

การประยุกต์ใช้แบบจำลองโครงข่ายประสาทเทียมเชิงลึกสำหรับการพยากรณ์ราคาพืชไร่ของประเทศไทย

ธนกร สุทธิสนธ์*

สาขาวิชาสถิติประยุกต์, คณะวิทยาศาสตร์, มหาวิทยาลัยราชภัฏอุบลราชธานี

* ผู้ประสานงานเผยแพร่ (Corresponding Author), E-mail: thanakon.s@ubru.ac.th

วันที่รับบทความ: 20 เมษายน 2565; วันที่ทบทวนบทความ: 17 ตุลาคม 2565; วันที่ตอบรับบทความ: 19 ตุลาคม 2565

วันที่เผยแพร่ออนไลน์: 23 ธันวาคม 2565

บทคัดย่อ: การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเลือกตัวแบบการพยากรณ์ที่มีความเหมาะสมกับข้อมูลอนุกรมเวลาราคาพืชไร่ 3 ชนิด ได้แก่ ราคามันสำปะหลัง ราคาข้าวโพดเลี้ยงสัตว์ และราคาข้าวเปลือกเจ้าหอมมะลิ เก็บรวบรวมข้อมูลจากเว็บไซต์ของสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ตั้งแต่เดือน มกราคม พ.ศ. 2545 ถึงเดือน มกราคม พ.ศ. 2565 จำนวน 241 ค่า สร้างตัวแบบการพยากรณ์ 3 วิธี คือ ตัวแบบโครงข่ายประสาทเทียม (Artificial Neural Network Model: ANN) ตัวแบบโครงข่ายประสาทเทียมเชิงลึก (Long Short-Term Memory: LSTM) และตัวแบบผสม (Hybrid Model) เปรียบเทียบความแม่นยำของตัวแบบด้วยเกณฑ์การประเมินประสิทธิภาพ 3 เกณฑ์ คือ ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) รากของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) และร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percent Error: MAPE) ผลการวิจัยพบว่า ตัวแบบผสมมีประสิทธิภาพการพยากรณ์ดีกว่าตัวแบบโครงข่ายประสาทเทียมและตัวแบบโครงข่ายประสาทเทียมเชิงลึกในทุกชุดข้อมูลและทุกเกณฑ์การวัดประสิทธิภาพการพยากรณ์ สรุปได้ว่า ตัวแบบผสมมีความเหมาะสมในการพยากรณ์ข้อมูลอนุกรมเวลาราคาพืชไร่ทั้ง 3 ชนิด

คำสำคัญ: ราคาพืชไร่; ตัวแบบโครงข่ายประสาทเทียมเชิงลึก; ตัวแบบโครงข่ายประสาทเทียม; ตัวแบบผสม

Application of a Long Short - Term Memory Deep Natural Network Model for Field Crops Price Forecasting in Thailand

Thanakon Sutthison*

Program of Applied Statistics, Faculty of Science, Ubon Ratchathani Rajabhat University

* Corresponding author, E-mail: thanakon.s@ubru.ac.th

Received: 20 April 2022; Revised: 17 October 2022; Accepted: 19 October 2022

Online Published: 23 December 2022

Abstract: The objective of this research was to select a forecasting model suitable for the time series data of three sets of field crop prices, namely cassava, corn and paddy rice prices. Two-hundred and forty-one values were collected from the website of the Office of Agricultural Economics, Ministry of Agriculture and Cooperatives from January 2002 to January 2022. Three forecasting models were constructed: an artificial neural network model (ANN), a long short-term memory (LSTM) deep neural network model and a hybrid model. The models' accuracy was compared by three performance evaluations: mean absolute error (MAE), root mean square error (RMSE) and mean absolute percent error (MAPE). The results showed that the hybrid model provided better forecast accuracy than the ANN model and the LSTM model in all data sets and all forecasting performance criteria. It can be concluded that the hybrid model was suitable for forecasting the time series data of three sets of field crop prices.

Keywords: Field crops prices; Long short-term memory (LSTM); Artificial neural network model (ANN); Hybrid model



1. บทนำ

ประเทศไทยตั้งอยู่ในเขตภูมิประเทศที่เอื้อต่อการทำเกษตร ทำให้ประชากรส่วนใหญ่ในประเทศประกอบอาชีพเกษตรกรรม [1] ปัจจุบันประเทศไทยมีจำนวนประชากรในภาคการเกษตรมากถึง 10.52 ล้านคน [2] และภาคการเกษตรสามารถสร้างรายได้ต่อระบบเศรษฐกิจของประเทศประมาณร้อยละ 8.631 ของผลิตภัณฑ์มวลรวมทั้งประเทศ [3] การผลิตสินค้าทางการเกษตรของประเทศไทย แบ่งเป็น 5 ชนิด [4] ได้แก่ (1) พืชไร่ ได้แก่ ข้าวมันสำปะหลัง ข้าวโพดเลี้ยงสัตว์ (2) พืชสวน ได้แก่ ลำไย ยางพารา ปาล์มน้ำมัน ทูเรียน กาแฟ เป็นต้น (3) พืชผัก ได้แก่ หอมหัวใหญ่ กระเทียม หอมแดง หน่อไม้ฝรั่ง เป็นต้น (4) ประมง ได้แก่ ปลานิล ปลาดุก กุ้งทะเลเพาะเลี้ยง และ (5) ไม้ดอก เช่น กล้วยไม้ ตามลำดับ จากการพิจารณาข้อมูลการผลิตสินค้าทางการเกษตรดังกล่าวข้างต้น พบว่า ประชากรไทยให้ความสำคัญในการผลิตสินค้าทางการเกษตรพืชไร่สูงที่สุด เมื่อเทียบกับการผลิตสินค้าทางการเกษตรอื่นๆ ทั้งนี้เนื่องจากประเทศไทยมีเนื้อที่เพาะปลูกพืชไร่จำนวนมาก โดยมีพื้นที่เพาะปลูกพืชไร่ที่สำคัญ ได้แก่ ข้าวประมาณ 56,832,000 ไร่ ข้าวโพดเลี้ยงสัตว์ ประมาณ 6,881,000 ไร่ มันสำปะหลัง ประมาณ 8,918,000 ไร่ [5] จากพื้นที่เพาะปลูกจำนวนมาก การหามาตรการควบคุมการผลิตเพื่อให้สอดคล้องกับอุปสงค์และอุปทานมีความสำคัญอย่างยิ่ง หากผลผลิตออกมามากเกินความต้องการอาจส่งผลทำให้ราคาสินค้าตกต่ำ และหากผลผลิตออกมาน้อยอาจส่งผลทำให้สินค้าขาดแคลนไม่เพียงพอต่อความต้องการในอนาคต

จากการทบทวนวรรณกรรมที่เกี่ยวข้อง พบว่าราคาผลผลิตทางการเกษตรเป็นสิ่งจูงใจให้เกษตรกรขยายพื้นที่เพาะปลูกพืชไร่ [6] ดังนั้น การทราบราคาผลผลิตทางการเกษตรล่วงหน้าจะเป็นประโยชน์อย่างยิ่งในการเป็นสารสนเทศที่สำคัญ สำหรับผู้ผลิตและผู้ส่งออกสามารถใช้ประกอบการวางแผนในการตัดสินใจกับเหตุการณ์ที่จะเกิดขึ้นในอนาคตได้ [7] การคาดการณ์ราคาผลผลิตทางการเกษตรมีหลายวิธี วิธีการพยากรณ์เชิงสถิติเป็นวิธีหนึ่งที่มีการประยุกต์ใช้ในการพยากรณ์ราคาผลผลิตทางการเกษตรในปัจจุบัน [8-10] จากการศึกษา พบว่าข้อมูลอนุกรมเวลาราคาผลผลิตทางการเกษตรรายเดือนมีการเปลี่ยนแปลงตามเวลาที่เปลี่ยนไปและมีความไม่แน่นอนแทรกอยู่ [4] หากทำการพยากรณ์ข้อมูลอนุกรมเวลาด้วยวิธีการพยากรณ์อย่างง่าย เช่น วิธีการปรับให้เรียบ วิธีบ็อกซ์ – เจนกินส์ [6, 7] และวิธีโครงข่ายประสาทเทียม [11] อาจเกิดความคลาดเคลื่อนได้ ด้วยเหตุนี้ นักวิจัยจำนวนมากพยายามเสนอตัวแบบการพยากรณ์อนุกรมเวลาที่มีความซับซ้อนมากขึ้น โดยมีจุดมุ่งหมายเพื่อให้ได้ค่าพยากรณ์ที่มีความแม่นยำสูง ปัจจุบันมีการนำเทคนิคการพยากรณ์เชิงลึก (Deep Learning) มาประยุกต์ใช้ในการพยากรณ์อนุกรมเวลาเนื่องจากมีความแม่นยำสูงเมื่อเทียบกับตัวแบบดั้งเดิม [12], [13] มีนักวิจัยบางท่านประยุกต์ใช้ตัวแบบการพยากรณ์แบบผสม (Hybrid Forecasting) โดยนำเอาเทคนิควิธีการแยกองค์ประกอบของข้อมูลอนุกรมเวลาเชิงประจักษ์ (Empirical Mode Decomposition: EMD) เพื่อลดความผันผวนของข้อมูลอนุกรมเวลาก่อนนำไปพยากรณ์ด้วยเทคนิควิธีการพยากรณ์



โครงข่ายประสาทเทียมแบบลึก (Long Short Term Memory Deep Learning: LSTM) พบว่า ค่าพยากรณ์มีความแม่นยำสูง [14, 15] ปัจจุบันมีการนำแนวคิดดังกล่าวไปประยุกต์ใช้ในการพยากรณ์อนุกรมเวลาหลากหลาย เช่น การพยากรณ์ราคาหุ้น [16] การพยากรณ์การฉายรังสีแสงอาทิตย์ [17] การพยากรณ์การผลิตน้ำมันดิบ [18, 19] การพยากรณ์ความเร็วลม [20] การพยากรณ์ข้อมูลทางการเงิน [21] เป็นต้น

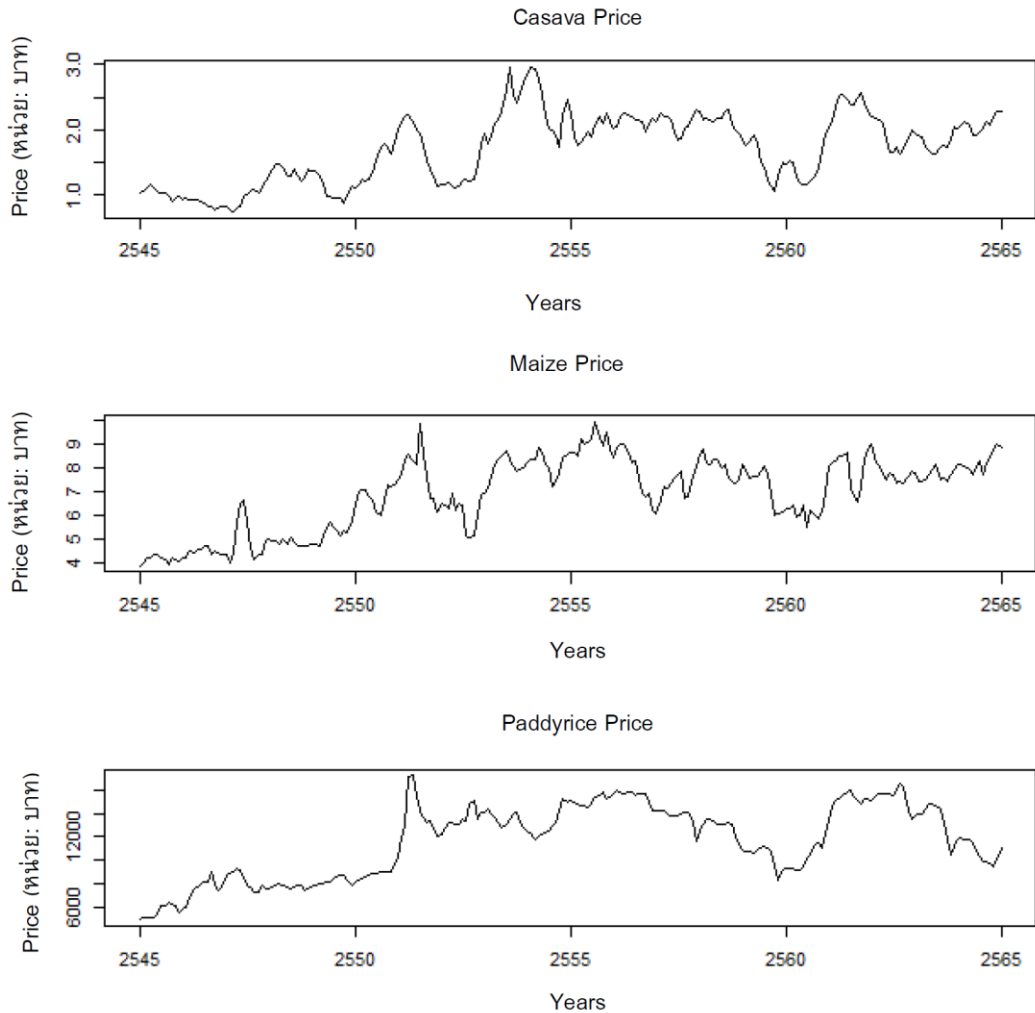
จากข้อค้นพบดังกล่าวข้างต้น พบว่า การพยากรณ์ราคาผลผลิตทางการเกษตรของประเทศไทย ส่วนใหญ่ ใช้วิธีการพยากรณ์พื้นฐานและข้อมูลมีความไม่แน่นอนแทรกอยู่ โดยเฉพาะราคาพืชไร่ ทำให้ค่าพยากรณ์มีความคลาดเคลื่อนสูง จากช่องว่างการวิจัยนี้จึงจำเป็นต้องมีการศึกษาเพิ่มเติมเพื่อเป็นสารสนเทศให้กับหน่วยงานที่เกี่ยวข้องนำไปใช้ในการตัดสินใจกับเหตุการณ์ที่เกิดขึ้นในอนาคต ด้วยเหตุนี้ผู้วิจัยจึงนำเสนอตัวแบบผสมโดยประยุกต์ใช้เทคนิควิธีการแยกองค์ประกอบของข้อมูลอนุกรมเวลาเชิงประจักษ์ร่วมกับเทคนิควิธีการพยากรณ์โครงข่ายประสาทเทียมแบบลึกพยากรณ์ราคาผลผลิตพืชไร่ของประเทศไทย เนื่องจากเป็นตัวแบบที่มีความแม่นยำสูง การศึกษาค้นคว้าครั้งนี้มุ่งศึกษาเฉพาะราคาพืชไร่ 3 ชนิด ประกอบด้วย ราคามันสำปะหลัง ราคาข้าวโพดเลี้ยงสัตว์ และราคาข้าวเปลือกเจ้าหอมมะลิ เหตุผลที่เลือกพยากรณ์ราคาพืชไร่ทั้ง 3 ชนิด เพราะเป็นพืชเศรษฐกิจประกอบกับมีพื้นที่เพาะปลูกจำนวนมาก หากสามารถสร้างตัวแบบการพยากรณ์ที่แม่นยำจะเป็นประโยชน์ต่อหน่วยงาน

ที่เกี่ยวข้อง สามารถใช้เป็นแนวทางในการวางแผนเพื่อประกอบการกำหนดแนวทางการบริหารจัดการราคาพืชไร่ได้อย่างมีประสิทธิภาพในอนาคต

2. วิธีการดำเนินการวิจัย

2.1 ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองครั้งนี้เป็นข้อมูลอนุกรมเวลาราคาพืชไร่ จำนวน 3 ชนิด ได้แก่ ราคามันสำปะหลัง (Casava Price) ราคาข้าวโพดเลี้ยงสัตว์ (Maize Price) และราคาข้าวเปลือกเจ้าหอมมะลิ (Paddyrice Price) ตั้งแต่เดือน มกราคม พ.ศ. 2545 ถึงเดือน มกราคม พ.ศ. 2565 ที่เผยแพร่ในเว็บไซต์ของสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ [5] จำนวน 241 ค่า แสดงในรูปที่ 1 ผู้วิจัยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่ 1 ข้อมูลฝึกฝน (Training Dataset) ตั้งแต่เดือนมกราคม พ.ศ. 2545 ถึงเดือนธันวาคม พ.ศ. 2562 จำนวน 216 ค่า สำหรับสร้างตัวแบบ ชุดที่ 2 ข้อมูลการทดสอบ (Testing Dataset) ตั้งแต่เดือนมกราคม พ.ศ. 2563 ถึงเดือนมกราคม 2565 จำนวน 25 ค่า การแบ่งข้อมูลดังกล่าวใช้เกณฑ์ร้อยละ 90 และร้อยละ 10 [22] เนื่องจากเก็บข้อมูลเป็นรายเดือนจำนวนคาบเวลาใน 1 ฤดูกาลเท่ากับ 12 เดือน ดังนั้นจึงแบ่งข้อมูลตามสัดส่วนดังกล่าว โดยมีวัตถุประสงค์เพื่อให้ครอบคลุมจำนวนคาบของฤดูกาลซึ่งมีผลต่อการทำนายของตัวแบบการพยากรณ์ทางสถิติ โดยการสร้างตัวแบบ มีรายละเอียดดังนี้



รูปที่ 1 ลักษณะการเคลื่อนไหวของข้อมูลอนุกรมเวลาราคาพืชไร่ จำนวน 3 ชุด ตั้งแต่เดือน มกราคม พ.ศ. 2545 ถึงเดือน มกราคม พ.ศ. 2565

2.2 วิธีการแยกองค์ประกอบของข้อมูลอนุกรมเวลาเชิงประจักษ์ (EMD)

วิธี EMD เป็นวิธีการแยกความถี่ของข้อมูลอนุกรมเวลาที่ไม่เป็นเชิงเส้นตรงและมีลักษณะเคลื่อนไหวไม่คงที่ (Non-Stationary) ออกจากกัน และมีการนำมาใช้ในการพยากรณ์ข้อมูลอนุกรมเวลา

อย่างแพร่หลาย [14, 15] นำเสนอโดย Huang [23] มีหลักการ คือ จะทำการลดสัญญาณแกว่งไวของข้อมูล (Fast Oscillation) โดยแยกเป็นฟังก์ชันภายใน (Intrinsic Mode Function: IMF) ด้วยเงื่อนไขดังนี้ (1) จำนวนจุดปลายสุด (Extrema Points) ประกอบด้วยจุดสูงสุดและจุดต่ำสุดและจำนวนจุดตัด



แกนศูนย์ต้องมีค่าเท่ากับหรือมีค่าต่างกันไม่เกินหนึ่ง (2) ณ จุดใด ๆ ของสัญญาณค่าเฉลี่ยขอบบนและขอบล่างจะมีค่าเท่ากับศูนย์ กระบวนการทำงานเพื่อสกัดชุดฟังก์ชันภายใน (IMF) ของวิธี EMD ประกอบด้วย 2 ขั้นตอน ดังนี้

1. การตรวจจับจุดปลายสุดของสัญญาณ (Extrema Detection) โดยการหาจุดสูงสุดและจุดต่ำสุดของแต่ละช่วงสัญญาณย่อย

2. การประมาณค่าขอบเขตพื้นผิว (Surface Interpolation) ของสัญญาณโดยการหาค่าประมาณสัญญาณขอบบนและสัญญาณขอบล่างและคำนวณหาสัญญาณเฉลี่ยตามสมการที่ (1) จากนั้นนำสัญญาณต้นฉบับลบออกด้วยสัญญาณเฉลี่ยจะได้สัญญาณผลลัพธ์ ขั้นตอนต่อไปตรวจสอบว่าสัญญาณผลลัพธ์นั้นเป็นฟังก์ชันภายใน (IMF) หรือไม่ ถ้าผลปรากฏว่าสัญญาณผลลัพธ์ที่ได้ไม่เป็นฟังก์ชันภายใน (IMF) ให้แทนที่สัญญาณต้นฉบับด้วยสัญญาณผลลัพธ์แล้วทำซ้ำกระบวนการดังกล่าวข้างต้น หากสัญญาณผลลัพธ์ที่ได้เป็นฟังก์ชันภายใน (IMF) ให้จัดเก็บสัญญาณผลลัพธ์นั้น จากนั้นทำการสกัดฟังก์ชันภายใน (IMF) ลำดับถัดไป สัญญาณต้นฉบับใหม่ที่ได้จากการนำสัญญาณต้นฉบับเดิมลบด้วยฟังก์ชันภายใน (IMF) แล้วทำกระบวนการข้างต้นซ้ำจนสัญญาณต้นฉบับเป็นสัญญาณฟังก์ชันเดียว

$$m(t) = \frac{e_{\max} + e_{\min}}{2} \quad (1)$$

เมื่อ $m(t)$ แทนสัญญาณเฉลี่ย e_{\max} , e_{\min} แทนสัญญาณขอบบนและสัญญาณขอบล่าง ซึ่งนำมาเขียนในรูปสมการได้ ดังสมการที่ (2)

2.3 ตัวแบบโครงข่ายประสาทเทียม (ANN)

ตัวแบบ ANN เป็นตัวแบบทางคณิตศาสตร์ที่จำลองขึ้นโดยอาศัยหลักการทำงานของสมองมนุษย์ เป็นเทคนิคการเรียนรู้โดยใช้เครื่อง (Machine Learning) ปัจจุบันมีการประยุกต์ใช้อย่างแพร่หลายในงานจำแนกประเภทและการพยากรณ์ [18] การประมวลผลของโครงข่ายประสาทเทียมจะเกิดขึ้นในหน่วยประมวลผลย่อย (Node) เมื่อมีข้อมูลนำเข้า (Input Data) เข้ามายังโครงข่าย ข้อมูลนำเข้าแต่ละค่าจะถูกนำมาคูณกับค่าถ่วงน้ำหนัก (Weight) แล้วนำมารวมกันเพื่อส่งผ่านไปยังฟังก์ชันการแปลง (Transfer Function) หรือฟังก์ชันกระตุ้น (Activation Function) และส่งข้อมูลผลลัพธ์ไปยังหน่วยประมวลผลย่อยถัดไป [24] โดยโครงข่ายประสาทเทียมเพอร์เซ็ปตรอนหลายชั้น (Multi - Layer Perceptron: MLP) จัดเป็นตัวแบบ ANN ที่นิยมใช้ในการพยากรณ์ข้อมูลอนุกรมเวลา [25] โครงสร้างของตัวแบบมีทั้งหมด 3 ชั้น ได้แก่ Input Layer, Hidden Layer และ Output Layer แต่ละชั้นจะประกอบด้วยโหนดที่มีจำนวนต่างกัน และโหนดในชั้น Input Layer จะเชื่อมไปข้างหน้า กับทุกโหนดในชั้นถัดไปตามลำดับจนถึง Output Layer ตัวแบบของ ANN [26] แสดงดังสมการที่ (3)

$$x(t) = \sum_{i=1}^n h_i(t) + r_N(t) \quad (2)$$

เมื่อ $x(t)$ แทน สัญญาณต้นฉบับ $h_i(t)$; $i=1, 2, 3, \dots, n$ เป็นชุดฟังก์ชันภายใน (IMFs) และ $r_N(t)$ เป็นส่วนเหลือ (Residual)

$$y_t = \alpha + \sum_{j=1}^n \alpha_j f\left(\sum_{i=1}^m \beta_{ij} y_{t-1} + \beta_j\right) + \varepsilon_t \quad (3)$$



เมื่อ y_t คือ ค่าสังเกตของข้อมูลอนุกรมเวลา ณ เวลา t , α คือ เวกเตอร์ของค่าถ่วงน้ำหนักระหว่างหน่วยประมวลผลย่อยจำนวน n หน่วยใน Hidden Layer กับหน่วยประมวลผลย่อยใน Output Layer, β_{ij} คือ ค่าถ่วงน้ำหนักระหว่างหน่วยประมวลผลย่อยแต่ละหน่วยใน Input Layer ซึ่งมีทั้งหมด m หน่วย กับหน่วยประมวลผลย่อยใน Hidden Layer โดยที่ $i = 1, 2, \dots, m; j = 1, 2, \dots, n; f$ คือ ฟังก์ชันกระตุ้นงานวิจัยนี้ใช้ฟังก์ชันกระตุ้น Sigmoid Logistic Function โดยมีสมการ ดังนี้ $f(x) = 1 / (1 + e^{-x})$ เนื่องจากทดลองใช้แล้วให้ค่าความแม่นยำกว่าฟังก์ชันกระตุ้นอื่น ๆ และ ε_t คือ ความคลาดเคลื่อน ณ เวลา t

2.4 ตัวแบบ Long Short-Term Memory (LSTM)

ตัวแบบ LSTM คือ โครงข่ายประสาทเทียมที่มีความสามารถด้านหน่วยความจำซึ่งเหมาะสำหรับการประมวลผลและคาดการณ์เหตุการณ์สำคัญที่มีช่วงเวลาค่อนข้างยาว [27] ซึ่งพัฒนาจากตัวแบบ Recurrent Neural Network (RNN) โดยมีหลักการทำงาน คือ สามารถเก็บ “สถานะ” หรือข้อมูลของแต่ละโหนดเอาไว้ เพื่อที่เวลาย้อนกลับไปที่จุดจะได้ทราบถึงที่มาของข้อมูลค่าดังกล่าวว่าเดิมเป็นค่าอะไร เทคนิค LSTM คือ มีฟังก์ชันพิเศษที่มีหน้าที่เสมือน “ประตู (Gate)” ที่คอยควบคุมข้อมูลที่จะเข้ามาในแต่ละโหนด ประกอบด้วย Forget Gate Layer, Input Gate Layer และ Output Gate Layer [14] แสดงดังรูปที่ 2

ขั้นตอนของ LSTM สามารถดำเนินการได้ดังนี้

1. Forget Gate Layer คือ Gate มีหน้าที่ในการกำหนดว่าข้อมูลที่เข้ามาใน Cell State ควรจะถูกเก็บไว้หรือควรที่จะทิ้งไป ข้อมูลที่ถูกตัดสินใจควรเก็บไว้

จะถูกประเมินจากข้อมูล Input ที่เข้ามาในโหนดนั้น ๆ รวมกับผลลัพธ์ที่ได้จากการคำนวณของโหนดก่อนหน้าผ่านฟังก์ชัน Sigmoid ดังสมการที่ (4)

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

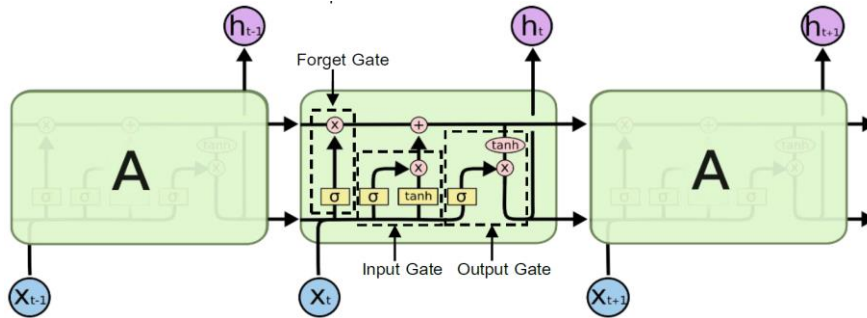
เมื่อ f_t คือ Forget Gate (มีค่าอยู่ระหว่าง 0 และ 1), σ คือ ฟังก์ชัน Sigmoid, W_f คือ ค่าน้ำหนักของ Matrices, h_{t-1} คือ ค่า Output ของ Cell State ก่อนหน้า, x_t คือ ค่า Input ที่เข้ามาใน Cell State ณ เวลา t และ b_f คือ ค่า Bias

2. Input Gate Layer คือ Gate มีหน้าที่รับข้อมูล Input เข้ามาใหม่แล้วจึงทำการบันทึกหรือเขียนข้อมูลลงไปในแต่ละโหนด การทำงานแบ่งออกเป็น 2 ส่วน คือ 1) ตรวจสอบการ Update Cell State เมื่อมีการรับข้อมูล Input โดยฟังก์ชัน Sigmoid ที่เป็นตัวควบคุมจะเรียกใช้ Input Gate เพื่อเลือกจะให้ Update Cell State หรือไม่ 2) ถ้า Input Gate เลือกทำการ Update Cell State ฟังก์ชัน Tanh จะทำการสร้าง Candidate Values (\tilde{C}_t) ขึ้นมาใน State ดังสมการที่ (5) และ (6)

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (6)$$

เมื่อ i_t คือ Input Gate, σ คือ ฟังก์ชัน Sigmoid, \tilde{C}_t คือ ค่า Candidate ของ Cell State ที่เวลา t , \tanh คือ ฟังก์ชัน Tanh, W_i, W_c คือ ค่าน้ำหนักของ Matrices, h_{t-1} คือ ค่า Output ของ Cell State ก่อนหน้า, x_t คือ ค่า Input ที่เข้ามาใน Cell State ณ เวลา t , b_i, b_c คือ ค่า Bias



รูปที่ 2 โครงสร้างการทำงานของตัวแบบ LSTM [28]

3. Output Gate Layer คือ Gate มีหน้าที่เตรียมทำการส่งออกข้อมูล (Output Data) ข้อมูลที่จะทำการ Output โดยดูจาก Cell State ที่ผ่านกระบวนการคำนวณต่าง ๆ เรียบร้อยแล้ว โดยฟังก์ชัน Sigmoid จะเป็นตัวเลือกว่าข้อมูลส่วนไหนใน Cell State ที่จะถูก Output จากนั้นนำค่า Cell State เข้าฟังก์ชัน Tanh (ค่าที่ได้มีค่าเป็น 1 หรือ -1) แล้วนำค่าที่ได้จากฟังก์ชัน Tanh มาทำการคำนวณกับค่า Output ที่ได้จาก Sigmoid Gate ได้ค่า Output ดังสมการที่ (7)

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

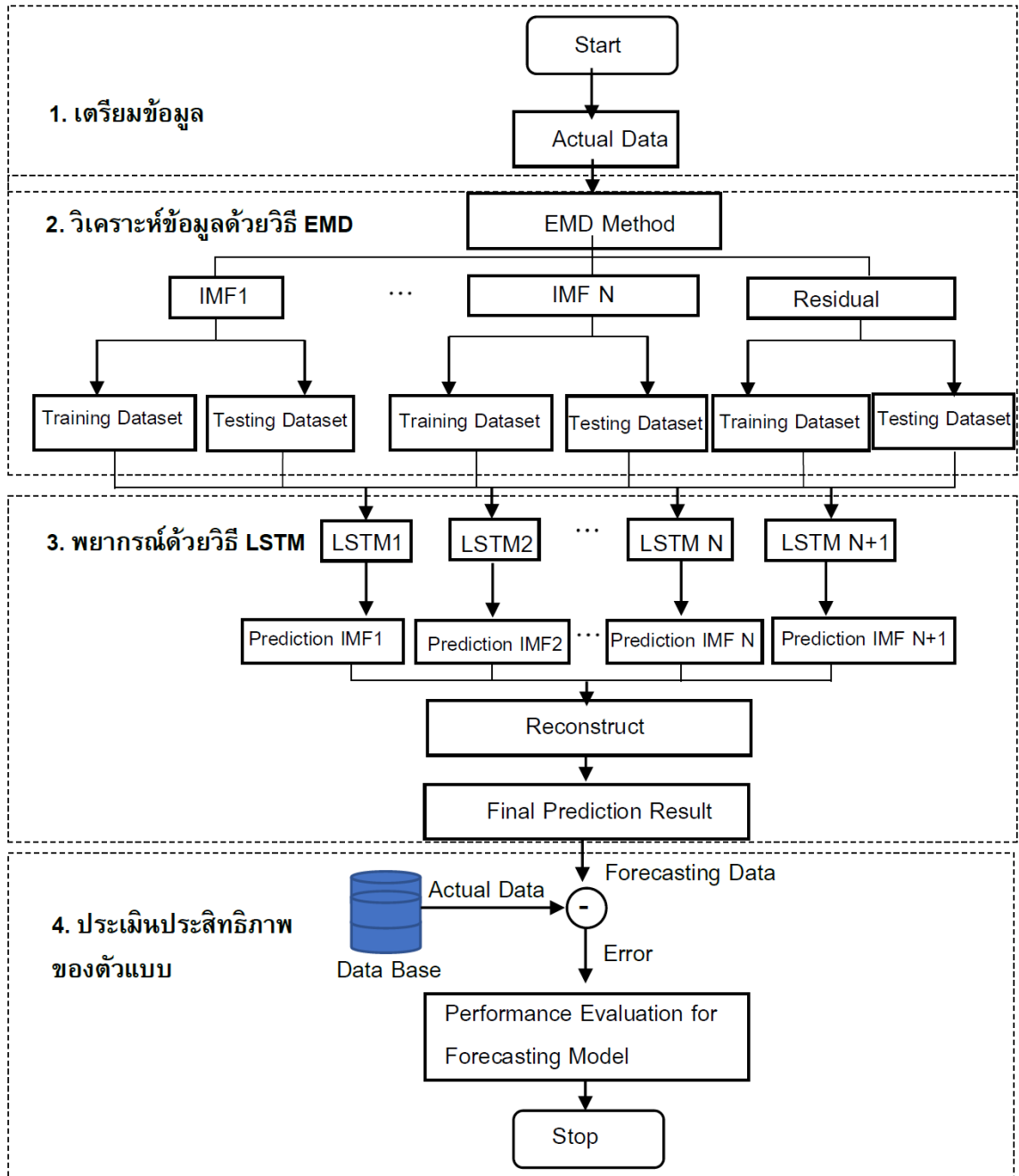
เมื่อ o_t คือ Output Gate, σ คือ ฟังก์ชัน Sigmoid, W_o คือ ค่านำหนักของ Matrices, h_{t-1} คือ ค่า Output ของ Cell State ก่อนหน้า (ที่ Time Stamp $t-1$), x_t คือ ค่า Input ที่เข้ามาใน Cell State ณ เวลา t , b_o คือ ค่า Bias

2.5 ตัวแบบผสม (Hybrid Model)

จากการวิเคราะห์ข้อมูลอนุกรมเวลาด้วยวิธี EMD เป็นวิธีการแยกข้อมูลอนุกรมเวลาออกเป็น ส่วน ๆ เพื่อลดความถี่และสัญญาณแกว่งไวของข้อมูล (Fast Oscillation) โดยแยกออกเป็น ฟังก์ชันภายใน (Intrinsic Mode Function: IMF) การสร้างตัวแบบ

ผสมจะนำองค์ประกอบที่แยกได้จากวิธี EMD มาทำการพยากรณ์ด้วยวิธี LSTM และนำค่าพยากรณ์มารวมกันในขั้นตอนสุดท้าย เรียกตัวแบบดังกล่าวว่า ตัวแบบ HybridEMD-LSTM มีขั้นตอนและวิธีการสร้างตัวแบบ แสดงดังรูปที่ 3

จากรูปที่ 3 สามารถอธิบายรายละเอียดของตัวแบบ HybridEMD -LSTM ได้ดังนี้ (1) เตรียมข้อมูลประกอบด้วยข้อมูลอนุกรมเวลาราคาฟิวเจอร์ 3 ชุด (2) วิเคราะห์ข้อมูลอนุกรมเวลาราคาฟิวเจอร์ทั้ง 3 ชุด ด้วยวิธี EMD เพื่อแยกองค์ประกอบของอนุกรมเวลาดันฉบับให้อยู่ในรูปของชุดฟังก์ชันภายใน (Intrinsic Mode Functions: IMFs) และค่าความคลาดเคลื่อนแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่ 1 ข้อมูลการเรียนรู้ (Training Dataset) และชุดที่ 2 ข้อมูลการทดสอบ (Testing Dataset) (3) นำองค์ประกอบของข้อมูลอนุกรมเวลาแต่ละส่วนจากขั้นตอนที่ 2 พยากรณ์ด้วยตัวแบบ LSTM (4) รวมค่าพยากรณ์จากขั้นตอนที่ 3 ได้ค่าพยากรณ์จากตัวแบบ HybridEMD-LSTM (5) คำนวณค่าความแม่นยำของการพยากรณ์ ด้วยมาตรวัดประสิทธิภาพของตัวแบบการพยากรณ์ MAE, RMSE และ MAPE ตามลำดับ



รูปที่ 3 กระบวนการในการพัฒนาตัวแบบ HybridEMD-LSTM



2.6 เกณฑ์การประเมินประสิทธิภาพของตัวแบบ

การประเมินประสิทธิภาพของตัวแบบ งานวิจัยนี้มีทั้งหมด 3 เกณฑ์ [27] ดังนี้

(1) ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) มีสูตรดังสมการที่ (8)

$$MAE = \sum_{t=1}^n |Y_t - \hat{Y}_t| / n \quad (8)$$

(2) รากของความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) มีสูตรดังสมการที่ (9)

$$RMSE = \sqrt{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2 / n} \quad (9)$$

(3) ร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) มีสูตรดังสมการที่ (10)

$$MAPE = 100 \times \sum_{t=1}^n |1 - \hat{Y}_t / Y_t| / n \quad (10)$$

โดยที่ค่า MAE, RMSE และ MAPE มีค่าน้อยเข้าใกล้ศูนย์ แสดงว่า ตัวแบบมีประสิทธิภาพความแม่นยำสูง

ตารางที่ 1 ค่าสถิติพื้นฐานราคาพืชไร่ 3 ชนิด

รายการ	ข้อมูล	จำนวน	ค่าเฉลี่ย	ค่าสูงสุด	ค่ามัธยฐาน	ค่าต่ำสุด	ส่วนเบี่ยงเบนมาตรฐาน
ราคามันสำปะหลัง	All sample	241	1.701	2.960	1.780	0.740	0.532
	Training Dataset	216	1.673	2.960	1.780	0.740	0.551
(หน่วย: บาท/กก.)	Testing Dataset	25	1.945	2.290	1.940	1.620	0.193
ราคาข้าวโพดเลี้ยงสัตว์	All sample	241	6.909	9.950	7.330	3.890	1.518
	Training Dataset	216	6.786	9.950	6.980	3.890	1.551
(หน่วย: บาท/กก.)	Testing Dataset	25	7.973	8.970	7.920	7.420	0.439
ราคาข้าวเปลือกเจ้าหอมมะลิ	All sample	241	7,827	13,259	7,827	4,579	1,667.590
	Training Dataset	216	7,750	13,259	7,750	4,579	1,729.277
(หน่วย : บาท/ตัน)	Testing Dataset	25	8,655	9,453	8,655	7,442	662.073

3. ผลการวิจัยและอภิปรายผล

3.1 ผลการวิเคราะห์สถิติพื้นฐานราคาพืชไร่ 3 ชนิด

(1) ราคามันสำปะหลังรายเดือน มีค่าเฉลี่ยเท่ากับ 1.701 บาทต่อกิโลกรัม ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.532 บาทต่อกิโลกรัม มีค่าสูงสุดเท่ากับ 2.960 บาทต่อกิโลกรัม และค่าต่ำสุดเท่ากับ 0.740 บาทต่อกิโลกรัม

(2) ราคาข้าวโพดเลี้ยงสัตว์ มีค่าเฉลี่ยเท่ากับ 6.909 บาทต่อกิโลกรัม ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.518 บาทต่อกิโลกรัม มีค่าสูงสุดเท่ากับ 9.950 บาทต่อกิโลกรัม และต่ำสุดเท่ากับ 3.890 ต่อกิโลกรัม

(3) ราคาข้าวเปลือกหอมมะลิ มีค่าเฉลี่ยเท่ากับ 7,827 บาทต่อตัน มีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1,667.590 บาทต่อตัน มีค่าสูงสุดเท่ากับ 13,259 บาทต่อตัน และค่าต่ำสุดเท่ากับ 4,579 บาทต่อตัน รายละเอียดดังตารางที่ 1



3.2 ผลการพยากรณ์ด้วยตัวแบบ ANN ตัวแบบ LSTM และตัวแบบ HybridEMD-LSTM

การพยากรณ์ข้อมูลอนุกรมเวลาราคาฟิวเจอร์ทั้ง 3 ชุด ผู้วิจัยทำการวิเคราะห์ข้อมูล โดยเรียงลำดับแต่ละตัวแบบมีรายละเอียด ดังนี้

3.2.1 การกำหนดตัวแบบการพยากรณ์แบบ ANN ผู้วิจัยประยุกต์ใช้โครงข่ายประสาทเทียมแบบเพอร์เซ็ปตรอนหลายชั้น (Multi - Layer Perceptron) ซึ่งโปรแกรมภาษาอาร์มีฟังก์ชัน MLP ซึ่งอยู่ใน Package “nnfor” สำหรับกำหนดข้อมูลนำเข้า (Input) และชั้นซ่อน (Hidden) และใช้ Sigmoid Logistic Function เป็น ฟังก์ชันกระตุ้น การกำหนดค่าพารามิเตอร์ของโครงข่ายประสาทเทียมในการเรียนรู้ในครั้งนี้เป็นแบบแพร่กลับ (Backpropagation Learning) และทำการปรับปรุงองคกอลิเทียมด้วยฟังก์ชัน Levenberg- Marquardt เพื่อให้การปรับค่าน้ำหนักและค่าเอนเอียงในเส้นเชื่อมแต่ละชั้นของโครงข่ายให้มีความรวดเร็วและได้ค่าที่มีความเหมาะสม จากนั้นทำการปรับค่าผลรวมของนิวรอนในชั้นซ่อนด้วยฟังก์ชันถ้ายอน คือ ฟังก์ชันซิกมอยด์แบบโลจิสติกส์ (Sigmoid Logistic Function) มีผลลัพธ์อยู่ในช่วง [-1,1] มีอัตราการเรียนรู้ (Learning Rate) เท่ากับ 0.0001 จำนวนรอบในการฝึกฝนการเรียนรู้ (Epoch) สูงสุด 1,000 รอบ และใช้ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error : MSE) เป็นเกณฑ์ในการเลือกนิวรอนในชั้นซ่อนที่เหมาะสมให้กับโครงข่ายการพยากรณ์ด้วย ตัวแบบ ANN จะใช้ข้อมูลชุดที่ 1 ข้อมูลฝึกฝน (Training Dataset) จำนวน 216 ค่าสังเกต สำหรับสร้างตัวแบบ และชุดที่ 2 ข้อมูลการทดสอบ (Testing Dataset) จำนวน 25 ค่าสังเกต

เพื่อให้ครอบคลุมจำนวนคาบของข้อมูลใน 1 ฤดูกาล ทำให้ตัวแบบมีความแกร่งและมีความแม่นยำในการพยากรณ์ การวิจัยครั้งนี้ผู้วิจัยได้กำหนดรูปแบบโครงสร้างสถาปัตยกรรมการเรียนรู้สำหรับตัวแบบ ANN ของข้อมูลแต่ละชุดดังนี้คือ ชุดที่ 1 ราคาเงินสำหรับปั๊ม มี Number of Input Nodes เท่ากับ 17, Hidden Layer Node เท่ากับ 5 และ Output Nodes เท่ากับ 1 โดยมีค่า MSE เท่ากับ 0.0011 ชุดที่ 2 ราคาข้าวโพด มี Number of Input Nodes เท่ากับ 13, Hidden Layer Node เท่ากับ 5 และ Output Nodes เท่ากับ 1 โดยมีค่า MSE เท่ากับ 0.0674 ชุดที่ 3 ราคาข้าวเปลือกเจ้าหอมมะลิ ข้าวโพด มี Number of Input Nodes เท่ากับ 5, Hidden Layer Node เท่ากับ 5 และ Output Nodes เท่ากับ 1 โดยมีค่า MSE เท่ากับ 0.0489 ผลการพยากรณ์และค่าการตรวจสอบความแม่นยำของค่าพยากรณ์ทั้ง 3 ชุด แสดงดังตารางที่ 2 และตารางที่ 3

3.2.2 การพยากรณ์ด้วยตัวแบบ LSTM ผู้วิจัยประยุกต์ใช้โปรแกรมภาษาไพธอน โดยแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่ 1 ข้อมูลการเรียนรู้ (Training Dataset) และ ชุดที่ 2 ข้อมูลทดสอบ (Testing Dataset) จากนั้นแบ่งข้อมูลชุดที่ 1 ข้อมูลการเรียนรู้ และชุดข้อมูลการทดสอบออกเป็นข้อมูลที่ใช้ในการ Input และ Output ของแบบจำลอง LSTM โดยแปลงค่าชุดข้อมูลให้อยู่ในรูปแบบของอาร์เรย์ 3 มิติ ได้แก่ Sample, Timesteps, Feature และแปลงข้อมูลให้อยู่ในรูปแบบ Normalize โดยปรับค่าให้อยู่ระหว่าง 0 ถึง 1 และนำข้อมูลไปสร้างตัวแบบ LSTM การพยากรณ์ด้วยตัวแบบ LSTM จะต้องมีการกำหนด Epochs Hidden Neuron และ Batch Size ที่เหมาะสม



งานวิจัยนี้กำหนด Epochs = 20, Batch Size = 1, Step = 1, Activation = Relu โดยกำหนดค่าของ Optimizer = Adam ทั้งสามชุดข้อมูล เนื่องจากทดลองใช้กับข้อมูล พบว่า มีค่าความแม่นยำมากที่สุด เมื่อได้ค่าพารามิเตอร์ที่เหมาะสม นำค่าดังกล่าวมา กำหนดตัวแบบและพยากรณ์ล่วงหน้า เมื่อได้ค่าพยากรณ์ทำการแปลงค่ากลับให้มีหน่วยเหมือนกับข้อมูลจริง ผลการพยากรณ์และค่าการตรวจสอบความแม่นยำของตัวแบบ แสดงดังตารางที่ 2 และ 3

3.2.3 การพยากรณ์ด้วยตัวแบบ HybridEMD-LSTM ที่ผู้วิจัยประยุกต์ใช้โปรแกรมภาษาไพธอน และดำเนินการตามขั้นตอนการสร้างตัวแบบผสม โดยการเตรียมข้อมูลอนุกรมเวลาและทำการวิเคราะห์ข้อมูลอนุกรมเวลาราคาพีซีไทร์ทั้ง 3 ชุด ด้วยวิธี EMD เพื่อแยกองค์ประกอบอนุกรมเวลาดัดฉบับให้อยู่ในรูปของชุดฟังก์ชันภายใน (Intrinsic Mode Functions : IMFs) และค่าความคลาดเคลื่อน โดยจะแสดงตัวอย่าง

การวิเคราะห์เพียง 1 ชุดข้อมูล ดังรูปที่ 4 (ข้อมูลชุดที่ 1 ราคามันสำปะหลัง) จากนั้นแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดที่ 1 ข้อมูลการเรียนรู้ (Training Dataset) และชุดที่ 2 ข้อมูลการทดสอบ (Testing Dataset) (ข้อมูลราคาข้าวโพดเลี้ยงสัตว์ และข้อมูลราคาข้าวเปลือกเจ้าหอมมะลิ สามารถทำได้ในลักษณะเดียวกัน) จากนั้นทำการพยากรณ์แต่ละองค์ประกอบ (IMFs) ด้วยตัวแบบ LSTM โดยกำหนด Epochs= 20, Batch Size =1, Step = 1, Activation = Relu และ Optimizer = Adam (ดังรูปที่ 5) โดยกำหนดให้มีค่าเท่ากันทั้ง 3 ชุด เนื่องจากทดลองใช้กับข้อมูล พบว่า มีค่าความแม่นยำมากที่สุด เมื่อได้ค่าพารามิเตอร์ที่เหมาะสม นำค่าดังกล่าวมา กำหนดตัวแบบและพยากรณ์ล่วงหน้า เมื่อได้ค่าพยากรณ์ทำการแปลงค่ากลับให้มีหน่วยเหมือนกับข้อมูลจริง ผลการพยากรณ์และค่าการตรวจสอบความแม่นยำของตัวแบบแสดงดังตารางที่ 2 และตารางที่ 3

ตารางที่ 2 ค่าจริงและค่าพยากรณ์ของตัวแบบ ANN ตัวแบบ LSTM และตัวแบบ HybridEMD-LSTM ของข้อมูลอนุกรมเวลาราคาพีซีไทร์ 3 ชุด

ข้อมูลจริง	ราคามันสำปะหลัง			ข้อมูลจริง	ราคาข้าวโพดเลี้ยงสัตว์		
	ANN	LSTM	HybridEMD-LSTM		ANN	LSTM	HybridEMD-LSTM
1.94	1.94	1.92	1.94	7.81	8.34	7.79	7.79
1.92	1.83	1.92	1.97	7.46	8.72	7.91	7.69
1.89	1.76	1.94	1.91	7.42	8.75	7.54	7.60
1.73	1.79	1.87	1.81	7.49	9.02	7.50	7.59
1.67	1.74	1.71	1.70	7.72	8.79	7.58	7.65
1.62	1.71	1.65	1.63	7.92	8.91	7.82	7.79
1.64	1.62	1.6	1.62	8.12	8.53	8.03	7.90



ตารางที่ 2 ค่าจริงและค่าพยากรณ์ของตัวแบบ ANN ตัวแบบ LSTM และตัวแบบ HybridEMD-LSTM ของข้อมูลอนุกรมเวลาราคาพีซีไทร์ 3 ชุด (ต่อ)

ข้อมูลจริง	ราคามันสำปะหลัง			ข้อมูลจริง	ราคาข้าวโพดเลี้ยงสัตว์		
	ANN	LSTM	HybridEMD-LSTM		ANN	LSTM	HybridEMD-LSTM
1.73	1.74	1.62	1.71	7.49	8.24	8.25	7.90
1.76	1.73	1.71	1.74	7.6	7.94	7.58	7.68
1.74	1.52	1.74	1.73	7.44	7.79	7.69	7.59
1.85	1.68	1.72	1.84	7.68	8.07	7.52	7.57
2.05	1.81	1.83	1.98	7.9	8.20	7.78	7.70
2.03	1.71	2.04	1.99	8.15	8.34	8.01	7.89
2.07	1.65	2.01	2.04	8.15	8.41	8.28	8.05
2.13	1.71	2.06	2.07	8.06	8.39	8.28	8.08
2.08	1.66	2.12	2.07	7.93	8.69	8.18	8.04
1.92	1.59	2.07	2.02	7.69	8.84	8.04	7.98
1.91	1.34	1.9	1.95	8.07	8.91	7.79	7.93
1.97	1.24	1.89	1.93	8.29	8.60	8.19	8.02
2.02	1.24	1.95	1.96	7.72	8.37	8.43	8.08
2.12	1.19	2	2.02	8.23	7.91	7.82	8.04
2.08	1.22	2.11	2.06	8.4	7.72	8.37	8.21
2.19	1.26	2.07	2.1	8.76	7.81	8.56	8.41
2.28	1.34	2.19	2.16	8.97	8.21	8.96	8.62
2.29	1.33	2.29	2.20	8.85	8.31	9.21	8.70



ตารางที่ 2 ค่าจริงและค่าพยากรณ์ของตัวแบบ ANN ตัวแบบ LSTM และตัวแบบ HybridEMD-LSTM ของข้อมูลอนุกรมเวลาราคาพีซีไร์ 3 ชุด (ต่อ)

ราคาข้าวเปลือกเจ้าหอมมะลิ			
ข้อมูลจริง	ANN	LSTM	HybridEMD-LSTM
7,756	7,701.58	7,556.22	7,753.22
8,166	7,706.02	7,802.41	7,926.32
8,655	7,703.92	8,194.02	8,347.37
9,453	7,736.45	8,673.45	8,752.46
9,319	7,721.09	9,484.68	8,930.10
9,052	7,725.20	9,345.95	9,053.33
8,910	7,742.33	9,072.55	9,047.00
9,297	7,740.51	8,928.78	8,973.33
9,419	7,747.43	9,323.28	9,306.37
8,804	7,749.00	9,449.38	8,929.94
8,281	7,747.64	8,822.20	8,640.52
8,408	7,752.49	8,305.56	8,425.55
9,001	7,754.28	8,429.60	8,464.21
9,148	7,755.97	9,020.79	8,684.39
9,312	7,758.69	9,170.39	8,787.12
9,068	7,760.41	9,338.74	8,857.95
8,742	7,762.46	9,088.83	8,716.23
8,474	7,765.05	8,760.15	8,485.86
8,257	7,767.04	8,494.42	8,220.59
7,508	7,769.08	8,282.22	7,941.69
7,442	7,771.27	7,570.11	7,535.46
7,586	7,773.21	7,508.87	7,417.88
7,693	7,775.30	7,642.80	7,471.66
7,788	7,777.41	7,743.07	7,561.04
7,934	7,779.42	7,832.64	7,836.76



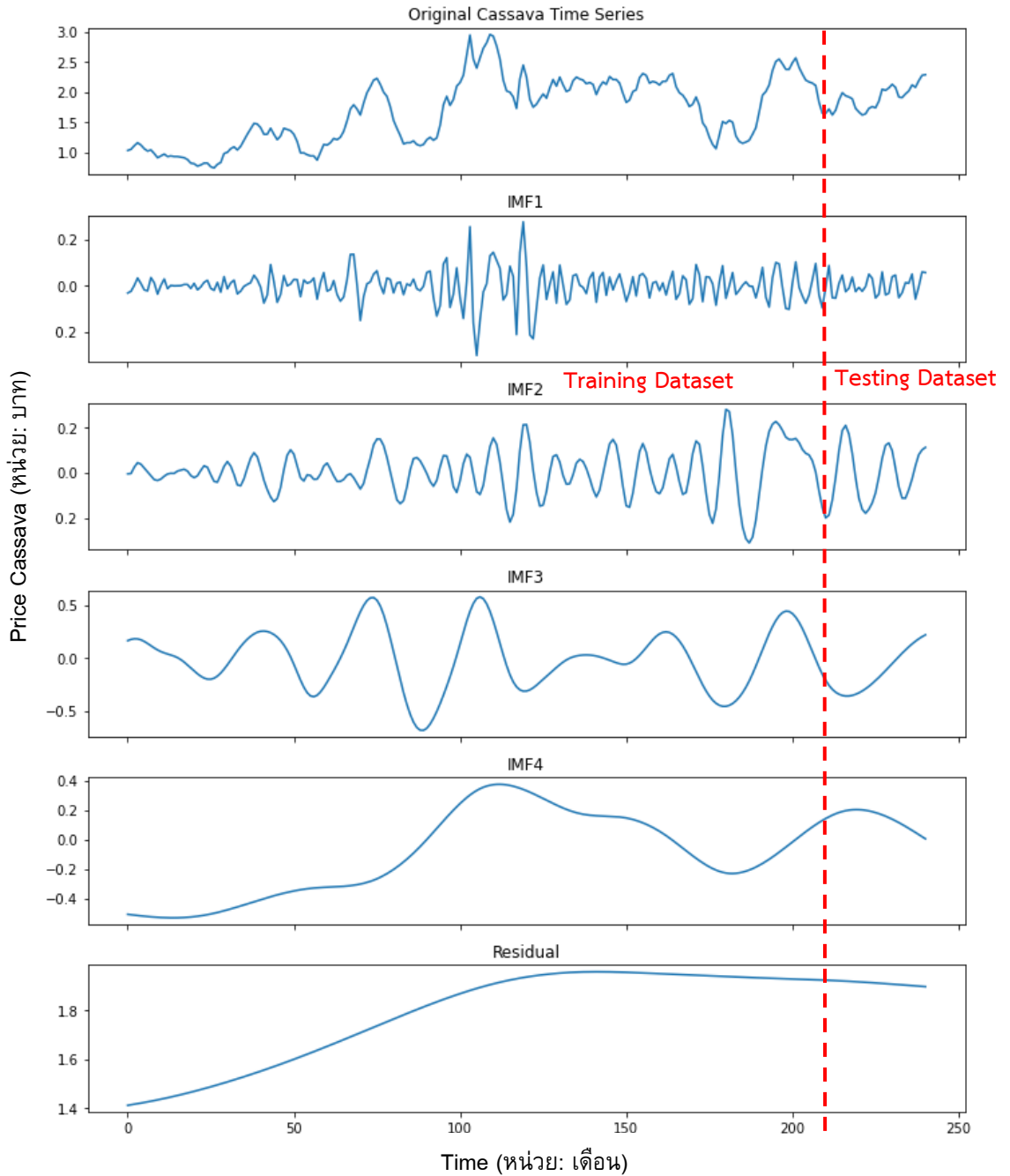
ตารางที่ 3 เปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์ข้อมูลอนุกรมเวลาราคาฟิวเจอร์ 3 ชุด

ข้อมูล	ตัวแบบ	MAE	RMSE	MAPE
ราคามันสำปะหลัง	ANN	0.3896	0.5161	0.1885
	LSTM	0.0672	0.0868	0.0346
	HybridEMD-LSTM	0.0456	0.0568	0.0227
ราคาข้าวโพดเลี้ยงสัตว์	ANN	0.6796	0.7743	0.0863
	LSTM	0.2172	0.2923	0.0276
	HybridEMD-LSTM	0.1872	0.2146	0.0234
ราคาข้าวเปลือกเจ้าหอมมะลิ	ANN	857.9788	1026.5670	0.0961
	LSTM	293.454	364.8786	0.0342
	HybridEMD-LSTM	229.894	298.1984	0.0265

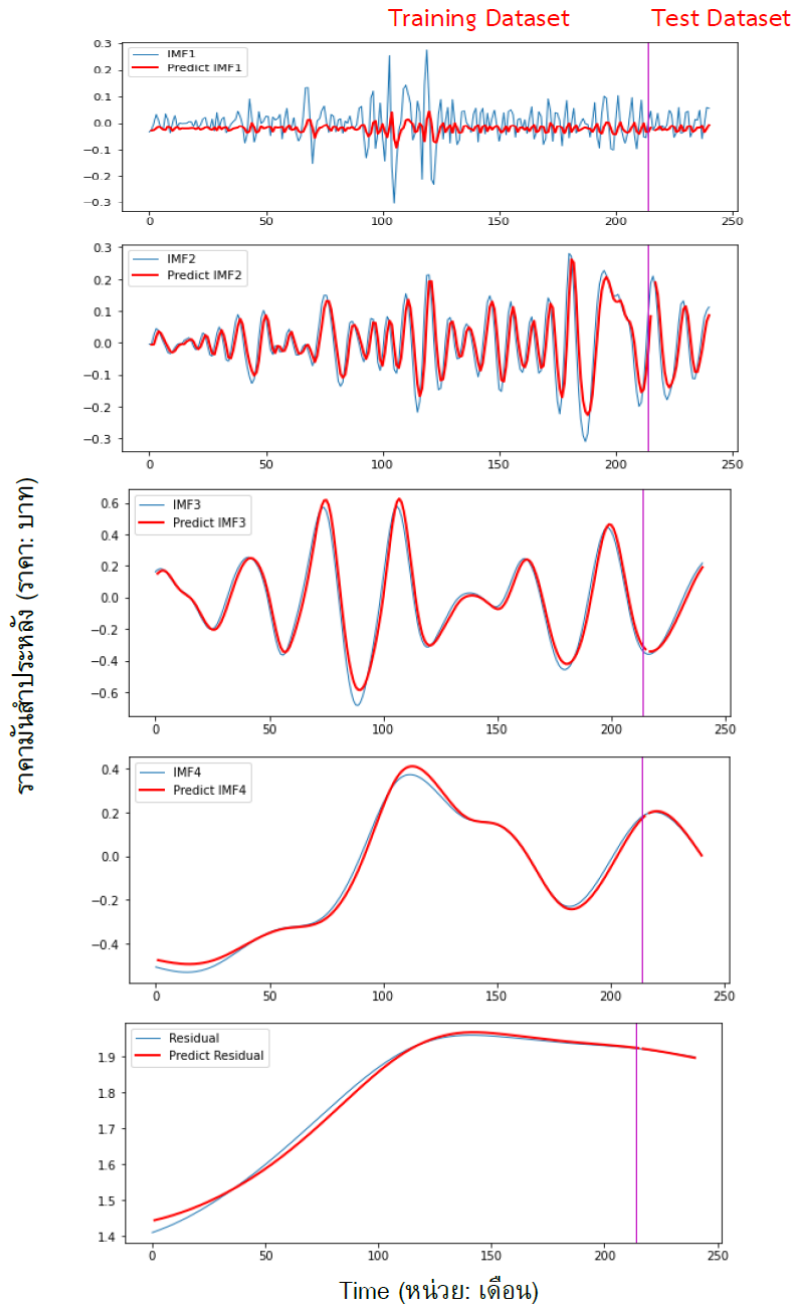
3.3 อภิปรายผลการวิจัย

การศึกษาค้นคว้าครั้งนี้ พบว่า วิธีการพยากรณ์ด้วยตัวแบบผสมระหว่างเทคนิควิธีการแยกองค์ประกอบของข้อมูลอนุกรมเวลาเชิงประจักษ์ (Empirical Mode Decomposition :EMD) กับวิธีการพยากรณ์โครงข่ายประสาทเทียมเชิงลึก (Long Short Term Memory : LSTM) เรียกว่า ตัวแบบ Hybrid EMD-LSTM มีความแม่นยำในการพยากรณ์ราคาฟิวเจอร์ทั้ง 3 ชุด มากกว่าตัวแบบโครงข่ายประสาทเทียม (Artificial Neural Network Model : ANN) และตัวแบบ LSTM ในทุกเกณฑ์การประเมินประสิทธิภาพ (ตารางที่ 3) ซึ่งให้ผลสอดคล้องกับการศึกษาที่ผ่านมา [14-21, 27] ทั้งนี้เนื่องจากตัวแบบ Hybrid EMD-LSTM ได้ประยุกต์ใช้วิธี EMD ซึ่งเป็นวิธีการแบ่งข้อมูลอนุกรมเวลาออกเป็น ส่วน ๆ เรียกว่า ฟังก์ชันภายใน (Intrinsic Mode Function: IMF) เป็นการลดสัญญาณแกว่งไวของข้อมูล (Fast Oscillation) และนำองค์ประกอบฟังก์ชันภายในไปพยากรณ์ด้วยตัวแบบ LSTM จากนั้นนำค่าการพยากรณ์แต่ละองค์ประกอบมารวมเป็นค่าพยากรณ์ในขั้นตอน

สุดท้าย ทำให้ได้ค่าพยากรณ์ที่มีความแม่นยำสูง ในส่วนของการประยุกต์ใช้งาน พบว่า ตัวแบบ ANN และตัวแบบ LSTM เป็นทางเลือกที่น่าสนใจเนื่องจากเป็นตัวแบบที่มีความซับซ้อนน้อยกว่าตัวแบบผสม ถ้าผู้ใช้งานต้องการตัวแบบพยากรณ์ที่มีความแม่นยำสูงใช้ได้ดีกับข้อมูลอนุกรมเวลาที่มีความผันผวนสูง แนะนำควรเลือกใช้ตัวแบบ Hybrid EMD-LSTM เพราะจะทำให้ได้ค่าพยากรณ์ที่มีความแม่นยำสูง การวิจัยนี้มุ่งศึกษาและนำเสนอตัวแบบพยากรณ์ข้อมูลอนุกรมเวลาราคาฟิวเจอร์ 3 ชุด ได้แก่ ราคามันสำปะหลัง ราคาข้าวโพดเลี้ยงสัตว์ และราคาข้าวเปลือกเจ้าหอมมะลิ ซึ่งเป็นทางเลือกหนึ่งที่สำคัญสำหรับหน่วยงานที่เกี่ยวข้องสามารถนำตัวแบบดังกล่าวนี้ ไปใช้สำหรับการทำนายราคาฟิวเจอร์ทั้ง 3 ชนิด จากนั้นนำเสนอสารสนเทศที่ได้นำมากำหนดเป็นแนวทางในการวางแผนหรือกำหนดเป็นกลยุทธ์เชิงนโยบายสำหรับแนะนำเกษตรกรเพื่อผลิตฟิวเจอร์ทั้ง 3 ชนิดนี้ การใช้งานอาจนำตัวแบบผสมที่พัฒนาขึ้นพัฒนาเป็นแอปพลิเคชันเพื่อให้ง่ายสำหรับการใช้งานต่อไป



รูปที่ 4 องค์ประกอบ (IMFs) ข้อมูลอนุกรมเวลาราคามันสำปะหลัง



รูปที่ 5 ค่าจริงและค่าพยากรณ์แต่ละองค์ประกอบ (IMFs) ข้อมูลอนุกรมเวลาราคามันสำปะหลังด้วยตัวแบบ HybridEMD-LSTM



4. บทสรุป

การศึกษาครั้งนี้ ผู้วิจัยนำเสนอวิธีการสร้างและคัดเลือกตัวแบบการพยากรณ์ที่เหมาะสมกับข้อมูลอนุกรมเวลาราคาพืชไร่ของประเทศไทย จำนวน 3 ชุด ได้แก่ ราคามันสำปะหลัง ราคาข้าวโพดเลี้ยงสัตว์และราคาข้าวเปลือกเจ้าหอมมะลิ ตั้งแต่เดือนมกราคม พ.ศ. 2545 ถึงเดือนมกราคม พ.ศ. 2565 จำนวน 241 ค่า จากเว็บไซต์ของสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ด้วยตัวแบบ ANN ตัวแบบ LSTM และตัวแบบ HybridEMD-LSTM ผลการวิจัย พบว่า ตัวแบบ HybridEMD-LSTM มีความแม่นยำมากกว่าตัวแบบ ANN และตัวแบบ LSTM ดังนั้นสรุปได้ว่า ตัวแบบผสมที่ทีมผู้วิจัยนำเสนอมีความเหมาะสมที่จะนำไปพยากรณ์ข้อมูลอนุกรมเวลาราคาพืชไร่ทั้ง 3 ชุด

5. กิตติกรรมประกาศ

ขอขอบคุณสาขาวิชาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏอุบลราชธานี ที่อนุเคราะห์สถานที่สำหรับการวิจัย และขอขอบคุณสำนักงานเศรษฐกิจการเกษตร กระทรวงเกษตรและสหกรณ์ ที่อนุเคราะห์ข้อมูลที่ใช้ในการทดลองวิจัยครั้งนี้

6. เอกสารอ้างอิง

- [1] K. Thongmeethip, Agricultural development in Thailand in terms of community development and quality of life, PSDS Journal of Development Studies, Puey Ungphakorn School of Development Studies, Thammasat University, 2021, 4(1), 132-162. (in Thai)
- [2] <http://www.nso.go.th> (Accessed on 20 March 2022)
- [3] <https://data.worldbank.org> (Accessed on 20 March 2022)
- [4] Office of agricultural economics, Agricultural statistics of Thailand 2020, Office of agricultural economics, Ministry of Agriculture and Cooperatives, Bangkok, Thailand, 2021.
- [5] <https://www.oae.go.th> (Accessed on 21 March 2022)
- [6] W. Riansut and K. Thongrit, Forecasting the price of field corn in Thailand, RUMTP Research Journal, 2017, 11(1), 1-14. (in Thai)
- [7] S. Deepradit and P. Rukson, The forecasting techniques comparison of field crops in Thailand, The Journal of Industrial Technology, 2021, 17(3), 214 – 231. (in Thai)
- [8] T. Xiong, C. Li, and Y. Bao, Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China, Neurocomputing, 2018, 275, 2831-2844.
- [9] T. Iizumi, Y. Shin, W. Kim, M. Kim and J. Choi, Global crop yield forecasting using seasonal climate information from a multi-model ensemble, Climate Services, 2018, 11, 13-23.



- [10] S.Y. Han, T.F.A. Bishop and P. Filippi, Data-driven, early-season forecasts of block sugarcane yield for precision agriculture, *Field Crops Research*, 2022, 276, 108360.
- [11] T. Sujjaviriyasup, Artificial neural network model for forecasting monthly price of maize in Thailand, *Srinakharinwirot Science Journal*, 2018, 34(1), 91-107. (in Thai)
- [12] H. Yin, D. Jin, Y.H. Gu, C.J. Park, S.K. Han and S.J. Yoo, STL-ATTLSTM: Vegetable price forecasting using STL and attention mechanism-based LSTM, *Agriculture*, 2020, 10(12), 612.
- [13] N. Filipović, S. Brdar, G. Mimić, O. Marko and V. Crnojević, Regional soil moisture prediction system based on Long Short-Term Memory network, *Biosystems Engineering*, 2022, 213, 30-38.
- [14] J. Zhang, Y. Meng, J. Wei, J. Chen and J. Qin, A novel hybrid deep learning model for sugar price forecasting based on time series decomposition, *Mathematical Problems in Engineering*, 2021, 6507688.
- [15] L. Yu, S. Liang, R. Chen and K.K. Lai, Predicting monthly biofuel production using a hybrid ensemble forecasting methodology, *International Journal of Forecasting*, 2022, 38(1), 3-20.
- [16] H. Rezaei, H. Faaljou and G. Mansourfar, Stock price prediction using deep learning and frequency decomposition, *Expert Systems with Applications*, 2021, 169, 114332.
- [17] B. Gao, X. Huang, J. Shi, Y. Tai and J. Zhang, Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks, *Renewable Energy*, 2020, 162, 1665-1683.
- [18] W. Liu, W.D. Liu and J. Gu, Forecasting oil production using ensemble empirical model decomposition based Long Short-Term Memory neural network, *Journal of Petroleum Science and Engineering*, 2020, 189, 107013.
- [19] Y.-X. Wu, Q.-B. Wu and J.-Q. Zhu. Improved EEMD-based crude oil price forecasting using LSTM networks, *Physica A: Statistical Mechanics and Its Applications*, 2019, 516, 114-124.
- [20] X. Yan, Y. Liu, Y. Xu and M. Jia, Multistep forecasting for diurnal wind speed based on hybrid deep learning model with improved singular spectrum decomposition, *Energy Conversion and Management*, 2020, 225, 113456.
- [21] J. Cao, Z. Li and J. Li, Financial time series forecasting model based on CEEMDAN and LSTM, *Physica A: Statistical Mechanics and Its Applications*, 2019, 519, 127-139.



- [22] Y. Dong, C. Zhang, M. Niu, S. Wang and S. Sun, Air pollution forecasting with multivariate interval decomposition ensemble approach, *Atmospheric Pollution Research*, 2021, 12(12), 101230.
- [23] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung and H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 1998, 454(1971), 903–995.
- [24] N. Kulvanich, Improving prediction accuracy of time series data using ARIMA-ANN hybrid model and regression analysis, *Thai Science and Technology Journal*, 2020, 28(12), 2101-2112. (in Thai)
- [25] M. Niu, S. Sun, J. Wu, L. Yu and J. Wang, An innovative integrated model using the singular spectrum analysis and nonlinear multi-layer perceptron network optimized by hybrid intelligent algorithm for short-term load forecasting, *Applied Mathematical Modelling*, 2016, 40(5-6), 4079-4093.
- [26] G.P. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing*, 2003, 50, 159-175.
- [27] Y. Lin, Y. Yan, J. Xu, Y. Liao, and F. Ma, Forecasting stock index price using the CEEMDAN-LSTM model, *The North American Journal of Economics and Finance*, 2021, 57, 101421.
- [28] <http://colah.github.io/posts/2015-08-Understanding-LSTMs/> (Accessed on 18 March 2022)