลี่ยนแปลงของราคาข้าวโพดในประเทศจีนโดยใช้ข้อมูลราคาข้าวโพดเลี้ยงสัตว์ตั้งแต่ปี พ.ศ. 2548 - 2559 และใช้แบบจำลองการถดถอยไม่เชิงเส้นตัวแปรเดียว (Univariate nonlinear) และการถดถอยเชิงเส้นหลายตัวแปร (Multivariate Linear Regression) ในการทำนายราคาข้าวโพดในปี พ.ศ. 2560 จากนั้นใช้ข้อมูลจากการวิเคราะห์ที่ได้ มาวิเคราะห์ร่วมกับข้อมูลการบริโภคข้าวโพดในประเทศจีน การผลิตข้าวโพด และปริมาณการนำเข้า-ส่งออกข้าวโพด ผลการวิเคราะห์พบว่าการรวมข้อมูลเหล่านี้ทำให้การพยากรณ์แม่นยำมากขึ้น และจากงานวิจัยของ Jantankaew & Soonthornphisaj (2023) ได้ใช้แบบจำลองการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression: MLR) แบบจำลองการถดถอยแบบริดจ์ (Ridge Regression) และแบบจำลองการถดถอยแบบลาสโซ (Lasso Regression) ในการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์ในประเทศ โดยใช้ข้อมูลรูปแบบรายเดือนของราคาข้าวโพดเลี้ยงสัตว์และมันสำปะหลังโรงงาน ตั้งแต่ปี พ.ศ. 2545 - 2562 ผลการทดลองชี้ให้เห็นว่าราคาขายข้าวโพดล่วงหน้าในตลาดชิคาโกและราคาขายมันสำปะหลัง มีความสัมพันธ์ในทิศทางเดียวกันกับราคาข้าวโพดในประเทศ

Silva et al. (2021) ทำนายราคาข้าวโพดและน้ำตาลรายวันในประเทศบราซิล โดยใช้ข้อมูลรายวันประกอบด้วยราคาน้ำตาล ตั้งแต่ปี พ.ศ. 2547 - 2562 และราคาข้าวโพด ตั้งแต่ปี พ.ศ. 2546 - 2562 มาวิเคราะห์โดยใช้แบบจำลองค่าเฉลี่ยเคลื่อนที่รวมแบบถดถอยอัตโนมัติหรืออาร์มา (Autoregressive Integrated Moving Average: ARIMA) ค่าเฉลี่ยเคลื่อนที่รวมอัตโนมัติแบบถดถอยตามฤดูกาล (Seasonal Autoregressive Integrated Moving Average: SARIMA) ซัพพอร์ตเวกเตอร์แบบถดถอย (Support Vector Regression: SVR) การเพิ่มประสิทธิภาพแบบปรับตัว (Adaptive Boosting) หรือเรียกว่า AdaBoost และแบบจำลองหน่วยความจำระยะสั้นระยะยาวหรือแอลเอสทีเอ็ม (Long Short-Term Memory: LSTM) พบว่า SVR ให้ผลลัพธ์ที่ดีที่สุด รองลงมาคือ LSTM และจากงานวิจัยของ Guo et al. (2022) ได้ทำนายราคาข้าวโพดในมณฑลเสฉวนของประเทศจีน โดยใช้อัลกอริทึมอาพริออริ (Apriori algorithm) เพื่อค้นหาปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงของราคา และใช้แบบจำลองหน่วยความจำระยะสั้นระยะยาวตามความสนใจ (Attention-based LSTM: AttLSTM) ร่วมกับ ARIMA และแบบจำลองโครงข่ายประสาทเทียมแบบการแพร่ย้อนกลับหรือพีพีเอเอ็นเอ็น (Backpropagation Artificial Neural Networks: BPANN) พบว่าแบบจำลอง AttLSTM-ARIMA-BPANN ให้ผลการคาดการณ์ที่ดีที่สุด

งานวิจัยของ Yin et al. (2020) ทำนายราคาผัก 5 ชนิดแบบรายเดือนในประเทศเกาหลี โดยใช้ข้อมูลราคาเฉลี่ยรายเดือน ข้อมูลสภาพอากาศ ข้อมูลทางอุตุนิยมวิทยาและอื่น ๆ ตั้งแต่ปี พ.ศ. 2555 -

2562 งานวิจัยนี้ใช้แบบจำลอง STL-AttLSTM เปรียบเทียบกับ LSTM และ AttLSTM พบว่าแบบจำลอง STL-AttLSTM ให้ผลลัพธ์ที่ดีที่สุด โดยมีค่าโดยมีค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) และค่าเฉลี่ยความคลาดเคลื่อนเปอร์เซ็นต์สัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) ที่ต่ำที่สุด และงานวิจัยของ Zhu et al. (2022) ได้ทำนายราคาสุกรโดยใช้แบบจำลอง STL-based Hybrid โดยผลการทดลองชี้ให้เห็นว่าแบบจำลองที่นำเสนอมีประสิทธิภาพสูงสุดในการทำนายราคาสุกร และจากงานวิจัยของ Sabu et al. (2020) ได้ทำนายราคาผลหมากทางการเกษตรของอินเดียในรัฐเกรละ โดยใช้ข้อมูลราคาขายผลหมากแบบรายเดือน ตั้งแต่ปี พ.ศ. 2550 - 2559 ทำนายโดยใช้แบบจำลอง SARIMA แบบจำลองโฮลท์-วินเทอร์ซีซันนัล (Holt-winter's Seasonal) และแบบจำลอง LSTM จากการเปรียบเทียบผลลัพธ์ของแต่ละแบบจำลองพบว่าแบบจำลอง LSTM มีประสิทธิภาพในการทำนายมากที่สุด

จากงานวิจัยของ Kraisornukhor et al. (2023) ได้ทำนายราคายางแผ่นรมควันชั้นสามล่งหน้าหนึ่งสัปดาห์โดยใช้แบบจำลองการเรียนรู้เชิงลึกพร้อมปัจจัยภายนอก ซึ่งใช้ข้อมูลราคาขายรายวันจากการยางแห่งประเทศไทย ตั้งแต่ปี พ.ศ. 2544 - 2566 โดยแบ่งข้อมูล 80% เป็นชุดข้อมูลฝึกสอน และ 20% เป็นชุดข้อมูลทดสอบ โดยงานวิจัยนี้ทำการเปรียบเทียบเพื่อหาแบบจำลองการทำนายที่ดีที่สุดระหว่างแบบจำลองโครงข่ายประตูกลับ (Gated Recurrent Units: GRU) แบบจำลอง LSTM แบบจำลอง GRU แบบสองทิศทาง (Bidirectional-GRU) และแบบจำลอง LSTM แบบสองทิศทาง (Bidirectional-LSTM) จากผลการทดลองพบว่า แบบจำลอง GRU มีประสิทธิภาพในการทำนายที่ดีที่สุด

จากงานวิจัยของ Nayak et al. (2024) ได้ทำนายราคาสินค้าเกษตรในประเทศอินเดีย โดยใช้ราคาสินค้าเกษตรร่วมกับปัจจัยภายนอกและข้อมูลสภาพอากาศ ข้อมูลพืชยอดนิยมรายสัปดาห์ของมะเขือเทศ หัวหอม และมันฝรั่ง จากรัฐบาลอินเดีย โดยแบ่งข้อมูล 90% เป็นชุดข้อมูลฝึกสอน และ 10% เป็นชุดข้อมูลทดสอบ งานวิจัยนี้ได้ทำการทดลองเชิงเปรียบเทียบที่ครอบคลุมทั้งแบบจำลองทางสถิติ เช่น แบบจำลองค่าเฉลี่ยเคลื่อนที่รวมแบบถดถอยอัตโนมัติพร้อมตัวแปรภายนอกหรืออาร์แม็กซ์ (Autoregressive Integrated Moving Average with Exogenous Variables: ARIMAX) แบบจำลองการเรียนรู้ของเครื่อง (Machine learning) และแบบจำลองการเรียนรู้เชิงลึก (Deep learning) ผลการวิจัยพบว่าแบบจำลองการเรียนรู้เชิงลึกที่ใช้ตัวแปรภายนอกให้ประสิทธิภาพในการทำนายที่ดีที่สุด

งานวิจัยนี้จึงได้เสนอการเปรียบเทียบประสิทธิภาพแบบจำลองสำหรับการทำนายราคาข้าวโพดเลี้ยงสัตว์ โดยนำแบบจำลองทางสถิติและแบบจำลองการเรียนรู้เชิงลึกที่ได้รับความนิยมจากการทบทวนวรรณกรรมข้างต้นมาทำการศึกษาทั้งหมด 4 แบบจำลอง ได้แก่ LSTM, GRU, ARIMA และ ARIMAX เพื่อค้นหาแบบจำลองที่สามารถทำนายราคาข้าวโพดได้อย่างมีประสิทธิภาพ งานวิจัยนี้จะช่วยให้เกษตรกรสามารถทราบแนวโน้มการเปลี่ยนแปลงราคาของข้าวโพดเลี้ยงสัตว์และสามารถวางแผนการเพาะปลูกในอนาคตได้อย่างมีประสิทธิภาพ

2. วิธีดำเนินการวิจัย

2.1 ชุดข้อมูล

ข้อมูลดิบที่ใช้ในงานวิจัยนี้เก็บรวบรวมมาจากสองแหล่งข้อมูล คือ สำนักงานเศรษฐกิจการเกษตร (Office of Agricultural Economics, 2023) และกระทรวงพาณิชย์ (Ministry of Commerce, 2023) โดยประกอบไปด้วยข้อมูล 6 รายการ รายการแรกคือ ราคาข้าวโพดเลี้ยงสัตว์ความชื้น 14.5% ที่เกษตรกรขายได้ (บาทต่อกิโลกรัม) ซึ่งต่อไปนี้จะเรียกว่า ราคาข้าวโพด และเป็นข้อมูลหลักที่ใช้ทำนาย สำหรับอีก 5 รายการที่เหลือจะเป็นคุณลักษณะอื่นที่ใช้ประกอบการทำนาย โดยงานวิจัยนี้ใช้รายงานของภาวะเศรษฐกิจการเกษตรปี พ.ศ. 2566 และแนวโน้มปี พ.ศ. 2567 (Fernando, 2024) ในการเลือกคุณลักษณะที่เกี่ยวข้องซึ่งประกอบไปด้วย จำนวนพื้นที่เพาะปลูกข้าวโพดรวมทั้งประเทศ (ไร่) ปริมาณผลผลิตข้าวโพดของเกษตรกรรวมทั้งประเทศ (ตัน) ปริมาณการนำเข้าและส่งออกข้าวโพดเลี้ยงสัตว์ของประเทศไทยที่ประเทศไทยนำเข้าหรือส่งออกสินค้าได้ (ตัน) มูลค่าของการนำเข้าและส่งออกข้าวโพดเลี้ยงสัตว์ของประเทศไทย (ล้านบาท) และราคาซื้อขายล่วงหน้าตลาดชิคาโก (บาทต่อกิโลกรัม)

ข้อมูลทุกข้อมูลยกเว้นข้อมูลราคาซื้อขายล่วงหน้าตลาดชิคาโกได้นำมาจากสำนักงานเศรษฐกิจการเกษตรและข้อมูลทั้งหมดอยู่ในรูปแบบข้อมูลรายเดือน งานวิจัยเลือกใช้ข้อมูลตั้งแต่ปี พ.ศ. 2558 - 2564 เนื่องจากเมื่อตรวจสอบข้อมูลทั้งหมดแล้วพบว่า ข้อมูลย้อนหลังก่อนปี พ.ศ. 2558 มีข้อมูลสูญหายจำนวนมากใน 1 ปี

2.2 ขั้นตอนการวิจัย

งานวิจัยนี้ใช้ภาษาไพทอน (Python) ในการทำวิจัย โดยมีขั้นตอนการวิจัยทั้งหมด 6 ขั้นตอนดังต่อไปนี้

1) รวบรวมข้อมูลที่อยู่ในรูปแบบแตกต่างกันซึ่งประกอบไปด้วยข้อมูลบนเว็บไซต์และไฟล์พีดีเอฟ ให้อยู่ในไฟล์ไมโครซอฟท์เอกซ์เซล (Microsoft excel)

2) สร้างคุณลักษณะใหม่จำนวน 2 คุณลักษณะที่คาดว่าจะมีความเกี่ยวข้องกับราคาของข้าวโพดในประเทศไทย จากข้อมูลประกอบไปด้วย ราคานำเข้าข้าวโพดเฉลี่ย ซึ่งเกิดจากการนำมูลค่าของการนำเข้าหารด้วยปริมาณการนำเข้าได้ข้อมูลในหน่วยบาทต่อกิโลกรัม และราคาส่งออกข้าวโพดเฉลี่ย ซึ่งเกิดจากการนำมูลค่าของการส่งออกหารด้วยปริมาณการส่งออกได้ข้อมูลในหน่วยบาทต่อกิโลกรัม

3) ตรวจสอบและแก้ไขข้อมูลที่เป็นค่าว่าง (Null value) และข้อมูลที่มีค่าผิดปกติ (Noisy data) โดยการแทนข้อมูลเหล่านี้ด้วยค่าเฉลี่ยของข้อมูลตำแหน่งถัดไปและก่อนหน้าที่ใกล้ที่สุด จากนั้นจัดเตรียมข้อมูลให้อยู่ในรูปแบบไฟล์ซีเอสวี (CSV) เพื่อนำไปใช้ต่อไป

4) การสร้างแบบจำลองมีคุณลักษณะที่เกี่ยวข้องจำนวนมากและมีจำนวนคุณลักษณะแตกต่างกันออกไปในแต่ละชุดข้อมูล อย่างไรก็ตาม คุณลักษณะบางอันอาจไม่ส่งผลต่อประสิทธิภาพของแบบจำลองหรืออาจลดประสิทธิภาพของแบบจำลอง ดังนั้นก่อนการสร้างแบบจำลองจึงต้องตรวจสอบว่ามีเฉพาะคุณสมบัติที่จำเป็นถูกเลือกเพื่อไปสร้างแบบจำลองเท่านั้น สัมประสิทธิ์สหสัมพันธ์ (Correlation coefficient) เป็นหนึ่งในการเลือกคุณลักษณะที่ได้รับความนิยม ซึ่งเป็นหลักการทางสถิติที่ใช้หาความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปร (Soontranon, 2023) กำหนดให้ค่า r แทนค่าสัมประสิทธิ์

สหสัมพันธ์ซึ่งจะมีค่าอยู่ระหว่าง -1 ถึง 1 จะสามารถตีความค่า r ได้ 3 กรณีคือ กรณีแรกค่า r เข้าใกล้ -1 หมายถึงตัวแปรทั้งสองตัวมีความสัมพันธ์กันในเชิงตรงกันข้าม กรณีที่ 2 ค่า r เข้าใกล้ 1 หมายความว่าตัวแปรทั้งสองตัวมีความสัมพันธ์ไปในทิศทางเดียวกันและกรณีสุดท้ายค่า $|r|$ มีค่าเข้าใกล้ 0 จะหมายความว่าตัวแปรทั้งสองมีความสัมพันธ์กันน้อยหรือไม่มีความสัมพันธ์กัน

5) ช่วงของข้อมูลในแต่ละคุณลักษณะมีความแตกต่างกัน ทำให้จำเป็นต้องปรับขนาดช่วงของข้อมูลเพื่อให้ข้อมูลเป็นมาตรฐานก่อน ซึ่งจะทำให้แบบจำลองทำงานได้ดีและให้ผลลัพธ์ที่มีความแม่นยำมากขึ้นสำหรับการปรับช่วงของข้อมูล งานวิจัยนี้ได้เลือกการปรับช่วงข้อมูลด้วยค่าต่ำสุด-สูงสุด (Min-Max normalization) (Ministry of Commerce, 2023) โดยการปรับแต่ละคุณลักษณะทุกคุณลักษณะรวมทั้งราคาข้าวโพดให้อยู่ในช่วง 0 ถึง 1 จากนั้นแบ่งข้อมูลออกเป็นชุดข้อมูลฝึกสอน (Training dataset) และชุดข้อมูลทดสอบ (Testing dataset) เนื่องจากมีจำนวนข้อมูลทั้งหมด 7 ปี งานวิจัยนี้จึงแบ่งการทดสอบออกเป็น 2 รูปแบบคือ ใช้ข้อมูลปี พ.ศ. 2558 - 2562 เป็นชุดข้อมูลฝึกสอน และใช้ข้อมูลปี พ.ศ. 2563 เป็นชุดข้อมูลทดสอบ และใช้ข้อมูลปี พ.ศ. 2559 - 2563 เป็นชุดข้อมูลฝึกสอน และใช้ข้อมูลปี พ.ศ. 2564 เป็นชุดข้อมูลทดสอบ โดยการสร้างแบบจำลองจะแยกสร้างแบบจำลองสำหรับชุดข้อมูลฝึกสอนแต่ละชุด และรายงานประสิทธิภาพแบบจำลองเฉลี่ยจากชุดข้อมูลทดสอบทั้ง 2 ชุด

6) งานวิจัยนี้จะนำชุดข้อมูลฝึกสอนมาสร้างแบบจำลองทั้งหมดสี่แบบจำลองแบ่งออกเป็นแบบจำลองทางสถิติ 2 แบบจำลอง และแบบจำลองการเรียนรู้เชิงลึก 2 แบบจำลอง ดังนี้

แบบจำลองทางสถิติแบบจำลองแรก คือ ARIMA เป็นแบบจำลองที่ประกอบไปด้วย 3 ส่วน คือ การถดถอยอัตโนมัติ (Autoregressive: AR) ใช้ตัวแปร p แทน ซึ่งคือจำนวนพจน์ที่บ่งบอกถึงส่วนที่เป็นค่าการถดถอยอัตโนมัติ อินทิเกรต (Integrated) ใช้ตัวแปร d แทน ซึ่งคือจำนวนค่าที่ทำให้ข้อมูลคงที่เพื่อทำให้ข้อมูลมีลักษณะนิ่ง (Stationary) และค่าเฉลี่ยเคลื่อนที่ (Moving average) หรือเรียกว่า MA ใช้ตัวแปร q แทน ซึ่งคือพารามิเตอร์ที่บ่งบอกถึงจำนวนของค่าที่ถูกใช้ในแบบจำลองค่าเฉลี่ยเคลื่อนที่ (Boonmana & Kulvanich, 2017; Bora, 2021) แบบจำลองทางสถิติแบบจำลองที่สอง คือ แบบจำลอง ARIMAX เป็นแบบจำลองที่นำแบบจำลอง ARIMA มาเพิ่มตัวแปรภายนอก (Exogenous) เข้าไปในแบบจำลอง ซึ่งตัวแปรภายนอกนี้จะเป็คุณลักษณะที่มีความเกี่ยวข้องกับตัวแปรที่ต้องการทำนาย (Mehandzhiyski, 2023)

อย่างไรก็ตาม เนื่องจากข้อมูลอนุกรมเวลาส่วนใหญ่อาจมีลักษณะไม่นิ่ง (Non-stationary) ทำให้จำเป็นต้องตรวจสอบก่อนว่าข้อมูลที่เลือกมานั้นมีลักษณะนิ่งหรือไม่ ไม่เช่นนั้นจะทำให้เกิดความผิดพลาดหลังการสร้างแบบจำลองได้ สำหรับการทดสอบความนิ่งของข้อมูลทำได้โดยใช้วิธีการทดสอบแบบเอ็ดเอฟ (Augmented Dickey-Fuller Test: ADF Test) หรือเรียกว่า ซึ่งเป็นการทดสอบประเภทการทดสอบรากหน่วย (Unit root test) ที่มีความเกี่ยวข้องกับความไม่นิ่งของอนุกรมเวลา โดยงานวิจัยนี้กำหนดระดับนัยสำคัญที่ 0.05 ดังนั้นหากผลลัพธ์ของ ADF Test มีค่าน้อยกว่าหรือเท่ากับระดับนัยสำคัญที่กำหนดจะหมายความว่าข้อมูลมีลักษณะนิ่ง และหากผลลัพธ์ของ ADF Test มีค่ามากกว่าระดับนัยสำคัญที่กำหนดจะหมายความว่าข้อมูลมีลักษณะไม่นิ่ง สำหรับการกำหนดค่า Integrated ของแบบจำลอง ARIMA และ ARIMAX จะกำหนดจากการ ADF Test ที่ทำให้ข้อมูลมีลักษณะนิ่ง (ALGOADDICT, 2019; Santra, 2023)

เพื่อช่วยกำหนดค่าอินทิเกรตสำหรับ ARIMA และ ARIMAX โดยการใช้วิธีทดสอบสมมติฐาน ซึ่งกำหนดให้สมมติฐานหลัก (H_0) แทนข้อมูลมีลักษณะไม่นิ่ง ซึ่งจะยอมรับสมมติฐานหลักเมื่อค่าที่คำนวณได้มี

ค่ามากกว่าค่าวิกฤต (Critical values) ณ ระดับนัยสำคัญที่กำหนด และสมมติฐานรอง (H1) แทนข้อมูลมีลักษณะหนึ่ง ซึ่งจะยอมรับสมมติฐานรองเมื่อค่าที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่าวิกฤต ณ ระดับนัยสำคัญที่กำหนด (Tejwin, 2022)

แบบจำลองการเรียนรู้เชิงลึกแบบจำลองแรก คือ แบบจำลอง LSTM เป็นโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent Neural Networks: RNN) ถูกเสนอโดย Hochreiter & Schmidhuber (1997) แบบจำลอง LSTM ถูกออกแบบมาเพื่อจัดการกับข้อมูลในระยะยาว ทำให้สามารถประมวลผลลำดับข้อมูลและทำนายผลข้อมูลในรูปแบบตามลำดับได้ เมื่อมีการป้อนข้อมูลลงในเซลล์แบบจำลอง LSTM จะเรียนรู้ข้อมูลที่ป้อนลงในเซลล์ จากนั้นจะส่งข้อมูลจากการประมวลผลเวลาปัจจุบันไปประมวลผลเวลาถัดไป เกต (Gates) ของแบบจำลองจะทำหน้าที่ควบคุมการไหลของข้อมูลเข้าและออกจากเซลล์หน่วยความจำ โดยแบบจำลอง LSTM มีเกตทั้งหมด 3 แบบ ได้แก่ 1) เกตที่ถูกลืม (Forget gate) ซึ่งมีฟังก์ชันกระตุ้นทำหน้าที่ปรับช่วงของข้อมูลตามต้องการ สำหรับตัดสินใจว่าควรจดจำข้อมูลใดและควรลืมข้อมูลใด ทำให้แต่ละเซลล์มีระดับความสำคัญที่แตกต่างกัน 2) เกตทางเข้า (Input gate) จะกำหนดว่าต้องเพิ่มค่าโหนด (Node) ของข้อมูลที่จะจดจำ และ 3) เกตทางออก (Output gate) กำหนดว่าเซลล์หน่วยความจำใดควรมีอิทธิพลต่อผลลัพธ์ที่มีอยู่ ทั้งนี้แต่ละเซลล์จะมีชั้นซ่อนอยู่เรียกว่าหน่วยความจำระยะสั้น (Short term memory) และมีสถานะเซลล์เรียกว่าหน่วยความจำระยะยาว (Long term memory) แบบจำลอง LSTM สามารถสร้างจากคุณลักษณะเดียวและหรือจากหลายคุณลักษณะได้ (Hull, 2022)

แบบจำลองการเรียนรู้เชิงลึกแบบจำลองที่สอง คือ แบบจำลอง GRU เสนอโดย Cho et al. (2014) โดยมีจุดมุ่งหมายเพื่อแก้ปัญหาความลาดชันที่หายไป (Vanishing gradient problem) ของ RNN นอกจากนี้ GRU ยังเป็นโครงข่ายประสาทเทียมแบบวนกลับประเภทหนึ่งที่คล้ายกับแบบจำลอง LSTM แต่ใช้หน่วยความจำน้อยกว่าและทำงานเร็วกว่า อย่างไรก็ตามความแตกต่างระหว่าง GRU และ LSTM คือแบบจำลอง GRU จะเปลี่ยนสถานะโดยใช้เกต 2 เกต ได้แก่ อัปเดตเกต (Update gate) จะกำหนดว่าจะรวมเวกเตอร์เข้าในสถานะที่ซ่อน (Hidden state) มากน้อยเพียงใด และรีเซ็ตเกต (Reset gate) จะกำหนดว่าควรลืมสถานะที่ซ่อนไว้ก่อนหน้านี้มากแค่ไหน ซึ่งเกตทั้งสองนี้ทำหน้าที่กำหนดข้อมูลที่สามารถผ่านไปยังทางออก (Output) ทำให้สามารถฝึกแบบจำลองได้โดยไม่ต้องลบข้อมูลที่ไม่เกี่ยวข้องกับการทำนาย (Chaiyadecha, 2022)

สำหรับการสร้างแบบจำลองการเรียนรู้เชิงลึกในงานวิจัยนี้ จะทำการค้นหาอัลกอริทึมการหาค่าที่เหมาะสมที่สุดสำหรับค่าถ่วงน้ำหนัก (Weight optimization algorithm) และฟังก์ชันกระตุ้น (Activation function) ที่เหมาะสม

สำหรับอัลกอริทึมการหาค่าที่เหมาะสมที่สุดสำหรับค่าถ่วงน้ำหนักจะทำการปรับค่าถ่วงน้ำหนักในการสอนแบบจำลองแต่ละรอบ งานวิจัยนี้เลือกทดสอบอัลกอริทึมการหาค่าที่เหมาะสมที่สุดจำนวน 3 อัลกอริทึมประกอบไปด้วย 1) อัดัม (Adaptive Moment Estimation: ADAM) 2) อาร์เอ็มเอสพรอป (Root Mean Square Propagation: RMSProp) และ 3) เอสจีดี (Stochastic Gradient Descent: SGD)

สำหรับฟังก์ชันกระตุ้นจะทำหน้าที่ในการรวมและแปลงข้อมูลเข้าจำนวนมากให้เป็นข้อมูลออกเพียงหนึ่งค่า (Chaiyadecha, 2022) โดยในงานวิจัยนี้เลือกทดสอบทั้งหมด 3 ฟังก์ชัน ประกอบไปด้วย 1) รีลู (Rectified Linear Unit: ReLU) 2) ซิกมอยด์ (Sigmoid) และ 3) ไฮเพอร์โบลิคแทงกเจนท์ (Hyperbolic

Ttangent Function: Tanh) โดยฟังก์ชันกระตุ้นทั้ง 3 สามารถแปลงข้อมูลด้วยสมการที่ (1) สำหรับฟังก์ชัน ReLU โดยมีผลลัพธ์อยู่ในช่วง $[0, \infty]$ สมการที่ (2) สำหรับฟังก์ชัน Sigmoid โดยมีผลลัพธ์อยู่ในช่วง $(0, 1)$ และสมการที่ (3) สำหรับฟังก์ชัน Tanh โดยมีผลลัพธ์อยู่ในช่วง $(-1, 1)$ เมื่อ x แทนผลรวมค่าของข้อมูลกับค่าถ่วงน้ำหนัก

$$f(x) = \begin{cases} 0; & x < 0 \\ x; & x \geq 0 \end{cases} \quad (1)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

$$f(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (3)$$

งานวิจัยนี้กำหนดพารามิเตอร์สำหรับฝึกสอนแบบจำลอง LSTM และ GRU ด้วยค่าเริ่มต้นดังแสดงดังตารางที่ 1 อย่างไรก็ตาม งานวิจัยนี้มีการกำหนดจำนวน Epochs โดยพิจารณาจากกราฟระหว่างการสร้างแบบจำลอง เมื่อค่าความผิดพลาดของแบบจำลองเริ่มคงที่จะหยุดการฝึกสอนแบบจำลอง

ตารางที่ 1. การกำหนดพารามิเตอร์สำหรับแบบจำลอง LSTM และ GRU

พารามิเตอร์	ค่าที่กำหนด
Units	200
Time Step	12
Kernel Regularizer	l2(0.01)
Loss Function	MSE

2.3 การวัดประสิทธิภาพแบบจำลอง

ในงานวิจัยนี้เลือกวัดประสิทธิภาพของแบบจำลองทางสถิติและแบบจำลองการเรียนรู้เชิงลึกด้วยวิธีการทางสถิติ 2 วิธี คือ *RMSE* และ *MAE* ซึ่งสามารถคำนวณได้จากสมการที่ (4) และสมการที่ (5)

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (4)$$

$$MAE = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N} \quad (5)$$

เมื่อ $\hat{y}_i$ แทนค่าที่ได้จากแบบจำลอง ณ เวลา i

y_i แทนค่าจริง ณ เวลา i

N แทนจำนวนข้อมูลทดสอบทั้งหมด

3. ผลการวิจัยและอภิปรายผล

3.1 การทดลองการเลือกคุณลักษณะ

การทดลองนี้จะค้นหาคุณลักษณะที่มีความสัมพันธ์กับราคาข้าวโพดที่ต้องการทำนายโดยการหาค่าสัมประสิทธิ์สหสัมพันธ์ ซึ่งผลลัพธ์ที่ได้แสดงดังตารางที่ 2 โดยเห็นได้ว่าคุณลักษณะที่เกี่ยวข้องกับราคาข้าวโพด 4 คุณลักษณะ ประกอบไปด้วย ปริมาณส่งออกข้าวโพดทั้งหมด ราคาข้าวโพดส่งออก ราคาซื้อขายล่วงหน้าตลาดชิคาโก และมูลค่าส่งออกข้าวโพดทั้งหมด จะมีค่าสัมบูรณ์ของค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าสูงซึ่งแตกต่างจากคุณลักษณะอื่น งานวิจัยนี้จึงเลือกใช้คุณลักษณะ 4 คุณลักษณะดังกล่าวเพื่อไปทำการทดสอบกับทั้งแบบจำลองทางสถิติและแบบจำลองการเรียนรู้เชิงลึกในการทดลองถัดไป

ตารางที่ 2. ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างราคาข้าวโพดและคุณลักษณะอื่น

คุณลักษณะ	ค่าสัมประสิทธิ์สหสัมพันธ์	ค่าสัมบูรณ์ของค่าสัมประสิทธิ์สหสัมพันธ์
ปริมาณส่งออกข้าวโพดทั้งหมด	-0.3200	0.3200
ราคาข้าวโพดส่งออก	0.3102	0.3102
ราคาซื้อขายล่วงหน้าตลาดชิคาโก	0.2786	0.2786
มูลค่าส่งออกข้าวโพดทั้งหมด	-0.2741	0.2741
ปริมาณนำเข้าข้าวโพดทั้งหมด	0.2293	0.2293
ปริมาณผลผลิตข้าวโพดในประเทศ	-0.1784	0.1784
มูลค่านำเข้าข้าวโพดทั้งหมด	0.1449	0.1449
พื้นที่เพาะปลูกข้าวโพด	0.1046	0.1046
ราคาข้าวโพดนำเข้า	0.0298	0.0298

3.2 การทดลองการทดสอบความนิ่งของข้อมูล

จากการทดสอบความนิ่งของข้อมูลด้วย ADF Test พบว่า ผลจาก ADF Test ได้ค่า 0.0434 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญที่ 0.05 สามารถสรุปได้ว่า ข้อมูลมีลักษณะนิ่ง เมื่อมีการใช้ค่า Integrated เป็น 1 ทำให้สามารถกำหนดค่า d เป็น 1 สำหรับแบบจำลองทางสถิติทั้ง ARIMA และ ARIMAX

3.3 การทดลองการทำนายราคาข้าวโพดด้วยแบบจำลองทางสถิติ

สำหรับแบบจำลองทางสถิติทั้ง 2 แบบจำลอง คือ ARIMA และ ARIMAX จะมีการพิจารณาตัวแปรที่เกี่ยวข้องจำนวน 3 ตัว คือ p , d และ q อย่างไรก็ตามงานวิจัยนี้ได้กำหนดค่า d เป็น 1 ตาม ADF Test จากหัวข้อ 3.1 การทดลองนี้จะทำการหาค่า p และ q ที่เหมาะสมสำหรับทั้ง 2 แบบจำลอง โดยทำการทดลองค่าตั้งแต่ 0 ถึง 5 โดยในการสร้างแบบจำลอง ARIMA ใช้ข้อมูลราคาข้าวโพดเพียงอย่างเดียว และ ARIMAX ใช้ข้อมูลราคาข้าวโพดและคุณลักษณะอื่นอีก 4 คุณลักษณะที่ถูกเลือกจากขั้นตอนที่ 3.1

ผลการทำนายของทั้ง 2 แบบจำลองแสดงดังตารางที่ 3 และ 4 โดยเป็นการแสดงค่าเฉลี่ยของชุดข้อมูลทดสอบ 2 ชุด จากตารางที่ 3 พบว่า แบบจำลอง ARIMA มีผลการทำนายที่ไม่แตกต่างกันในแต่ละแบบจำลอง โดย ARIMA (1,1,2) ได้ประสิทธิภาพในการทำนายดีที่สุด โดยมีค่า RMSE และ MAE อยู่ที่ 0.3213 และ 0.2496 ตามลำดับ จากตารางที่ 4 พบว่า แบบจำลอง ARIMAX มีผลการทำนายที่ไม่แตกต่าง

กันในแต่ละแบบจำลองโดย ARIMAX (4,1,4) มีค่า MAE น้อยที่สุดอยู่ที่ 0.0941 และ ARIMAX (5,1,4) มีค่า RMSE น้อยที่สุดอยู่ที่ 0.1145 ดังนั้นงานวิจัยนี้จึงเลือกแบบจำลอง ARIMA (1,1,2) ARIMAX (4,1,4) และ ARIMAX (5,1,4) เพื่อนำไปเปรียบเทียบกับแบบจำลองการเรียนรู้เชิงลึกต่อไป

ตารางที่ 3. ประสิทธิภาพของแบบจำลอง ARIMA เฉลี่ยของชุดข้อมูลทดสอบ

แบบจำลอง	วิธีการทางสถิติ		แบบจำลอง	วิธีการทางสถิติ	
	RMSE	MAE		RMSE	MAE
ARIMA (1,1,1)	0.1099	0.0881	ARIMA (3,1,4)	0.1190	0.0966
ARIMA (1,1,2)	0.0937	0.0728	ARIMA (3,1,5)	0.1474	0.1255
ARIMA (1,1,3)	0.1445	0.1166	ARIMA (4,1,1)	0.1116	0.0895
ARIMA (1,1,4)	0.1280	0.1055	ARIMA (4,1,2)	0.1192	0.0951
ARIMA (1,1,5)	0.1362	0.1139	ARIMA (4,1,3)	0.1239	0.1009
ARIMA (2,1,1)	0.0992	0.0778	ARIMA (4,1,4)	0.1226	0.0995
ARIMA (2,1,2)	0.1034	0.0823	ARIMA (4,1,5)	0.1187	0.0970
ARIMA (2,1,3)	0.1135	0.0913	ARIMA (5,1,1)	0.1153	0.0937
ARIMA (2,1,4)	0.1446	0.1228	ARIMA (5,1,2)	0.1236	0.0995
ARIMA (2,1,5)	0.1473	0.1258	ARIMA (5,1,3)	0.1393	0.1157
ARIMA (3,1,1)	0.1109	0.0879	ARIMA (5,1,4)	0.1101	0.0869
ARIMA (3,1,2)	0.1098	0.0871	ARIMA (5,1,5)	0.1489	0.1284
ARIMA (3,1,3)	0.1442	0.1074	-	-	-

ตารางที่ 4. ประสิทธิภาพของแบบจำลอง ARIMAX เฉลี่ยของชุดข้อมูลทดสอบ

แบบจำลอง	ปี พ.ศ. 2563		แบบจำลอง	ปี พ.ศ. 2564	
	RMSE	MAE		RMSE	MAE
ARIMAX (1,1,1)	0.1401	0.1131	ARIMAX (3,1,4)	0.2157	0.2003
ARIMAX (1,1,2)	0.1656	0.1403	ARIMAX (3,1,5)	0.2467	0.2316
ARIMAX (1,1,3)	0.1697	0.1438	ARIMAX (4,1,1)	0.1490	0.1297
ARIMAX (1,1,4)	0.1950	0.1771	ARIMAX (4,1,2)	0.1448	0.1247
ARIMAX (1,1,5)	0.1977	0.1795	ARIMAX (4,1,3)	0.1576	0.1379
ARIMAX (2,1,1)	0.1610	0.1413	ARIMAX (4,1,4)	0.1176	0.0941
ARIMAX (2,1,2)	0.1907	0.1747	ARIMAX (4,1,5)	0.1747	0.1574
ARIMAX (2,1,3)	0.1633	0.1442	ARIMAX (5,1,1)	0.1447	0.1245
ARIMAX (2,1,4)	0.1815	0.1620	ARIMAX (5,1,2)	0.1466	0.1250
ARIMAX (2,1,5)	0.1950	0.1764	ARIMAX (5,1,3)	0.1286	0.1047
ARIMAX (3,1,1)	0.1560	0.1360	ARIMAX (5,1,4)	0.1145	0.0955
ARIMAX (3,1,2)	0.1594	0.1397	ARIMAX (5,1,5)	0.2142	0.2006
ARIMAX (3,1,3)	0.1612	0.1417	-	-	-

3.4 การทดลองการทำนายราคาข้าวโพดด้วยแบบจำลองการเรียนรู้เชิงลึก

สำหรับแบบจำลองการเรียนรู้เชิงลึกทั้งสองแบบจำลอง คือ LSTM และ GRU โดยแต่ละแบบจำลองจะทดลองกับอัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนักจำนวน 3 อัลกอริทึมคือ ADAM, RMSProp และ SGD และฟังก์ชันกระตุ้น 3 ฟังก์ชันคือ ReLU, Sigmoid และ Tanh เพื่อหาว่าอัลกอริทึมการหาค่าเหมาะที่สุด สำหรับค่าถ่วงน้ำหนักและฟังก์ชันกระตุ้นที่เหมาะสมที่สุด นอกจากนี้ในการทดลองนี้จะแบ่งข้อมูลที่ใช้ฝึกสอนออกเป็น 2 รูปแบบคือ รูปแบบที่ 1 ใช้เฉพาะข้อมูลราคาข้าวโพดเพื่อสร้างแบบจำลอง และรูปแบบที่ 2 ใช้ข้อมูลราคาข้าวโพดและคุณลักษณะอื่นอีก 4 คุณลักษณะที่ถูกเลือกจากขั้นตอนที่ 3.1 เพื่อสร้างแบบจำลอง สำหรับส่วนนี้จะแบ่งผลการทดลองออกเป็น 2 หัวข้อย่อยตามแบบจำลองที่ใช้คือ LSTM และ GRU

3.4.1 ผลการทดลองสำหรับแบบจำลอง LSTM

ในการทดลองของแบบจำลอง LSTM มีการสร้างจากชุดข้อมูล 2 รูปแบบ คือ LSTM ที่สร้างจากข้อมูลราคาข้าวโพดเท่านั้นเรียกว่า LSTM รูปแบบที่ 1 และ LSTM ที่สร้างจากข้อมูลราคาข้าวโพดร่วมกับคุณลักษณะอื่นเรียกว่า LSTM รูปแบบที่ 2 สำหรับผลการใช้แบบจำลอง LSTM รูปแบบที่ 1 และ 2 ไปทำนายชุดข้อมูลทดสอบแสดงตามตารางที่ 5 และ 6 ตามลำดับ

ตารางที่ 5. ประสิทธิภาพของแบบจำลอง LSTM รูปแบบที่ 1 เฉลี่ยของชุดข้อมูลทดสอบ

อัลกอริทึมการหาค่าเหมาะที่สุด สำหรับค่าถ่วงน้ำหนัก	ฟังก์ชันกระตุ้น	วิธีการทางสถิติ	
		RMSE	MAE
ADAM	ReLU	0.1671	0.1498
ADAM	Sigmoid	0.1844	0.1622
ADAM	Tanh	0.1021	0.0903
RMSProp	ReLU	0.1243	0.1044
RMSProp	Sigmoid	0.1865	0.1638
RMSProp	Tanh	0.1125	0.0965
SGD	ReLU	0.3738	0.3635
SGD	Sigmoid	0.1976	0.1760
SGD	Tanh	0.2542	0.2385

จากตารางที่ 5 ผลการทดลองแบบจำลอง LSTM ที่สร้างจากข้อมูลราคาข้าวโพดพบว่า อัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนัก ADAM และฟังก์ชันกระตุ้น Tanh ให้ผลลัพธ์ที่มีประสิทธิภาพดีที่สุด โดยมีค่า RMSE คือ 0.1021 และมีค่า MAE คือ 0.0903 สำหรับการเปรียบเทียบระหว่างอัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนักพบว่า SGD จะให้ผลลัพธ์ที่มีประสิทธิภาพต่ำที่สุด และสำหรับการเปรียบเทียบระหว่างฟังก์ชันกระตุ้นพบว่า Sigmoid จะให้ผลลัพธ์ที่มีประสิทธิภาพต่ำที่สุด เมื่อใช้งานร่วมกับ ADAM หรือ RMSProp

ตารางที่ 6. ประสิทธิภาพของแบบจำลอง LSTM รูปแบบที่ 2 เฉลี่ยของชุดข้อมูลทดสอบ

อัลกอริทึมการหาค่าเหมาะที่สุด สำหรับค่าถ่วงน้ำหนัก	ฟังก์ชันกระตุ้น	วิธีการทางสถิติ	
		RMSE	MAE
ADAM	ReLU	0.3034	0.2590
ADAM	Sigmoid	0.1703	0.1463
ADAM	Tanh	0.3648	0.3248
RMSProp	ReLU	0.1637	0.1492
RMSProp	Sigmoid	0.1599	0.1358
RMSProp	Tanh	0.3003	0.2757
SGD	ReLU	0.3417	0.3309
SGD	Sigmoid	0.2055	0.1841
SGD	Tanh	0.2088	0.1926

จากตารางที่ 6 ผลการทดลองแบบจำลอง LSTM ที่สร้างจากข้อมูลราคาข้าวโพดร่วมกับคุณลักษณะอื่นพบว่า อัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนัก RMSProp และฟังก์ชันกระตุ้น Sigmoid ให้ผลลัพธ์ที่มีประสิทธิภาพดีที่สุดโดยมีค่า RMSE คือ 0.1599 และมีค่า MAE คือ 0.1358 สำหรับการเปรียบเทียบระหว่างอัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนักพบว่า ADAM และ SGD จะให้ผลลัพธ์ที่มีประสิทธิภาพต่ำกว่า RMSProp ในภาพรวม และสำหรับการเปรียบเทียบระหว่างฟังก์ชันกระตุ้นพบว่า Sigmoid จะให้ผลลัพธ์ที่มีประสิทธิภาพดีที่สุด

3.4.2 ผลการทดลองสำหรับแบบจำลอง GRU

ในการทดลองของแบบจำลอง GRU มีการสร้างจากชุดข้อมูล 2 รูปแบบคือ GRU ที่สร้างจากข้อมูลราคาข้าวโพดเท่านั้นเรียกว่า GRU รูปแบบที่ 1 และ GRU ที่สร้างจากข้อมูลราคาข้าวโพดร่วมกับคุณลักษณะอื่นเรียกว่า GRU รูปแบบที่ 2 สำหรับผลการใช้แบบจำลอง GRU รูปแบบที่ 1 และ 2 ไปทำนายชุดข้อมูลทดสอบแสดงตามตารางที่ 7 และ 8 ตามลำดับ

ตารางที่ 7. ประสิทธิภาพของแบบจำลอง GRU รูปแบบที่ 1 เฉลี่ยของชุดข้อมูลทดสอบ

อัลกอริทึมการหาค่าเหมาะที่สุด สำหรับค่าถ่วงน้ำหนัก	ฟังก์ชันกระตุ้น	วิธีการทางสถิติ	
		RMSE	MAE
ADAM	ReLU	0.0905	0.0765
ADAM	Sigmoid	0.1134	0.0981
ADAM	Tanh	0.0965	0.0821
RMSProp	ReLU	0.0789	0.0644
RMSProp	Sigmoid	0.0814	0.0668
RMSProp	Tanh	0.0930	0.0666
SGD	ReLU	0.1582	0.1324
SGD	Sigmoid	0.1853	0.1627
SGD	Tanh	0.1700	0.1451

จากตารางที่ 7 ผลการทดลองแบบจำลอง GRU ที่สร้างจากข้อมูลราคาข้าวโพดพบว่า อัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนัก RMSProp และฟังก์ชันกระตุ้น ReLU ให้ผลลัพธ์ที่มีประสิทธิภาพดีที่สุดโดยมีค่า RMSE คือ 0.0789 และมีค่า MAE คือ 0.0644 สำหรับการเปรียบเทียบระหว่างอัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนักพบว่า SGD จะให้ผลลัพธ์ที่มีประสิทธิภาพต่ำที่สุด ส่วน ADAM และ RMSProp ให้ผลลัพธ์ใกล้เคียงกัน และสำหรับการเปรียบเทียบระหว่างฟังก์ชันกระตุ้นพบว่า Sigmoid ให้ผลลัพธ์ที่มีประสิทธิภาพต่ำที่สุด

ตารางที่ 8. ประสิทธิภาพของแบบจำลอง GRU รูปแบบที่ 2 เฉลี่ยของชุดข้อมูลทดสอบ

อัลกอริทึมการหาค่าเหมาะที่สุด สำหรับค่าถ่วงน้ำหนัก	ฟังก์ชันกระตุ้น	วิธีการทางสถิติ	
		RMSE	MAE
ADAM	ReLU	0.0780	0.0662
ADAM	Sigmoid	0.0996	0.0808
ADAM	Tanh	0.1046	0.0826
RMSProp	ReLU	0.0864	0.0721
RMSProp	Sigmoid	0.0943	0.0812
RMSProp	Tanh	0.1286	0.1060
SGD	ReLU	0.1180	0.0951
SGD	Sigmoid	0.1934	0.1683
SGD	Tanh	0.1447	0.1296

จากตารางที่ 8 ผลการทดลองแบบจำลอง GRU ที่สร้างจากข้อมูลราคาข้าวโพดร่วมกับคุณลักษณะอื่นพบว่า อัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนัก ADAM และฟังก์ชันกระตุ้น ReLU ให้ผลลัพธ์ที่มีประสิทธิภาพดีที่สุดโดยมีค่า RMSE คือ 0.0780 และมีค่า MAE คือ 0.0662 สำหรับการเปรียบเทียบระหว่างอัลกอริทึมการหาค่าเหมาะที่สุดสำหรับค่าถ่วงน้ำหนักพบว่า SGD จะให้ผลลัพธ์ที่มีประสิทธิภาพต่ำที่สุด และสำหรับการเปรียบเทียบระหว่างฟังก์ชันกระตุ้นพบว่า Tanh ให้ผลลัพธ์ที่มีประสิทธิภาพต่ำที่สุด

3.5 การเปรียบเทียบแบบจำลองทั้งหมด

ในส่วนนี้จะทำการเปรียบเทียบผลการทดลองทั้งหมดโดยผลลัพธ์ที่ได้แสดงตามตารางที่ 9 โดยพบว่าแบบจำลองการเรียนรู้เชิงลึกที่ใช้ข้อมูลราคาข้าวโพดและคุณลักษณะอื่นอีก 4 คุณลักษณะเพื่อสร้างแบบจำลองมีประสิทธิภาพในการทำนายที่ดีกว่าแบบจำลองทางสถิติ อย่างไรก็ตาม แบบจำลอง ARIMA ซึ่งใช้เฉพาะข้อมูลราคาข้าวโพดเพื่อสร้างแบบจำลองกลับมาประสิทธิภาพในการทำนายที่ดีกว่า LSTM เมื่อพิจารณาเปรียบเทียบระหว่างการใช้เฉพาะข้อมูลราคาข้าวโพดเพื่อสร้างแบบจำลอง และการใช้ข้อมูลราคาข้าวโพดและคุณลักษณะอื่นอีก 4 คุณลักษณะเพื่อสร้างแบบจำลองพบว่า การใช้เฉพาะข้อมูลราคาข้าวโพดเพื่อสร้างแบบจำลองจะให้ประสิทธิภาพในการทำนายที่สูงกว่า ยกเว้นแบบจำลอง GRU ที่ให้ประสิทธิภาพการทำนายใกล้เคียงกัน และเมื่อพิจารณาผลลัพธ์รายปีของแบบจำลอง GRU ตามตารางที่ 10 พบว่าแบบจำลอง GRU รูปแบบที่ 1 จะมีประสิทธิภาพการทำนายที่ดีกว่าในการทำนายปี พ.ศ. 2563 แต่

แบบจำลอง GRU รูปแบบที่ 2 จะมีประสิทธิภาพการทำนายที่ดีกว่าในการทำนายปี พ.ศ. 2564 อย่างไรก็ตาม ประสิทธิภาพการทำนายของ GRU รูปแบบที่ 1 มีการกระจายตัวมากกว่า ดังนั้นในงานวิจัยนี้จึงเลือก GRU รูปแบบที่ 2 เป็นแบบจำลองที่มีประสิทธิภาพมากที่สุด

ตารางที่ 9. ประสิทธิภาพการทำนายเฉลี่ยสองปีของแบบจำลองแบบต่าง ๆ

แบบจำลอง	วิธีการทางสถิติ	
	RMSE	MAE
ARIMA (1,1,2)	0.0937	0.0728
ARIMAX (4,1,4)	0.1176	0.0941
ARIMAX (5,1,4)	0.1145	0.0955
LSTM รูปแบบที่ 1	0.1021	0.0903
LSTM รูปแบบที่ 2	0.1599	0.1358
GRU รูปแบบที่ 1	0.0789	0.0644
GRU รูปแบบที่ 2	0.0780	0.0662

ตารางที่ 10. การเปรียบเทียบวิธีการทางสถิติแบบจำลอง GRU แยกรายปี

แบบจำลอง	พ.ศ. 2563		พ.ศ. 2564	
	RMSE	MAE	RMSE	MAE
GRU รูปแบบที่ 1	0.0739	0.0591	0.0840	0.0698
GRU รูปแบบที่ 2	0.0787	0.0685	0.0774	0.0639

4. สรุปผลการวิจัย

งานวิจัยนี้ค้นหาแบบจำลองที่เหมาะสมสำหรับการทำนายราคาข้าวโพดในประเทศไทย ชุดข้อมูลประกอบไปด้วยข้อมูลราคาข้าวโพดย้อนหลัง 7 ปี และข้อมูลคุณลักษณะอื่นที่เกี่ยวข้องอีก 9 คุณลักษณะ แบบจำลองที่เลือกประกอบไปด้วย ARIMA ARIMAX LSTM และ GRU จากการทดลองต่าง ๆ พบว่า ข้อมูลคุณลักษณะอื่นที่เลือกมา 4 คุณลักษณะที่เกี่ยวข้องกับราคาข้าวโพดมากที่สุดประกอบไปด้วย ปริมาณส่งออกข้าวโพดทั้งหมด ราคาข้าวโพดส่งออก ราคาซื้อขายล่วงหน้าตลาดชิคาโก และมูลค่าส่งออกข้าวโพดทั้งหมด สำหรับผลการทำนายพบว่า แบบจำลอง GRU ที่ใช้ข้อมูลราคาข้าวโพดและคุณลักษณะอื่นอีก 4 คุณลักษณะเพื่อสร้างแบบจำลองมีประสิทธิภาพในการทำนายที่ดีที่สุด การพัฒนาต่อยอดในอนาคตสามารถปรับปรุงแบบจำลอง GRU ในส่วนการปรับพารามิเตอร์อื่น ๆ เพิ่มเติม หรือการทำแบบจำลอง GRU ไปรวมกับแบบจำลองอื่นเป็นแบบจำลองแบบผสมผสานที่จะเพิ่มประสิทธิภาพในการทำนายต่อไป

เอกสารอ้างอิง (References)

- ALGOADDICT. (2019, June 22). *How ARIMA works in time series forecasting*. <https://shorturl.at/bKpnM> (in Thai)
- Boonmana, C., & Kulvanich, N. (2017). A comparative prediction accuracy of hybrid time series models. *Journal of Science and Technology*, 25(2), 177-190. (in Thai)

- Bora, N. (2021, November 9). *Understanding ARIMA models for machine learning*. Capital One. <https://www.capitalone.com/tech/machine-learning/understanding-arma-models/>
- Chaiyadecha, S. (2022, April 27). *Time series and stationary test using Python with time series data*. Medium. <https://lengyi.medium.com/time-series-stationary-python-adf-929ea3c64887> (in Thai)
- Cho, K., Merrienboer, B. V., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning phrase representations using RNN encoder-decoder for statistical machine translation*. arXiv. <https://arxiv.org/abs/1406.1078>
- Fernando, J. (2024). *The correlation coefficient: What it is and what it tells investors*. Investopedia. <https://www.investopedia.com/terms/c/correlationcoefficient.asp>
- Ge, Y., & Wu, H. (2019). Prediction of corn price fluctuation based on multiple linear regression analysis model under big data. *Neural Computing and Applications*, 32(22), 16843-16855. <https://doi.org/10.1007/s00521-018-03970-4>
- Guo, Y., Tang, D., Tang, W., Yang, S., Tang, Q., Feng, Y., & Zhang, F. (2022). Agricultural price prediction based on combined forecasting model under spatial-temporal influencing factors. *Sustainability*, 14(17), Article 10483. <https://doi.org/10.3390/su141710483>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hull, G. (2022, December 19). *Building a neural network zoo from scratch: The long short-term memory network*. Medium. <https://medium.com/@CallMeTwitch/building-a-neural-network-zoo-from-scratch-the-long-short-term-memory-network-1cec5cf31b7>
- Jantankaew, P., & Soonthornphisaj, N. (2023). Data analytics for maize price prediction using regression algorithms. *KKU Research Journal (Graduate Studies)*, 23(2), 92-106. (in Thai)
- Kraisornukhor, P., Wongchaisuwat, P., Paoprasert, N., & Mungwattana, A. (2023). Multi-step rubber price prediction using deep learning models with external factors. *Proceedings of 2023 IEEE 11th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)* (Vol. 11, pp. 1125-1129). IEEE. <https://doi.org/10.1109/ITAIC58329.2023.10408839>
- Mehandzhiyski, V. (2023). *What is an ARIMAX model?*. 365 Data Science. <https://365datascience.com/tutorials/python-tutorials/arimax/>
- Ministry of Commerce. (2023, November 14). *Historical prices: Maize (feed) – Chicago futures market price (THB/kg)*. Ministry of Commerce. <https://mex.moc.go.th/page/dit/checkpricedetail/type/W/catid/7/itemid/W16037> (in Thai)

- Nayak, G. H. H., Alam, M. W., Singh, K. N., Avinash, G., Kumar, R. R., Ray, M., & Deb, C. K. (2024). Exogenous variable driven deep learning models for improved price forecasting of TOP crops in India. *Scientific Reports*, 14(1), Article 17203. <https://doi.org/10.1038/s41598-024-68040-3>
- Office of Agricultural Economics Region 1, Chiang Mai. (2023, October 5). *This year, livestock corn production in 6 provinces of Northern Thailand totals 880,000 tons; Chiang Mai is the major production area, and farmers are preparing for harvest*. <https://shorturl.at/3FGul> (in Thai)
- Office of Agricultural Economics. (2023, November 14). *Agricultural economic data*. Office of Agricultural Economics, Ministry of Agriculture and Cooperatives. <https://shorturl.at/U0uW1> (in Thai)
- Office of Agricultural Research and Development Region 5, Department of Agriculture, Ministry of Agriculture and Cooperatives. (2019). *Knowledge management of post-rice corn production technology in the central region*. <https://www.doa.go.th/oard5/wp-content/uploads/2019/09/km62.pdf> (in Thai)
- Sabu, K. M., & Kumar, T. K. M. (2020). Predictive analytics in agriculture: Forecasting prices of arecanuts in Kerala. *Procedia Computer Science*, 171, 699-708. <https://doi.org/10.1016/j.procs.2020.04.076>
- Santra, R. (2023, May 12). *Tests for stationarity in time series — Dickey Fuller Test & Augmented Dickey Fuller (ADF) Test*. Medium. <https://medium.com/@ritusantra/tests-for-stationarity-in-time-series-dickey-fuller-test-augmented-dickey-fuller-adf-test-d2e92e214360>
- Silva, R. F., Barreira, B. L., & Cugnasca, C. E. (2021). Prediction of corn and sugar prices using machine learning, econometrics, and ensemble models. *Engineering Proceedings*, 9(1), Article 31. <https://doi.org/10.3390/engproc2021009031>
- Soontranon, N. (2023). *The difference between normalization and standardization*. https://www.nerd-data.com/normalization_standardization (in Thai)
- Tejwin. (2022, January 4). *ARIMA-GARCH model (Part 1)*. <https://www.tejwin.com/en/insight/arima-garch-modelpart-1/>
- Yin, H., Jin, D., Gu, Y. H., Park, C. J., Han, S. K., & Yoo, S. J. (2020). STL-ATTLSTM: Vegetable price forecasting using STL and attention mechanism-based LSTM. *Agriculture*, 10(12), Article 612. <https://doi.org/10.3390/agriculture10120612>

Zhu, H., Xu, R., & Deng, H. (2022). A novel STL-based hybrid model for forecasting hog price in China. *Computers and Electronics in Agriculture*, 198, Article 107068. <https://doi.org/10.1016/j.compag.2022.107068>