

**แบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนพุนลอยน้ำ โดยอาศัยวิธีการ
หน่วยความจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตติ้ง
Floating Solar Power Generation Forecasting Model Based On
Long Short-Term Memory And Extreme Gradient Boosting Methods**

อดิเรก ปันบุญ¹ และ วลัยยา ร่มสายหยุด²

¹แขนงวิชาเทคโนโลยีดิจิทัล สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช,

E-Mail adirek.pb@gmail.com

²สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช, E-Mail walisa.rom@stou.ac.th

Adirek Panboon¹ and Walisa Romsaiyud²

¹Digital Technology, School of Science and Technology, Sukhothai Thammathirat Open University,

E-Mail adirek.pb@gmail.com

²School of Science and Technology, Sukhothai Thammathirat Open University,

E-Mail walisa.rom@stou.ac.th

วันที่รับบทความ 28 พฤษภาคม 2568

วันที่แก้ไขบทความ 29 พฤษภาคม 2568

วันที่ตอบรับบทความ 25 มิถุนายน 2568

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ (1) ศึกษาวิธีการและสร้างแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนพุนลอยน้ำ โดยอาศัยวิธีการหน่วยความจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตติ้ง และ (2) ประเมินประสิทธิภาพของแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนพุนลอยน้ำ โดยอาศัยวิธีการหน่วยความจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตติ้ง วิธีการวิจัยดำเนินการตามกระบวนการเรียนรู้เชิงลึก (Deep Learning) ซึ่งประกอบด้วย 3 ขั้นตอนหลัก ดังนี้ ขั้นตอนแรก เก็บรวบรวมข้อมูลการผลิตพลังงานไฟฟ้าและข้อมูลจากเซนเซอร์ของโรงไฟฟ้าพลังงานแสงอาทิตย์บนพุนลอยน้ำ ขนาดกำลังการผลิต 45 เมกะวัตต์ ในช่วงเดือนกุมภาพันธ์ ถึง ตุลาคม พ.ศ. 2566 จำนวน 6,511 ระเบียบ และ 11 คุณลักษณะ ขั้นตอนที่สอง การคัดเลือกคุณลักษณะและการจำแนกประเภท โดยประยุกต์ใช้วิธีการแบบผสมผสานด้วยอัลกอริทึมการเรียนรู้เชิงลึก 2 อัลกอริทึม ได้แก่ 1) หน่วยความจำระยะสั้นแบบยาว (LSTM) สำหรับการจดจำข้อมูลในระยะยาวตามอนุกรมเวลา และ 2) เอ็กซ์ตรีมเกรเดียนต์บูตติ้ง (XGBoost) สำหรับการเรียนรู้จากชุดข้อมูลที่ไม่น่าสนใจและการพยากรณ์ที่มีประสิทธิภาพสูง และขั้นตอนที่สาม การประเมินประสิทธิภาพของแบบจำลองโดยใช้ตัวชี้วัด ได้แก่ การวัดค่าเฉลี่ยของผลต่างสัมบูรณ์ (MAE) ค่าเฉลี่ยความผิดพลาดกำลังสอง (MSE) และรากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (RMSE)

สำหรับค่าผลลัพธ์ที่ได้จากแบบจำลองการพยากรณ์ ได้ค่า MAE เฉลี่ย 0.0577, MSE เฉลี่ย 0.0143 และ RMSE เฉลี่ย 0.1196 ซึ่งแสดงถึงค่าที่เหมาะสมในสถานการณ์จริง

คำสำคัญ: การพยากรณ์, พลังงานแสงอาทิตย์บนทุ่นลอยน้ำ, หน่วยความจำระยะสั้นแบบยาว, เอ็กซ์ตรีมเกรเดียนต์บูสติง, วิธีการพยากรณ์แบบผสม

ABSTRACT

The purposes of this research were to (1) build a forecasting model for floating solar energy generation based on Long Short-Term Memory and Extreme Gradient Boosting methods and (2) to evaluate the performance of a model based on Long Short-Term Memory and Extreme Gradient Boosting methods. The research methodology follows the deep learning pipeline, which consists of three main steps: In the first step, the data were collected from the electricity generation records and sensor readings obtained from a floating solar power facility with a capacity of 45 megawatts during February–October 2023, which consisted of 6,511 examples and 11 features. In the second step, feature extraction and classification were applied using the hybrid model by using two deep learning algorithms: 1) Long Short-Term Memory (LSTM) for memorizing long-term dependencies of time-series data, and 2) Extreme Gradient Boosting (XGBoost) for learning from uncertain data and predicting high performance; and In the third step, model evaluation was assessed using metrics including the mean absolute error (MAE), mean square error (MSE) and root mean square error (RMSE) for the indicated values from the forecasting model.

The experiment result shows an average MAE of 0.0577, an average MSE of 0.0143, and an average RMSE of 0.1196 that represent suitable values in a real situation.

KEYWORDS: Forecasting, Floating Solar Power, Long Short-Term Memory, Extreme Gradient Boosting, Hybrid Forecasting Method

1. บทนำ

ประเทศไทยใช้ก๊าซธรรมชาติเป็นเชื้อเพลิงหลักในการผลิตพลังงานไฟฟ้า รองลงมาคือถ่านหิน ซึ่งเชื้อเพลิงจากฟอสซิลเป็นพลังงานที่มีอยู่อย่างจำกัด หากใช้อย่างต่อเนื่อง อาจไม่เพียงพอต่อความต้องการที่กำลังเพิ่มขึ้นในอนาคต แม้ว่าการใช้เชื้อเพลิงฟอสซิลยังคงเป็นที่ต้องการในปัจจุบัน แต่กลับส่งผลต่อการเปลี่ยนแปลงสภาพภูมิอากาศ เกิดเป็นภาวะโลกร้อน จากการปล่อยก๊าซเรือนกระจกเข้าสู่ชั้นบรรยากาศ ในปี พ.ศ. 2564 ประเทศไทยได้แสดงเจตนารมณ์ในการมีส่วนร่วมลดการปล่อยก๊าซเรือนกระจก ต่อที่ประชุมรัฐภาคีสมาชิกรายหนึ่งของอนุสัญญาสหประชาชาติว่าด้วยการเปลี่ยนแปลงสภาพภูมิอากาศ หรือ COP26 เพื่อยกระดับการแก้ปัญหาภูมิอากาศ โดยมีเป้าหมายความเป็นกลางทางคาร์บอน (Carbon Neutrality) ภายในปี พ.ศ. 2593 การปล่อยก๊าซเรือนกระจกสุทธิเป็นศูนย์ (Net Zero Emissions) ในปี พ.ศ. 2608 และยกระดับเป้าหมายการลดก๊าซเรือนกระจก (Nationally Determined Contribution) ภายในปี พ.ศ. 2573 มีการจัดตั้งโมเดลเศรษฐกิจสู่การพัฒนาที่ยั่งยืน (Bio-Circular-Green Economy: BCG) ปี พ.ศ. 2562 เพื่อส่งเสริมเศรษฐกิจหมุนเวียน สนับสนุนการใช้พลังงานทดแทนและพลังงานทางเลือก ในปัจจุบันประเทศไทยมีการสนับสนุนการใช้พลังงานหมุนเวียน (renewable energy) ในการผลิตพลังงานไฟฟ้า เช่น พลังงานน้ำ พลังงานแสงอาทิตย์ พลังงานลม พลังงานความร้อนใต้พิภพ และพลังงานชีวมวล เป็นต้น (กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน กระทรวงพลังงาน, 2564)

จากศักยภาพพื้นที่ของประเทศไทย ที่มีปริมาณความเข้มแสงอาทิตย์สูงเพียงพอ ทำให้การก่อสร้างและพัฒนาโรงไฟฟ้าพลังงานหมุนเวียนด้วยพลังงานแสงอาทิตย์เติบโตอย่างรวดเร็ว สอดคล้องกับแผนพัฒนาพลังงาน ทดแทนและพลังงานทางเลือก พ.ศ. 2561 – 2580 (AEDP2018) และแผนพัฒนากำลังผลิตไฟฟ้าของประเทศไทย พ.ศ. 2561 – 2580 ฉบับ

ปรับปรุงครั้งที่ 1 (PDP2018 Revision 1) ที่มุ่งเน้นการพัฒนาที่ยั่งยืน สนับสนุนการผลิตและการใช้พลังงานทดแทนตามศักยภาพของแหล่งเชื้อเพลิงในพื้นที่ โดยมีเป้าหมายการผลิตพลังงานไฟฟ้าจากพลังงานแสงอาทิตย์ประเภทติดตั้งบนพื้นดิน (solar farm) ติดตั้งบนหลังคา (solar rooftop) และติดตั้งบนทุ่นลอยน้ำ (floating solar) รวมทั้งหมด 14,864 เมกะวัตต์ ภายในปี พ.ศ. 2580 (สำนักงานนโยบายและแผนพลังงาน กระทรวงพลังงาน, 2561)

การผลิตไฟฟ้าจากพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ หรือ โซลาร์เซลล์ลอยน้ำ (floating solar) เป็นแผงโซลาร์เซลล์ที่ติดตั้งบนโครงสร้างที่ลอยอยู่บนแหล่งน้ำ โดยทั่วไปจะเป็นอ่างเก็บน้ำหรือทะเลสาบ โรงไฟฟ้าประเภทนี้มีจำนวนเพิ่มมากขึ้นในประเทศจีน ญี่ปุ่น สหรัฐอเมริกา ฝรั่งเศส อินเดีย และเกาหลีใต้ เป็นต้น ถือเป็นนวัตกรรมที่หลายประเทศในยุโรปและภูมิภาคเอเชียนำมาใช้ในการผลิตไฟฟ้า (Cazzaniga & Rosa-Clot, 2021) ด้วยข้อดีในการลดการระเหยของน้ำและการที่มีโครงสร้างแบบลอยตัวบนผิวน้ำ สามารถช่วยระบบระบายความร้อนและเพิ่มประสิทธิภาพของแผงโซลาร์เซลล์ได้ดีขึ้น (World Bank, 2018) สำหรับประเทศไทย ได้มีการก่อสร้างและพัฒนาโรงไฟฟ้าพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ บนพื้นที่อ่างเก็บน้ำในพื้นที่ภาคตะวันออกเฉียงเหนือ ขนาดกำลังการผลิต 45 เมกะวัตต์ สามารถผลิตไฟฟ้าจากพลังงานแสงอาทิตย์ ช่วยเสริมความมั่นคงทางพลังงานไฟฟ้าให้กับประเทศได้อีกทางหนึ่ง แต่ข้อจำกัดที่สำคัญของโรงไฟฟ้าพลังงานแสงอาทิตย์ คือ ค่ากำลังไฟฟ้าที่ผลิตได้มีความไม่แน่นอน เนื่องจากปัจจัยด้านภูมิอากาศในช่วงเวลานั้น ๆ มีความผันผวนไม่คงที่ เช่น ความเข้มแสงอาทิตย์ อุณหภูมิอากาศ อุณหภูมิบนแผง ความชื้น ความเร็วลม เป็นต้น ล้วนเป็นปัจจัยที่ไม่สามารถควบคุมและหลีกเลี่ยงได้ ส่งผลกระทบต่อคุณภาพและเสถียรภาพของระบบไฟฟ้า จึงจำเป็นต้องปรับปรุงความแม่นยำในการพยากรณ์ล่วงหน้าแบบวันต่อวัน การมีระบบพยากรณ์ที่เชื่อถือได้จะช่วยให้การวางแผนบริหารจัดการกำลังไฟฟ้ามีประสิทธิภาพดีขึ้น

เพื่อเพิ่มเสถียรภาพให้กับการวางแผนบริหารจัดการกำลังไฟฟ้าและลดข้อจำกัดด้านภูมิอากาศ จึงได้มีการนำเสนอวิธีการพยากรณ์พลังงานแสงอาทิตย์ด้วยวิธีการทางปัญญาประดิษฐ์ (artificial intelligence method) (Ye et al., 2022) ผ่านวิธีการเรียนรู้ของเครื่อง (machine learning) เช่น การใช้โครงข่ายประสาทเทียม (Artificial Neuron Network: ANN) งานวิจัยของ (ภาณุพงศ์ น้ำแก้ว, 2563) นำเสนอโครงข่ายประสาทเทียมแบบหลายชั้น (multilayer network) โดยมีรูปแบบการทำงานแบบป้อนไปข้างหน้า (feed forward) มีการเรียนรู้แบบแพร่ย้อนกลับ (back propagation) มาสร้างแบบจำลองการพยากรณ์กำลังไฟฟ้าของระบบเซลล์แสงอาทิตย์ในวันถัดไปโดยใช้ข้อมูลป้อนเข้าจากอุปกรณ์วัดที่ติดตั้งบนหลังคาที่פקอาศัย ต่อมาได้มีการประยุกต์ใช้แบบจำลองการเรียนรู้เชิงลึก (deep learning) เช่น โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network: RNN) หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory: LSTM) และโครงข่ายประตูกลับ (Gated Recurrent Unit: GRU) สำหรับใช้จัดการข้อมูลที่มีลักษณะเป็นข้อมูลที่มีลำดับหรืออนุกรมเวลา งานวิจัยของ (Jebli et al., 2021) ได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองทั้ง 3 แบบจำลองในการพยากรณ์พลังงานแสงอาทิตย์ พบว่า แบบจำลอง RNN และ LSTM มีความเหมาะสมสำหรับการพยากรณ์ เนื่องจากความสามารถในการจัดการข้อมูลที่เป็นลำดับและปัญหาการถดถอยอนุกรมเวลา เมื่อเทียบกับแบบจำลอง GRU ที่ให้ประสิทธิภาพน้อยกว่าเล็กน้อย นอกจากนี้ยังมีการนำเสนอแบบจำลอง Convolutional Neural Network (CNN) มาใช้พยากรณ์ผสมกับแบบจำลอง LSTM เพื่อเพิ่มประสิทธิภาพในการพยากรณ์ให้กับข้อมูล เช่น งานวิจัยของ (Lim et al., 2022) ใช้แบบจำลอง CNN ในการจำแนกสภาพอากาศ ผสมกับแบบจำลอง LSTM ในการพยากรณ์กำลังไฟฟ้าตามสภาพอากาศ เพิ่มความแม่นยำในการพยากรณ์ให้สูงขึ้น

ดังนั้นในงานวิจัยนี้จึงได้นำเสนอแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ โดยอาศัยวิธีการหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory: LSTM) และเอ็กซ์ตรีมเกรเดียนต์บูสติง (Extreme Gradient Boosting: XGBoost) ที่มีความสามารถในการจัดการกับข้อมูลที่ไม่เป็นไม่เชิงเส้นและแยกแยะลักษณะความผันผวนของข้อมูลอนุกรมเวลาได้ดี และเพื่อให้ได้ประโยชน์สูงสุดจากแบบจำลองทั้งสองจึงได้นำเสนอวิธีการพยากรณ์แบบผสม เพื่อปรับปรุงคุณภาพและความแม่นยำในการพยากรณ์โดยรวมให้ดีขึ้น ทำให้การวางแผนการผลิตไฟฟ้าจากพลังงานแสงอาทิตย์มี

ประสิทธิภาพมากขึ้น และสามารถนำไปใช้ในการวางแผนและบริหารจัดการการผลิตไฟฟ้าเพื่อรองรับการเติบโตของโรงไฟฟ้าพลังงานแสงอาทิตย์ที่กำลังจะเกิดขึ้นต่อไปในอนาคต

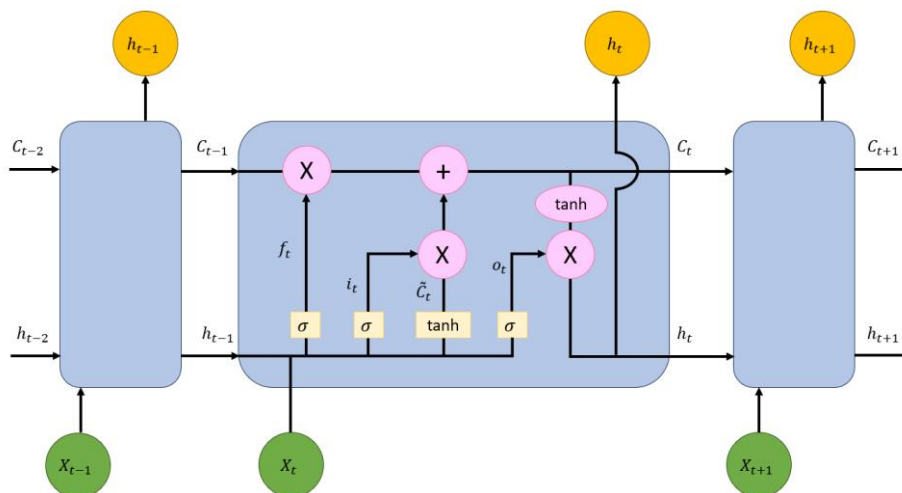
2. วัตถุประสงค์ของการวิจัย

2.1 เพื่อศึกษาวิธีการและสร้างแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนพุ่มลอยน้ำ โดยอาศัยวิธีการหน่วยความจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตตึง

2.2 เพื่อประเมินประสิทธิภาพของแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนพุ่มลอยน้ำ โดยอาศัยวิธีการหน่วยความจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตตึง

3. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory: LSTM) ถูกพัฒนาจากโครงข่ายประสาทแบบวนซ้ำ (Recurrent neural network: RNN) รู้จักในปี ค.ศ. 1997 (Hochreiter & Schmidhuber, 1997) เพื่อแก้ปัญหาการลดลงของเกรเดียนต์ (vanishing gradient) ด้วยการเพิ่มหน่วยความจำระยะยาว (long term memory) คำนวณร่วมกับหน่วยความจำระยะสั้น (short term memory) โดยที่หน่วยความจำระยะยาวใน LSTM สามารถเก็บความสัมพันธ์ของข้อมูลที่มีความยาวได้ดี มีความเหมาะสมในการแก้ปัญหาที่เกี่ยวข้องกับข้อมูลตามลำดับ (sequential data) มีการเพิ่มประตูสัญญาณ (gate) สำหรับการตัดสินใจของ cell state จำนวน 3 ประตู ได้แก่ forget gate (f_t), input gate (i_t) และ output gate (o_t) (Saxena, 2021) แสดงดังรูปที่ 1 คำนวณได้ดังสมการที่ 1 – 6



รูปที่ 1 โครงสร้างของหน่วยความจำระยะสั้นแบบยาว

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

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i) \quad (2)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

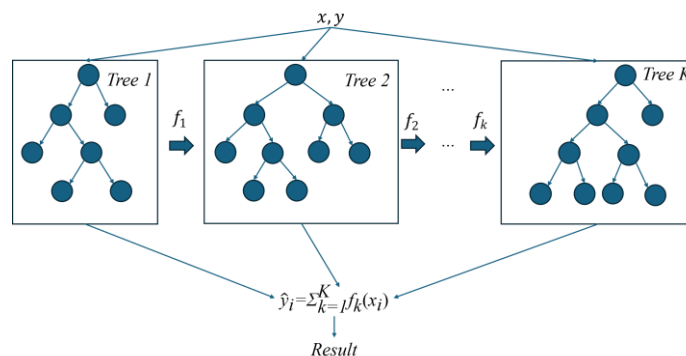
$$C_t = \sigma(f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t) \quad (5)$$

$$h_t = \tanh(C_t) \cdot o_t \quad (6)$$

โดยที่

f_t	คือ	Forget Gate
i_t	คือ	Input Gate
o_t	คือ	Output Gate
c_t	คือ	ค่า Cell State ณ เวลา t
h_t	คือ	ค่า ค่า Hidden State ณ เวลา t
$\tilde{C}_t$	คือ	ค่า Candidate Values ของ Cell State ณ เวลา t
$\tanh$	คือ	ฟังก์ชัน tanh
σ	คือ	ฟังก์ชันซิกมอยด์
W	คือ	ค่าน้ำหนักของ Matrices
h_{t-1}	คือ	ค่าน้ำหนักของ Output ของ Cell State ก่อนหน้า ณ เวลา t-1
x_t	คือ	ค่า Input ที่เข้ามาใน Cell State ณ เวลา t
b	คือ	ค่า Bias

3.2 เอ็กซ์ตรีมเกรเดียนต์บูสตีง (Extreme Gradient Boosting: XGBoost) เป็นเทคนิคการเรียนรู้ของเครื่องที่พัฒนามาจากอัลกอริทึมต้นไม้ความเร็วแบบไล่ระดับ (gradient boosted tree) จัดอยู่ในกลุ่ม ensemble learning โดยแบบจำลองได้เลือกใช้ต้นไม้ตัดสินใจมาฝึกสอนต่อกันจำนวนหลาย ๆ ต้น เพื่อเรียนรู้จากค่าความผิดพลาดของกระบวนการทำงานก่อนหน้า ส่งผลต่อการพยากรณ์ที่ให้ค่าความแม่นยำมากขึ้น โดยแบบจำลองจะถูกหยุดการเรียนรู้ เมื่อพบว่าการเรียนรู้ของต้นไม้ตัดสินใจต่อเนื่องจนมีความลึกมากเพียงพอ จนไม่เหลือค่าความผิดพลาดจากต้นไม้ตัดสินใจต้นก่อนหน้าให้เรียนรู้ (Chen & Guestrin, 2016) ถูกนำมาใช้ในการแก้ปัญหาการถดถอยและการจำแนกประเภท สามารถตรวจจับความสัมพันธ์ของข้อมูลที่มีความเป็นลำดับ รวมถึงข้อมูลที่ไม่เป็นเชิงเส้น แสดงดังรูปที่ 2 คำนวณได้ดังสมการที่ 7



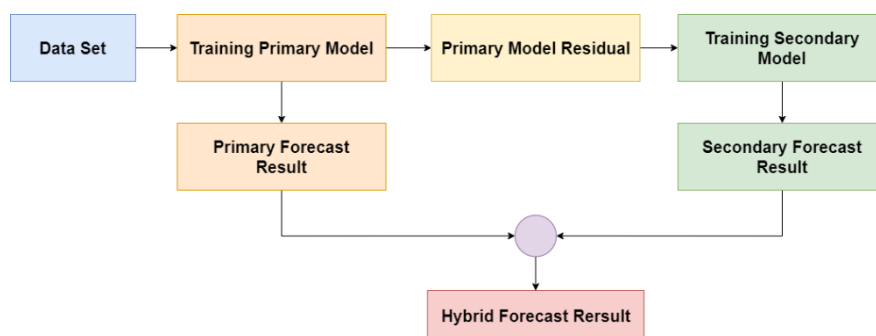
รูปที่ 2 โครงสร้างของเอ็กซ์ตรีมเกรเดียนต์บูสตีง

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad (7)$$

โดยที่

$\hat{y}_i$	คือ	ค่าพยากรณ์ของข้อมูลลำดับที่ i
x_i	คือ	Feature Vector
K	คือ	จำนวนต้นไม้ในการตัดสินใจ
$f_k(x_i)$	คือ	ค่าพยากรณ์ของลำดับที่ k

3.3 วิธีการพยากรณ์แบบผสม (Hybrid Forecasting Method) ถูกพัฒนาเพื่อลดข้อจำกัดของวิธีการหนึ่งด้วยข้อดีของอีกวิธีการหนึ่ง ช่วยแก้ปัญหาสมมติฐานเชิงเส้น เนื่องจากข้อมูลอนุกรมเวลาอาจเป็นแบบเชิงเส้นหรือไม่เป็นเชิงเส้นก็ได้ ช่วยให้การพยากรณ์มีประสิทธิภาพมากขึ้น ในกรณีที่วิธีการพยากรณ์อนุกรมเวลาแบบดั้งเดิมมีข้อจำกัด (Berberich, 2020) โดยเป็นการสร้างแบบจำลองเดี่ยวและคัดเลือกแบบจำลองเดี่ยวที่มีความเหมาะสมและแม่นยำมากที่สุด มาสร้างแบบจำลองผสม เมื่อได้แบบจำลองจากวิธีพยากรณ์หลักให้นำมาหาค่าส่วนเหลือ (Residual) จากแบบจำลองพยากรณ์หลัก จากนั้นนำค่าส่วนเหลือที่ได้เป็นข้อมูลนำเข้าให้กับแบบจำลองด้วยวิธีพยากรณ์รอง จากนั้นนำค่าผลลัพธ์จากวิธีพยากรณ์หลักมารวมกับค่าพยากรณ์จากวิธีพยากรณ์รอง จะได้ค่าผลลัพธ์สุดท้าย (อุรสา จันทรภา, 2564) แสดงดังรูปที่ 3



รูปที่ 3 วิธีการพยากรณ์แบบผสม

จากทฤษฎีที่เกี่ยวข้อง ผู้วิจัยพบว่า วิธีการที่เหมาะสมกับชุดข้อมูลสำหรับสร้างแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ มีดังนี้

1) **หน่วยความจำระยะสั้นแบบยาว (LSTM)** เนื่องจากวิธีการนี้ มีการเพิ่มประตู (gates) ที่ช่วยควบคุมการไหลของข้อมูล ทำให้สามารถรักษาข้อมูลที่สำคัญไว้ได้นานขึ้นและลดปัญหาการลดลงของเกรเดียนต์ รวมถึงจัดการกับลำดับข้อมูลอนุกรมเวลาที่มีความยาวไม่คงที่ มีความสามารถในการเรียนรู้รูปแบบระยะยาวและจัดการกับความซับซ้อนของข้อมูล ให้ผลลัพธ์การพยากรณ์ที่แม่นยำกว่าเมื่อเทียบกับแบบจำลองทางสถิติแบบดั้งเดิม จึงเหมาะกับชุดข้อมูลการผลิตพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ ที่เป็นข้อมูลประเภทอนุกรมเวลาและไม่เป็นเชิงเส้น

2) **เอ็กซ์ตรีมเกรเดียนต์บูตติ้ง (XGBoost)** เนื่องจากวิธีการนี้มีการปรับปรุงประสิทธิภาพในการคำนวณแบบขนาน (Parallel Processing) และการจัดการหน่วยความจำที่มีประสิทธิภาพ รวมถึงมีความเร็วในการฝึกสอนแบบจำลองและการพยากรณ์ที่มีความแม่นยำสูง มีความสามารถในการตรวจจับความสัมพันธ์ของข้อมูลที่มีความเป็นลำดับ และข้อมูลที่ไม่เป็นเชิงเส้น รวมถึงจัดการกับข้อมูลที่มีความซับซ้อนและเปลี่ยนแปลงตลอดเวลาได้ดี สามารถตรวจจับแนวโน้ม ฤดูกาล และ

ความสัมพันธ์ที่ซับซ้อนของข้อมูลได้ จึงเหมาะสมกับการพยากรณ์ข้อมูลที่มีหลายปัจจัยเข้ามาเกี่ยวข้อง เช่น ข้อมูลจาก เซนเซอร์ตรวจวัดสภาพอากาศและสภาพแวดล้อมในกระบวนการผลิตพลังงานไฟฟ้าจากพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ

3) *วิธีการพยากรณ์แบบผสม (Hybrid Forecasting Method)* เนื่องจากวิธีการนี้มีความสามารถในการพยากรณ์ข้อมูลแบบอนุกรมเวลาและไม่เป็นเชิงเส้น โดยการผสมผสานวิธีการเรียนรู้ของเครื่องและการพยากรณ์ทางสถิติ มาใช้ในการสร้างตัวแบบจำลองการพยากรณ์ที่เหมาะสม ช่วยเพิ่มประสิทธิภาพการพยากรณ์ให้สูงขึ้นกว่าการใช้แบบจำลองแบบเดี่ยว จึงเหมาะกับการพยากรณ์ปริมาณการผลิตพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ ที่เป็นข้อมูลประเภทอนุกรมเวลาและไม่เป็นเชิงเส้น

4. วิธีดำเนินการวิจัย

วิธีการวิจัยครั้งนี้ประกอบด้วย 3 ขั้นตอนหลัก คือ การเก็บรวบรวมข้อมูล การคัดเลือกคุณลักษณะและการจำแนกประเภท การประเมินประสิทธิภาพของแบบจำลอง มีรายละเอียดดังนี้

4.1 **การเก็บรวบรวมข้อมูล** โดยใช้ข้อมูลค่ากำลังไฟฟ้าที่ผลิตได้และข้อมูลจากอุปกรณ์เซนเซอร์ที่อยู่บนฐานข้อมูลของ โรงไฟฟ้าพลังงานแสงอาทิตย์บนทุ่นลอยน้ำแห่งหนึ่ง ขนาดกำลังการผลิต 45 เมกะวัตต์ ตั้งแต่เดือนกุมภาพันธ์ ถึงเดือนตุลาคม พ.ศ. 2566 จำนวน 6,511 ระเบียบ ซึ่งถูกเก็บรวบรวมข้อมูลเป็นรายชั่วโมง (hourly) และมี 11 คุณลักษณะ โดยมีรายละเอียดคุณลักษณะดังตารางที่ 1 ทั้งนี้ ข้อมูลที่ได้เป็นข้อมูลดิบจากเซนเซอร์โดยตรง ซึ่งอาจมีค่าผิดปกติ (outlier) เช่น ค่าอุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 3 ที่มีค่าสูงที่สุดถึง 100 องศาเซลเซียส ซึ่งอาจเกิดจากปัจจัยเฉพาะขณะทำการวัด แต่งานวิจัยนี้ได้เลือกใช้ข้อมูลดังกล่าวเพื่อสร้างแบบจำลองที่ทนทานต่อความผันผวนของข้อมูลจริง

ตารางที่ 1 รายละเอียดคุณลักษณะ

คุณลักษณะ	รายละเอียด	ช่วงข้อมูล	หน่วยวัด
Date	วันที่	1/2/2023 - 31/10/2023	-
Time	เวลา	00:00 - 23:00	-
Solar_active_power	ค่ากำลังไฟฟ้า	0.00 - 47.61961	เมกะวัตต์ (MW)
Obliquity_radiation	ค่าความเข้มแสงอาทิตย์	0.00 - 1321.00	วัตต์ต่อตารางเมตร (W/m ²)
Environment_temperature	อุณหภูมิสภาพแวดล้อม	16.5 - 37.1	องศาเซลเซียส (°C)
Environment_humidity	ค่าความชื้น	82.0 - 99.0	ร้อยละ (%)
Wind_speed	ความเร็วลม	0.00 - 42.00	เมตรต่อวินาที (m/s)
Wind_direction	ทิศทางลม	0.00 - 360	องศา (Degree)
Panel_temperature_1	อุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 1	12.9 - 74.1	องศาเซลเซียส (°C)
Panel_temperature_2	อุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 2	10.4 - 69.3	องศาเซลเซียส (°C)
Panel_temperature_3	อุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 3	10.0 - 100	องศาเซลเซียส (°C)

4.2 การคัดเลือกคุณลักษณะและการจำแนกประเภท

4.2.1 การแปลงข้อมูล เป็นการทำความสะอาดข้อมูลเพื่อตัดข้อมูลที่ผิดปกติออก ข้อมูลที่ขาดหายไป (missing values) จัดกลุ่มของคุณลักษณะ และทำการแปลงข้อมูลซึ่งเป็นการทำให้ข้อมูลอยู่ในรูปแบบที่พร้อมนำเอาไปใช้ในการวิเคราะห์ข้อมูลตามอัลกอริทึมที่เลือกใช้

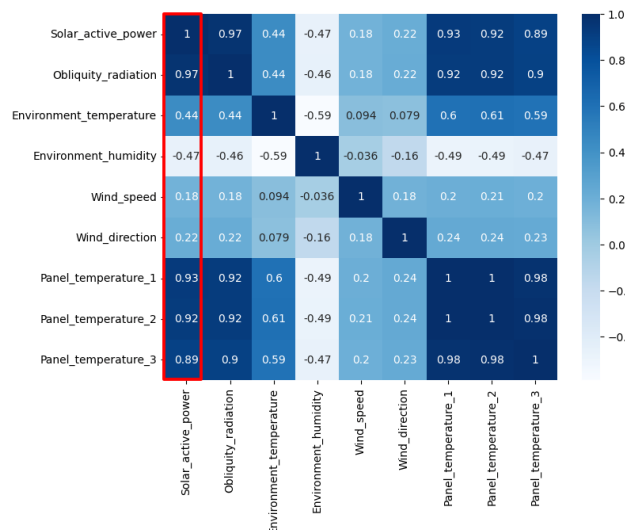
4.2.2 การคัดเลือกคุณลักษณะ เป็นวิธีการที่ช่วยลดจำนวนคุณลักษณะ โดยหาความสัมพันธ์จากคุณลักษณะทั้งหมดที่มีเพื่อดูความเกี่ยวข้องกัน ดังนั้นจึงใช้การเลือกคุณลักษณะเพื่อกำจัดเฉพาะคุณลักษณะที่มีความเกี่ยวข้องกันเท่านั้น (เกศินี ช่อนกลิ่น, 2566) เนื่องจากชุดข้อมูลมีลักษณะเป็นตัวเลข (numeric) และผลลัพธ์เป้าหมาย (target result) เป็นตัวเลข ผู้วิจัยจึงเลือกใช้การคัดเลือกคุณลักษณะด้วยวิธีสหสัมพันธ์เพียร์สัน (Pearson's correlation) เพื่อหาคุณลักษณะที่สัมพันธ์กันจำนวนได้ดังสมการที่ 8

$$r_{xy} = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (8)$$

โดยที่

r_{xy}	คือ	สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน
x	คือ	ตัวแปรอิสระหรือตัวแปรต้น
y	คือ	ตัวแปรตาม
n	คือ	ขนาดข้อมูล

งานวิจัยนี้ได้คัดเลือกคุณลักษณะที่มีความสัมพันธ์กับค่ากำลังไฟฟ้า โดยใช้เกณฑ์การคัดเลือกที่มีค่าตั้งแต่ 0.5 – 1.0 จึงจะถือว่าคุณลักษณะนั้นมีความสัมพันธ์กัน แสดงดังรูปที่ 4



รูปที่ 4 ค่าความสัมพันธ์ของข้อมูลด้วยค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

จากรูปที่ 4 พบว่าคุณลักษณะที่มีความสัมพันธ์กับค่ากำลังไฟฟ้า (Solar_active_power) ตามเกณฑ์ที่กำหนด (ค่าสัมประสิทธิ์ ≥ 0.5) ในกรอบสีแดง ได้แก่ ค่าความเข้มแสงอาทิตย์ (0.97), อุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 1 (0.93), อุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 2 (0.92) และอุณหภูมิบนแผงเซลล์แสงอาทิตย์ชุดที่ 3 (0.89) ดังนั้น คุณลักษณะทั้ง 4 นี้จึงถูกเลือกเป็นข้อมูลป้อนเข้าแบบจำลอง

4.2.3 การปรับขนาดข้อมูลให้มาตรฐาน (scaling data) เนื่องจากคุณลักษณะในงานวิจัยนี้มีช่วงข้อมูลที่แตกต่างกันอย่างมาก สังเกตได้จากค่าอุณหภูมิบนแผง เทียบกับค่าความเข้มแสงอาทิตย์ที่มีช่วงข้อมูลกว้างกว่า เมื่อพบว่าข้อมูลส่วนใหญ่ที่เป็นตัวเลขมีความแตกต่างจากข้อมูลตัวเลขอื่น จำเป็นต้องมีการปรับขนาดข้อมูลให้สมมาตร (Liu et al., 2021) ผู้วิจัยใช้วิธีการทำให้เป็นมาตรฐาน (normalization) เป็นการแปลงค่าข้อมูลให้อยู่ในช่วงสั้น ๆ ตั้งแต่ 0 – 1 หรือที่เรียกว่าการทำให้เป็นมาตรฐานต่ำสุด-สูงสุด (min-max normalization) คำนวณได้ดังสมการที่ 9

$$X^* = \frac{X - X_{min}}{X - X_{max}} \quad (9)$$

โดยที่

X^*	คือ	ค่าที่จะทำให้เป็นมาตรฐานของตัวแปร x
x	คือ	ค่าดั้งเดิมของตัวแปร x
X_{min}	คือ	ค่าต่ำสุดของตัวแปร x
X_{max}	คือ	ค่าสูงสุดของตัวแปร x

4.2.4 การแบ่งข้อมูลรูปแบบ Sliding Window เป็น การแบ่งชุดข้อมูลฝึกสอนก่อนนำเข้าแบบจำลอง จุดเริ่มต้นของข้อมูลจะถูกขยับออกไปเรื่อย ๆ แต่จำนวนชุดข้อมูลทั้งหมดยังคงเดิม (Judith Foster et al., 2017) ผู้วิจัยใช้การแบ่งข้อมูลทุก ๆ 24 ค่า เพื่อพยากรณ์ค่าถัดไป

4.2.5 การแบ่งข้อมูลเพื่อประเมินประสิทธิภาพแบบจำลอง งานวิจัยครั้งนี้ได้เลือกใช้วิธีการแบ่งข้อมูลตามลำดับเวลา (sequential split) ซึ่งเป็นวิธีการที่เหมาะสมสำหรับการพยากรณ์ข้อมูลอนุกรมเวลาที่มีปัจจัยด้านฤดูกาลเข้ามาเกี่ยวข้อง ข้อมูลนำเข้าแบบจำลองสำหรับการพยากรณ์ ถูกแบ่งออกเป็น 3 ชุด ได้แก่ ชุดข้อมูลฝึกสอน (training data) ชุดข้อมูลตรวจสอบ (validation data) และชุดข้อมูลทดสอบ (testing data) ในอัตราส่วน 80:10:10 ตามลำดับ สำหรับข้อมูลที่ใช้ในงานวิจัยนี้ครอบคลุมช่วงเดือนกุมภาพันธ์ถึงตุลาคม ซึ่งกินระยะเวลา 9 เดือน และครอบคลุมการเปลี่ยนแปลงของฤดูกาลในประเทศไทย (เช่น การเปลี่ยนผ่านจากฤดูร้อนเข้าสู่ฤดูฝน) ชุดข้อมูลสำหรับฝึกสอน (training data) ซึ่งมีขนาดใหญ่ถึงร้อยละ 80 ของข้อมูลทั้งหมด จึงเพียงพอให้แบบจำลองได้เรียนรู้รูปแบบความสัมพันธ์ระหว่างปัจจัยต่าง ๆ (เช่น ความเข้มแสง, อุณหภูมิ) กับกำลังการผลิตไฟฟ้าที่เปลี่ยนแปลงไปตามฤดูกาล

4.2.6 การจำแนกประเภท งานวิจัยนี้ได้เลือกอัลกอริทึมการเรียนรู้เชิงลึก 2 อัลกอริทึม ได้แก่ อัลกอริทึมหน่วยความจำระยะสั้นแบบยาว (LSTM) และอัลกอริทึมเอ็กซ์ตรีมเกรเดียนต์บูสติง (XGBoost) และประยุกต์ใช้วิธีการพยากรณ์แบบผสม (Hybrid Forecasting Method) เพื่อเพิ่มประสิทธิภาพให้กับแบบจำลอง ดังนี้

1) วิธีการหน่วยความจำระยะสั้นแบบยาว (LSTM) ต้องอาศัยค่าไฮเปอร์พารามิเตอร์ (hyperparameter) ซึ่งผู้วิจัยได้ทำการกำหนดจำนวนชั้นซ่อน 2 ชั้น จำนวนโหนด 32 โหนด แม้ว่าหลักการทั่วไปจะแนะนำว่า จำนวนโหนดในชั้นซ่อนไม่ควรเกิน 2 เท่าของจำนวนโหนดชั้นนำเข้า (ศรีรักษ์ ศรีทองชัย, 2566) แต่จากการทดลอง พบว่าการใช้จำนวนโหนดที่ 32 ให้ประสิทธิภาพที่ดีที่สุดในการเรียนรู้รูปแบบที่ซับซ้อนของข้อมูลอนุกรมเวลา และเพื่อป้องกันการเกิดภาวะ Overfitting ซึ่งอาจ

เกิดขึ้นได้ในแบบจำลองที่มีความซับซ้อน จึงได้กำหนดค่า dropout เท่ากับ 0.2 กำหนด batch size เท่ากับ 32 ใช้ฟังก์ชันกระตุ้น (activation function) คือ Relu ใช้ optimizer คือ Adam (Adaptive Moment Estimation) กำหนดให้ค่าความคลาดเคลื่อน Mean Absolute Error (MAE) เป็นฟังก์ชันการสูญเสีย (loss function) และกำหนดรอบการสอน (epoch) เท่ากับ 100 ฟังก์ชัน callback early stopping เท่ากับ 50 แสดงดังตารางที่ 2

ตารางที่ 2 ไฮเปอร์พารามิเตอร์สำหรับวิธีการหน่วยความจำระยะสั้นแบบยาว

Hyperparameter	LSTM
Model	Sequential
Activation	Relu
Input	Input Array
Hidden Node	32 neurons
Hidden Layer	(Dropout)
Output	Output Array
Loss Calculation	MAE
Optimizer	Adam

2) วิธีการเอ็กซ์ตรีมเกรเดียนต์บูสตีง (Extreme Gradient Boosting: XGBoost) งานวิจัยนี้ได้กำหนดค่าพารามิเตอร์โดยยึดถือหลักการสร้างแบบจำลองพื้นฐาน (Baseline Model) ที่มีประสิทธิภาพและน่าเชื่อถือ ผู้วิจัยจึงอ้างอิงและปรับปรุงชุดพารามิเตอร์จากงานวิจัยของ (Li et al., 2019) และ (Obiora et al., 2021) ซึ่งมีชุดข้อมูลที่สอดคล้องกับงานวิจัยนี้และมีบริบทของปัญหาที่คล้ายคลึงกัน แสดงดังตารางที่ 3

ตารางที่ 3 พารามิเตอร์สำหรับวิธีการเอ็กซ์ตรีมเกรเดียนต์บูสตีง

Class	Parameter	Value
General Parameters	Default	-
Booster Parameters	colsample_bytree	0.7
Booster Parameters	learning_rate	0.01
Booster Parameters	max_depth	3
Booster Parameters	n_estimators	1000
Early Stopping	early_stopping_rounds	20
Learning Task Parameters	eval_metric	MAE

3) วิธีการพยากรณ์แบบผสม (Hybrid Forecasting Method) เพื่อเพิ่มประสิทธิภาพให้กับแบบจำลองขั้นสุดท้าย วิธีการพยากรณ์แบบผสมจึงถูกนำมาใช้ในขั้นตอนนี้ ขั้นตอนการทำงานของแบบจำลองขั้นสุดท้ายด้วยวิธีการพยากรณ์แบบผสม แสดงดังรหัสเทียม (Pseudo Code)

Algorithm: Hybrid Forecasting (LSTM-XGBoost)**Input:** Dataset (D)**Output:** Hybrid Forecast Result (F_H)

// --- Primary Model (LSTM) Processing ---

1. $Data_{train} \leftarrow D$ // Assuming entire dataset D is used for training in this context
2. $Model_{lstm} \leftarrow Train_Primary_Model (Data_{train})$
3. $F_{lstm} \leftarrow Generate_Primary_Forecast (Model_{lstm}, Data_{train,features})$
4. $R_{lstm} \leftarrow Compute_Residual(Data_{train,actuals}, F_{lstm})$ // $R_{lstm} = Actual - F_{lstm}$

// --- Secondary Model (XGBoost) Processing (Forecasting Residuals) ---

5. $Model_{XGBoost} \leftarrow Train_Secondary_Model (R_{lstm}, Data_{train,features_for_residual})$
6. $F_{Residual} \leftarrow Generate_Secondary_Forecast (Model_{XGBoost}, Data_{train,features_for_residual})$

// --- Hybrid Forecast Composition ---

7. $F_H \leftarrow Combine_Forecasts (F_{lstm}, F_{Residual})$ // $F_H = F_{lstm} + F_{Residual}$
8. **Return** F_H

4.3 การประเมินประสิทธิภาพแบบจำลอง

งานวิจัยครั้งนี้ ได้มุ่งเน้นการวัดผลในด้านความแม่นยำของแบบจำลองในการพยากรณ์ให้ใกล้เคียงกับค่าจริงมากที่สุด เพื่อนำไปใช้งานในเชิงปฏิบัติ จึงดำเนินการทดสอบประสิทธิภาพของแบบจำลองด้วยการวัดค่าเฉลี่ยของผลต่างสัมบูรณ์ (MAE) การหาค่าเฉลี่ยความผิดพลาดกำลังสอง (MSE) และการหาค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (RMSE) ซึ่งเป็นเครื่องมือที่ได้รับความนิยมในการนำมาประเมินประสิทธิภาพของแบบจำลองแบบการถดถอย คำนวณได้ดังสมการที่ 10 – 12

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (10)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (12)$$

โดยที่

y_t	คือ	ค่าจริงที่ t ไต ๆ
$\hat{y}_t$	คือ	ค่าที่ได้จากการพยากรณ์ที่ t ไต
n	คือ	ขนาดข้อมูล

จากสมการ 10 - 12 เป็นการประเมินความสามารถของแบบจำลองที่ได้ผ่านการฝึกสอนแล้ว โดยการทดสอบประสิทธิภาพของแบบจำลองจะถูกตัดสินด้วยความแม่นยำของผลลัพธ์ที่ได้จากแบบจำลอง หากค่าความผิดพลาดจากค่าจริงน้อย คือ มีค่าเข้าใกล้ 0 แสดงว่าแบบจำลองสามารถทำนายได้ถูกต้อง (วฤชาย รมสายหยุด, 2566)

5. ผลการวิจัยและอภิปรายผลการวิจัย

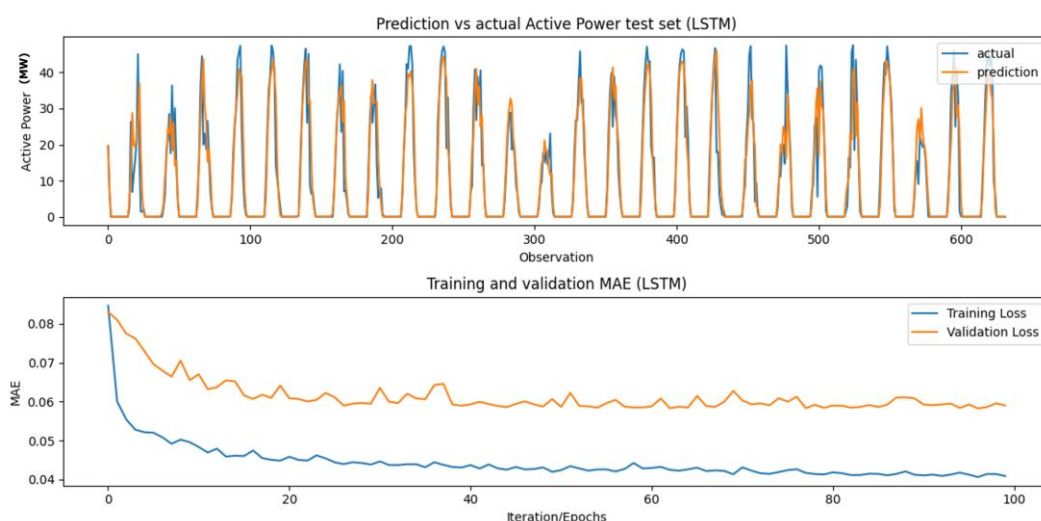
ในการศึกษาและสร้างแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ โดยอาศัยวิธีการหน่วยความจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตติ้ง ซึ่งใช้ข้อมูลการผลิตพลังงานไฟฟ้าและข้อมูลจากเซนเซอร์ที่บันทึกอยู่บนฐานข้อมูลของโรงไฟฟ้าพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ ได้รับการประเมินประสิทธิภาพของแบบจำลองด้วยค่าเฉลี่ยของผลต่างสัมบูรณ์ (MAE) ค่าเฉลี่ยความผิดพลาดกำลังสอง (MSE) และรากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (RMSE) ผลการประเมินประสิทธิภาพในการสร้างแบบจำลอง แสดงดังตารางที่ 4

ตารางที่ 4 เปรียบเทียบประสิทธิภาพของแบบจำลองทั้ง 3 วิธี

Methods	MAE	MSE	RMSE
LSTM	0.0585	0.0148	0.1215
XGBoost	0.0630	0.0147	0.1211
Hybrid	0.0577	0.0143	0.1196

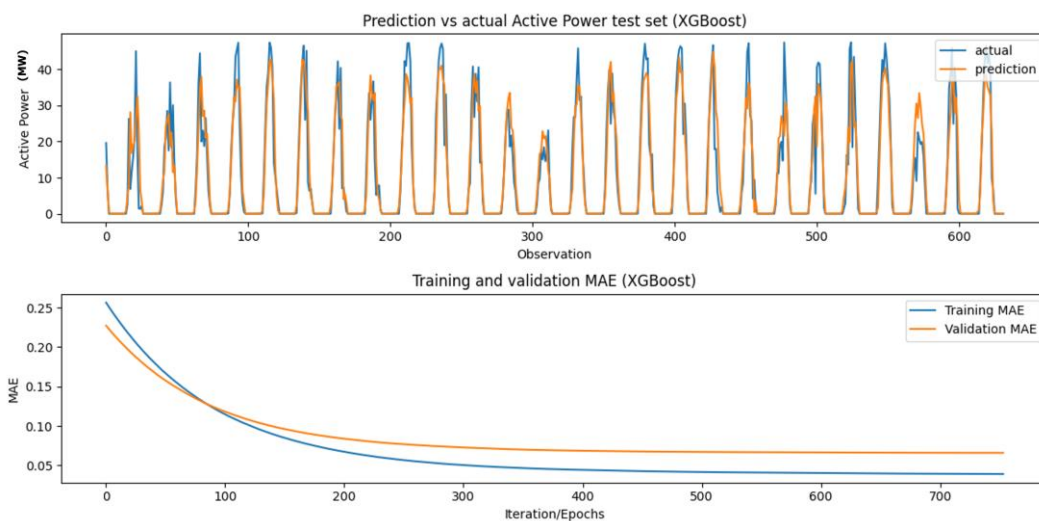
จากตารางที่ 4 แสดงให้เห็นว่าแบบจำลองที่สร้างจากวิธีการพยากรณ์แบบผสม (Hybrid) มีค่า MAE เท่ากับ 0.0577 ค่า MSE เท่ากับ 0.01430 และค่า RMSE เท่ากับ 0.1196 ซึ่งดีกว่าแบบจำลอง LSTM และ XGBoost

สำหรับผลการพยากรณ์ด้วยชุดข้อมูลทดสอบ (testing data) ที่ได้จากขั้นตอนการแบ่งข้อมูล ซึ่งครอบคลุมเวลา 632 ชั่วโมง ถูกแสดงในรูปที่ 5-7



รูปที่ 5 ผลการพยากรณ์ด้วยชุดข้อมูลทดสอบที่สร้างจาก LSTM

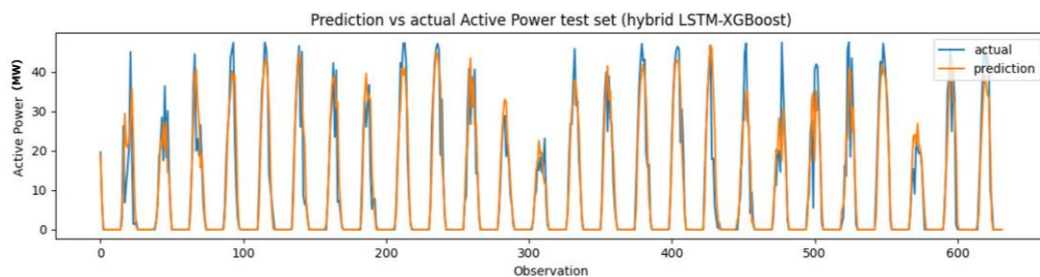
จากรูปที่ 5 ผลการพยากรณ์ด้วยชุดข้อมูลทดสอบที่สร้างจาก LSTM กราฟบนเส้นสีน้ำเงิน คือ ค่ากำลังไฟฟ้าจริงและเส้นสีส้ม คือ ค่ากำลังไฟฟ้าจากการพยากรณ์ จากชุดข้อมูลทดสอบ พบว่า ค่าพยากรณ์ที่ได้มีความใกล้เคียงกับค่ากำลังไฟฟ้าจริงตั้งแต่ชั่วโมงที่ 0 - ชั่วโมงที่ 632 โดยที่ LSTM สามารถติดตามรูปแบบของข้อมูลจริงได้อย่างแม่นยำ โดยเฉพาะในช่วงที่มีการเปลี่ยนแปลงอย่างรวดเร็วของกำลังไฟฟ้าขาขึ้น (Active Power) กราฟล่างเส้นสีน้ำเงิน คือ ค่าสูญเสียจากข้อมูลชุดฝึกสอนและเส้นสีส้ม คือ ค่าสูญเสียจากชุดข้อมูลการตรวจสอบ จากการเรียนรู้แต่ละรอบด้วย loss function พบว่า เมื่อมีการฝึกสอนแบบจำลอง ค่า training loss มีแนวโน้มลดลง เช่นเดียวกับ ค่า validation loss ที่เริ่มมีค่าคงที่ (stabilize) หลังผ่านไปประมาณ 20-30 epochs ซึ่งบ่งชี้ว่าแบบจำลองได้เรียนรู้รูปแบบสำคัญส่วนใหญ่แล้ว การฝึกฝนต่อไปอาจไม่ช่วยลดความคลาดเคลื่อนในชุดข้อมูลตรวจสอบและเสี่ยงต่อการเกิด overfitting



รูปที่ 6 ผลการพยากรณ์ด้วยชุดข้อมูลทดสอบที่สร้างจาก XGBoost

จากรูปที่ 6 ผลการพยากรณ์ด้วยชุดข้อมูลทดสอบที่สร้างจาก XGBoost กราฟบนเส้นสีน้ำเงิน คือ ค่ากำลังไฟฟ้าจริงและเส้นสีส้ม คือ ค่ากำลังไฟฟ้าจากการพยากรณ์ จากชุดข้อมูลทดสอบ พบว่าค่าพยากรณ์ที่ได้มีความใกล้เคียงกับค่ากำลังไฟฟ้าจริงตั้งแต่ชั่วโมงที่ 100 - ชั่วโมงที่ 200 โดยแบบจำลองสามารถจับแนวโน้มทั่วไปของข้อมูลได้ แต่มีความล่าช้า (lag effect) ในการตอบสนองต่อการเปลี่ยนแปลงอย่างฉับพลันของข้อมูล และให้ค่าความแม่นยำที่ต่ำกว่าในหลายจุด กราฟล่างเส้นสีน้ำเงิน คือค่า training MAE และเส้นสีส้ม คือ ค่า validation MAE จากการเรียนรู้แต่ละรอบของแบบจำลอง พบว่า เมื่อมีการฝึกสอนแบบจำลอง ค่า training MAE และ ค่า validation MAE มีแนวโน้มลดลงอย่างรวดเร็วในช่วงแรก และค่อย ๆ ชะลอตัวลง ประกอบกับจำนวน iteration สูงถึง 700 รอบ แสดงให้เห็นว่าต้นไม้แต่ละต้นที่เพิ่มเข้าไปในช่วงหลัง ช่วยปรับปรุงประสิทธิภาพของแบบจำลองได้ทีละน้อย จนกระทั่งประสิทธิภาพบนชุดข้อมูลตรวจสอบไม่ดีขึ้นและเข้าเกณฑ์การหยุดก่อนกำหนด ซึ่งยังอยู่ในระดับที่ยอมรับได้ และไม่แสดงสัญญาณของการเกิด overfitting ที่รุนแรง

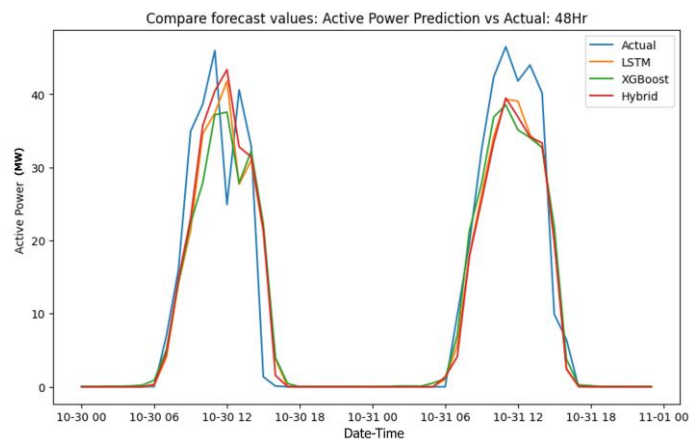
เมื่อพิจารณาผลการประเมินแบบจำลองด้วยค่า MAE แสดงให้เห็นถึงความแตกต่างที่ชัดเจนระหว่างแบบจำลองทั้งสอง โดยแบบจำลอง LSTM ให้ค่าสุดท้ายที่ 0.04-0.06 ในขณะที่แบบจำลอง XGBoost ให้ค่าสุดท้าย 0.05-0.06 นอกจากนี้แบบจำลอง LSTM ยังแสดงการลู่เข้าที่เร็วกว่าและมีเสถียรภาพสูงกว่าตลอดกระบวนการฝึกสอน หากพิจารณาในบริบทของข้อมูลที่มีค่ากำลังไฟฟ้าที่อยู่ในช่วง 0-45 เมกะวัตต์ ผลการประเมินด้วยค่า MAE ให้ค่าต่ำ แสดงให้เห็นว่าแบบจำลองสามารถพยากรณ์ได้แม่นยำ (สรุณี และคณะ., 2566) คิดเป็นร้อยละความคลาดเคลื่อนเพียง ร้อยละ 0.16-0.30 ของค่าเฉลี่ยของข้อมูล หรือประมาณร้อยละ 0.1-0.15 ของช่วงข้อมูลทั้งหมด



รูปที่ 7 ผลการพยากรณ์ด้วยชุดข้อมูลทดสอบที่สร้างจาก LSTM-XGBoost

จากรูปที่ 7 กำหนดเส้นสีน้ำเงิน คือ ค่ากำลังไฟฟ้าจริง และเส้นสีส้ม คือ ค่ากำลังไฟฟ้าจากการพยากรณ์ จากชุดข้อมูลทดสอบ 632 ชั่วโมง พบว่าแบบจำลองที่สร้างโดยอาศัยวิธีการแบบผสม ด้วยหน่วยจำระยะสั้นแบบยาวและเอ็กซ์ตรีมเกรเดียนต์บูตติ้ง (LSTM-XGBoost) มีค่าพยากรณ์ใกล้เคียงกับค่ากำลังไฟฟ้าจริง จากกราฟเปรียบเทียบค่าพยากรณ์กับค่าจริงแสดงให้เห็นว่าวิธีการแบบผสม มีความสามารถในการติดตามรูปแบบข้อมูลที่ใกล้เคียงกับแบบจำลอง LSTM แต่มีความเสถียรมากกว่าในบางช่วง และสามารถจับรูปแบบข้อมูลในช่วงที่มีการเปลี่ยนแปลงของกำลังไฟฟ้าอย่างทันทีทันใด

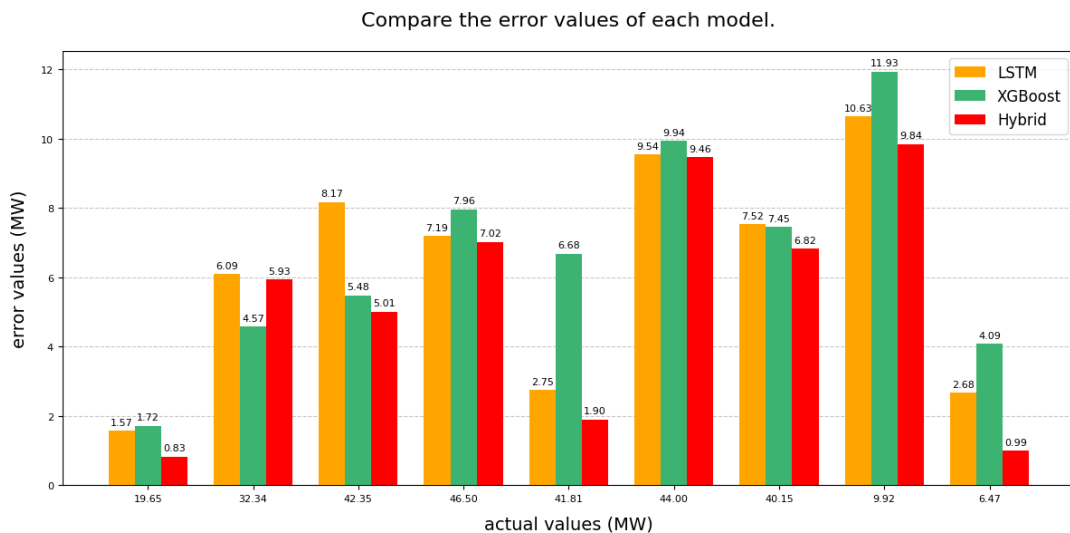
สำหรับการประเมินผลที่สำคัญที่สุดของแบบจำลองที่สร้างโดยอาศัยวิธีการแบบผสม มุ่งเน้นที่ประสิทธิภาพของแบบจำลองโดยรวมเมื่อทำงานเสร็จสมบูรณ์แล้วกับชุดข้อมูลทดสอบ ซึ่งเป็นผลจากกระบวนการฝึกสอนแบบจำลองที่แยกจากกัน การสร้างกราฟการเรียนรู้แบบรวมจึงไม่สามารถดำเนินการได้ในทางปฏิบัติและไม่สื่อถึงประสิทธิภาพที่แท้จริง ดังนั้น ผู้วิจัยเลือกนำเสนอผลลัพธ์สุดท้ายบนชุดข้อมูลทดสอบ เพื่อแสดงประสิทธิภาพของแบบจำลองแบบผสม



รูปที่ 8 เปรียบเทียบผลการพยากรณ์ของแบบจำลองทั้ง 3 วิธี

จากรูปที่ 8 แสดงการเปรียบเทียบผลการพยากรณ์ของแบบจำลองทั้ง 3 วิธี ในช่วงเวลา 48 ชั่วโมง ตั้งแต่วันที่ 30 ตุลาคม เวลา 00:00 น. ถึง วันที่ 1 พฤศจิกายน เวลา 00:00 น. โดยแกน X แสดงวันและเวลา (รูปแบบ เดือน-วัน ชั่วโมง) และแกน Y คือ Active Power (MW) แสดงให้เห็นลักษณะเฉพาะของแต่ละแบบจำลองได้อย่างชัดเจน จากกราฟ พบว่า แบบจำลอง LSTM แสดงความแม่นยำสูงในการพยากรณ์ในช่วงที่มีค่ากำลังไฟฟ้าสูงสุด แต่มีแนวโน้มการประมาณต่ำกว่าค่าจริงเล็กน้อยในบางจุด ซึ่งอาจส่งผลกระทบต่อวางแผนการผลิตไฟฟ้าในช่วงความต้องการไฟฟ้าสูงสุด (peak load) ในขณะที่แบบจำลอง

XGBoost แสดงผลการพยากรณ์ที่เสถียรมากกว่า แต่มีความล่าช้าในการตอบสนองต่อการเปลี่ยนแปลงอย่างฉับพลัน โดยเฉพาะในช่วงการขึ้นและลงของโหลด ทำให้เกิดความคลาดเคลื่อนในช่วงเวลาที่สำคัญ ส่วนแบบจำลองที่สร้างด้วยวิธีการพยากรณ์แบบผสม แสดงประสิทธิภาพที่สมดุลกว่า สามารถจับค่ากำลังไฟฟ้าสูงสุดได้แม่นยำใกล้เคียงกับแบบจำลอง LSTM และเสถียรมากกว่าในช่วงที่มีการเปลี่ยนแปลงน้อย



รูปที่ 9 กราฟเปรียบเทียบค่าความคลาดเคลื่อน (error) ของแบบจำลองทั้ง 3 วิธี

จากรูปที่ 9 เพื่อให้เห็นความแตกต่างของค่าความคลาดเคลื่อน (error) ที่พยากรณ์ได้ในเชิงตัวเลข ผู้วิจัยได้เลือกข้อมูลบางส่วนมาจากช่วงเวลาแสดงในรูปที่ 8 แสดงเป็นกราฟเปรียบเทียบค่าความคลาดเคลื่อนของแบบจำลองแต่ละวิธี โดยจะเห็นได้อย่างชัดเจนว่า ค่าความคลาดเคลื่อนของวิธีการแบบผสม (แท่งสีแดง) มีค่าต่ำที่สุดในกรณีส่วนใหญ่ (8 จาก 9 กรณี) แสดงให้เห็นว่าแบบจำลองแบบผสมให้ความแม่นยำและเสถียรภาพในการพยากรณ์สูงที่สุดในชุดการทดสอบนี้

6. สรุปผลการวิจัย

งานวิจัยนี้ได้นำเสนอแบบจำลองการพยากรณ์การผลิตพลังงานแสงอาทิตย์บนพุ่มลอยน้ำ โดยประยุกต์ใช้วิธีการแบบผสมด้วยอัลกอริทึมการเรียนรู้เชิงลึก 2 ชนิด ได้แก่ วิธีการหน่วยความจำระยะสั้นแบบยาว (LSTM) และเอ็กซ์ตรีมเกรเดียนต์บูสติ่ง (XGBoost) ผลการวิจัยชี้ให้เห็นว่าแบบจำลองแบบผสม (LSTM-XGBoost) มีประสิทธิภาพในการพยากรณ์สูงกว่าแบบจำลองแบบเดี่ยว

เมื่อพิจารณาในเชิงปริมาณจากข้อมูลประสิทธิภาพในการสร้างแบบจำลอง พบว่าแบบจำลองที่ประยุกต์ใช้วิธีการพยากรณ์แบบผสม (Hybrid) ให้ประสิทธิภาพสูงสุด โดยมีค่า MAE เท่ากับ 0.0577 ค่า MSE เท่ากับ 0.0143 และค่า RMSE เท่ากับ 0.1196 ซึ่งเป็นค่าความคลาดเคลื่อนที่ต่ำที่สุดในทุกตัวชี้วัด เพื่อความชัดเจนในการเปรียบเทียบประสิทธิภาพ เมื่อเทียบกับแบบจำลองเดี่ยวที่ดีที่สุดในกลุ่มคือ LSTM (MAE 0.0585, RMSE 0.1215) แบบจำลองแบบผสมสามารถลดค่า MAE ลงได้ คิดเป็นร้อยละ 1.37 และลดค่า RMSE ลงได้ คิดเป็นร้อยละ 1.56 เมื่อเทียบกับแบบจำลอง XGBoost (MAE 0.0630, RMSE 0.1211) แบบจำลองแบบผสมให้ค่า MAE ที่ต่ำกว่า คิดเป็นร้อยละ 8.41 และ RMSE ต่ำกว่า คิดเป็นร้อยละ 1.24

จากผลการประเมินประสิทธิภาพพบว่า การผสมผสานความสามารถของแบบจำลองหลายชนิดสามารถลดข้อจำกัดของแต่ละวิธีได้จริง โดยแบบจำลองหลัก (LSTM) ทำหน้าที่จับรูปแบบอนุกรมเวลาที่ซับซ้อนของข้อมูลการผลิตไฟฟ้า จากนั้นแบบจำลอง

รอง (XGBoost) จะเข้ามาเรียนรู้และพยากรณ์ในส่วนของคุณค่าความคลาดเคลื่อนส่วนเหลือ (residual) ที่แบบจำลองหลักไม่สามารถจับรูปแบบได้ สอดคล้องกับงานวิจัยของ (อุรสา จันทระ, 2564) ที่พบว่าการสร้างแบบจำลองแบบผสมช่วยเพิ่มประสิทธิภาพในการพยากรณ์ได้อย่างมีนัยสำคัญ แต่เนื่องจากลักษณะเฉพาะของชุดข้อมูลนำเข้ามีความแตกต่างกัน อาจไม่สามารถนำผลการวิจัยมาเปรียบเทียบกันได้โดยตรง

อย่างไรก็ตาม แบบจำลองที่พัฒนาขึ้น แสดงให้เห็นถึงความหลากหลายของเทคนิคการพยากรณ์แต่ละวิธี สามารถช่วยปรับปรุงประสิทธิภาพในการจับรูปแบบการเปลี่ยนแปลงของข้อมูลที่ซับซ้อน โดยเฉพาะในช่วงที่ข้อมูลมีการเปลี่ยนแปลงอย่างรวดเร็ว ซึ่งเป็นผลมาจากการผสมผสานความสามารถในการจดจำรูปแบบข้อมูลอนุกรมเวลาของ LSTM กับความสามารถในการจัดการข้อมูลที่ไม่เป็นเชิงเส้นของ XGBoost

ผลการวิจัยนี้ เป็นพื้นฐานสำคัญสำหรับนำไปประยุกต์ใช้ในการพัฒนาระบบจัดการพลังงานอัจฉริยะและการวางแผนการผลิตไฟฟ้าจากพลังงานแสงอาทิตย์บนทุ่นลอยน้ำ รวมถึงระบบงานจริงที่ต้องการค่าพยากรณ์ที่มีความแม่นยำสูง ทั้งนี้ ในการนำไปใช้งานจริงอาจต้องพิจารณาจากความต้องการเฉพาะของระบบ ความรวดเร็วในการประมวลผล ข้อจำกัดด้านทรัพยากร และระดับความแม่นยำที่ต้องการ

7. ข้อเสนอแนะ

7.1 พิจารณาข้อมูลคุณลักษณะ (Feature) อื่น ๆ ที่เป็นปัจจัยหรือส่งผลต่อการพยากรณ์ เช่น ปริมาณรังสีดวงอาทิตย์ในระนาบของแผง (Plane of Array Irradiance) รวมถึงการเก็บรวบรวมข้อมูลในอดีตในปริมาณมาก จะช่วยให้แบบจำลองสามารถเรียนรู้ได้อย่างแม่นยำและประสิทธิภาพดียิ่งขึ้น

7.2 ควรมีการปรับปรุงไฮเปอร์พารามิเตอร์ในแบบจำลองการเรียนรู้ XGBoost และ LSTM โดยใช้วิธีการค้นหาที่เป็นระบบ เช่น Grid Search หรือ Bayesian Optimization และพิจารณาใช้ฟังก์ชันกระตุ้นและฟังก์ชันการสูญเสียชนิดอื่น ๆ เพื่อเปรียบเทียบและหาแบบจำลองที่มีประสิทธิภาพสูงสุด

7.3 ประยุกต์ใช้แบบจำลองการเรียนรู้เชิงลึกอื่น ๆ ที่มีความสามารถในการพยากรณ์ข้อมูลแบบลำดับหรืออนุกรมเวลา เช่น GRU หรือแบบจำลองที่มีกลไก Attention เพื่อเพิ่มความสามารถในการจับความสัมพันธ์ที่สำคัญในข้อมูล

7.4 แบบจำลองที่ได้จากการวิจัยนี้ สามารถนำไปพัฒนาเขียนเป็นโปรแกรมสำหรับระบบการพยากรณ์พลังงานแสงอาทิตย์บนทุ่นลอยน้ำแบบรายชั่วโมง หรือรายวัน เพื่อช่วยให้ผู้ควบคุมระบบ สามารถบริหารจัดการการผลิตพลังงานไฟฟ้าจากพลังงานแสงอาทิตย์ได้อย่างมีประสิทธิภาพ

8. เอกสารอ้างอิง

กรมพัฒนาพลังงานทดแทนและอนุรักษ์พลังงาน กระทรวงพลังงาน. (ม.ป.ป.). การจัดการก๊าซเรือนกระจก. เข้าถึงได้จาก

<https://www.dede.go.th/articles?id=191>

เกศินี ช่อนกลิ่น. (2566). แบบจำลองการพยากรณ์ราคาอนุพันธ์โดยใช้เทคนิคโครงข่ายประสาทเทียมแบบสังวัตนาการ 1 มิติ

หน่วยความจำระยะสั้นแบบยาวและซัพพอร์ตเวกตอร์รีเกรสชัน. (วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต). กรุงเทพฯ:

สถาบันเทคโนโลยีไทย-ญี่ปุ่น.

ภาณุพงศ์ น้ำแก้ว. (2563). การพยากรณ์กำลังไฟฟ้าของระบบเซลล์แสงอาทิตย์แบบติดตั้งอุปกรณ์วัดด้วยโครงข่าย

ประสาทเทียม. (วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต). กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย

วฤชัย ร่มสายหยุด. (2566). การเรียนรู้ของเครื่องสำหรับการวิเคราะห์ข้อมูลเชิงทำนายและการประยุกต์. นนทบุรี:

มหาวิทยาลัยสุโขทัยธรรมาธิราช

- ศรีรักษ์ ศรีทองชัย. (2565). แบบจำลองการพยากรณ์ค่า PM2.5 โดยใช้โครงข่ายประสาทเทียมแบบ LSTM ในพื้นที่ กรุงเทพมหานคร. *Journal of Engineering and Digital Technology (JEDT)*, 10(1), 1-9.
- สรุจภูมิ จิตตานนท์ และจักรกฤษ เต็มฤทธิกุล. (2566). การพยากรณ์การใช้พลังงานภายในอาคาร SGtech อย่างชาญฉลาด ด้วยการเรียนรู้เชิงลึก. (วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต). พิษณุโลก: มหาวิทยาลัยนเรศวร.
- สำนักนโยบายและแผนพลังงาน กระทรวงพลังงาน. (2561). แผนพัฒนากำลังผลิตไฟฟ้าของประเทศ (Power Development Plan : PDP). เข้าถึงได้จาก <https://www.eppo.go.th/index.php/th/electricity/pdp>
- อุรชา จันทรรภา. (2564). การพยากรณ์มูลค่าการส่งออกรถยนต์ อุปกรณ์ และส่วนประกอบของประเทศไทย. (วิทยานิพนธ์ วิศวกรรมศาสตรมหาบัณฑิต). กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย
- Berberich, D. (2020). *Hybrid Methods for Time Series Forecasting*. Retrieved from <https://www.inovex.de/de/blog/hybrid-time-series-forecasting/>
- Cazzaniga, R., & Rosa-Clot, M. (2021). The booming of floating PV. *Solar Energy*, 219, 3–10.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Foster, J., Liu, X., & McLoone, S. (2017). Adaptive sliding window load forecasting. *2017 28th Irish Signals and Systems Conference (ISSC)*, 1–6.
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.
- Jebli, I., Belouadha, F.-Z., Kabbaj, M. I., & Tilioua, A. (2021). Deep Learning based Models for Solar Energy Prediction. *Advances in Science, Technology and Engineering Systems Journal*, 6(1), 349–355.
- Li, C., Chen, Z., Liu, J., Li, D., Gao, X., Di, F., Li, L., & Ji, X. (2019). Power Load Forecasting Based on the Combined Model of LSTM and XGBoost. *Proceedings of the 2019 the International Conference on Pattern Recognition and Artificial Intelligence*, 46–51.
- Lim, S.-C., Huh, J.-H., Hong, S.-H., Park, C.-Y., & Kim, J.-C. (2022). Solar Power Forecasting Using CNN-LSTM Hybrid Model. *Energies*, 15(21), 8233.
- Liu, C.-H., Gu, J.-C., & Yang, M.-T. (2021). A Simplified LSTM Neural Networks for One Day-Ahead Solar Power Forecasting. *IEEE Access*, 9, 17174–17195.
- Obiora, C. N., Ali, A., & Hasan, A. N. (2021). Implementing Extreme Gradient Boosting (XGBoost) Algorithm in Predicting Solar Irradiance. *2021 IEEE PES/IAS PowerAfrica*, 1–5.
- Obiora, C. N., Ali, A., & Hasan, A. N. (2022). Efficient Method of Finding the Best Time Interval for Predicting Short Term Solar Radiation Using CNN-LSTM Model. *2022 11th International Conference on Power Science and Engineering (ICPSE)*, 159–163.
- Saxena, S. (2021, March 16). What is LSTM? Introduction to Long Short-Term Memory. Retrieved from <https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/>
- World Bank Group. (2018, OCTOBER 30). Where Sun Meets Water: Floating Solar Market Report. Retrieved from <https://www.worldbank.org/en/topic/energy/publication/where-sun-meets-water>
- Ye, H., Yang, B., Han, Y., & Chen, N. (2022). State-Of-The-Art Solar Energy Forecasting Approaches: Critical Potentials and Challenges. *Frontiers in Energy Research*, 10, 875790.

ประวัติผู้เขียนบทความ



นายอดิเรก ปั่นบุญ นักศึกษาหลักสูตรวิทยาศาสตรมหาบัณฑิต แขนงวิชาเทคโนโลยีดิจิทัล
สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช
โทรศัพท์ 08-6431-7255
E-Mail: adirek.pb@gmail.com
งานวิจัยที่สนใจ: Machine Learning for data analytics



รองศาสตราจารย์ ดร.วณิชยา ร่มสายหยุด อาจารย์ประจำหลักสูตรวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยสุโขทัยธรรมาธิราช
โทรศัพท์ 0-2504-8267
E-Mail: walisa.rom@stou.ac.th
ความเชี่ยวชาญ: Machine learning for data analytics, Big data management