



การประยุกต์ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยาน
แห่งชาติของประเทศไทย

นางสาวณัฐชานันท์ ลายเงินมณี

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ปีการศึกษา 2568

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

การประยุกต์ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยาน
แห่งชาติของประเทศไทย



นางสาวณัฐชานันท์ ลายเงินมณี

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาหลักสูตร

วิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาการคอมพิวเตอร์

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ปีการศึกษา 2568

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ



ใบรับรองโครงการค้นคว้าอิสระ

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง การประยุกต์ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำสำหรับการพยากรณ์จำนวนนักท่องเที่ยว
ในอุทยานแห่งชาติของประเทศไทย

โดย นางสาวณัฐชานันท์ ลายเงินมณี

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชา
วิทยาการคอมพิวเตอร์

คณบดีบัณฑิตวิทยาลัย / หัวหน้าภาควิชา

คณะกรรมการสอบการค้นคว้าอิสระ

.....

ประธานกรรมการ

(รองศาสตราจารย์ ดร.ธนภัทร์ อนุศาสน์อมรกุล)

.....

อาจารย์ที่ปรึกษา

(ผู้ช่วยศาสตราจารย์ ดร.ลือพล พิพานเมฆาภรณ์)

.....

กรรมการ

(ดร.มารุต บุณรัช)

.....

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.ลือพล พิพานเมฆาภรณ์)

ชื่อ : นางสาวณัฐชานันท์ ลายเงินมณี
 ชื่อการค้นคว้าอิสระ : การประยุกต์ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติของประเทศไทย
 สาขาวิชา : วิทยาการคอมพิวเตอร์
 มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
 อาจารย์ที่ปรึกษาการค้นคว้าอิสระหลัก : ผู้ช่วยศาสตราจารย์ ดร.ลือพล พิพานเมฆาภรณ์
 ปีการศึกษา : 2568

บทคัดย่อ

อุทยานแห่งชาติของประเทศไทยเป็นแหล่งท่องเที่ยวที่มีความสำคัญทั้งด้านการอนุรักษ์ทรัพยากรธรรมชาติและการส่งเสริมเศรษฐกิจของประเทศ ข้อมูลคาดการณ์ที่ใกล้เคียงความเป็นจริง โดยเฉพาะจำนวนนักท่องเที่ยว จึงมีบทบาทสำคัญต่อการวางแผนด้านบุคลากร การจัดสรรทรัพยากร และการให้บริการให้เหมาะสมกับความต้องการจริง ในทางปฏิบัติทั่วไปข้อมูลจำนวนนักท่องเที่ยวในอดีตมักถูกนำมาใช้ประกอบเพื่อสนับสนุนการตัดสินใจในการวางแผนและสนับสนุนการจัดสรรงบประมาณสำหรับการดูแลและบริหารจัดการอุทยาน หรือการดำเนินงานใด ๆ ตามภารกิจของหน่วยงาน อย่างไรก็ตาม วิธีการพยากรณ์ที่ใช้อยู่ในปัจจุบันมักอ้างอิงข้อมูลของปีที่ผ่านมาเป็นหลัก ซึ่งอาจไม่สะท้อนแนวโน้มที่ได้รับผลกระทบจากปัจจัยอื่น ๆ เช่น ฤดูกาล สภาพอากาศ เป็นต้น ซึ่งในงานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองสำหรับการพยากรณ์จำนวนนักท่องเที่ยวรายเดือนในอุทยานแห่งชาติของประเทศไทย โดยใช้ข้อมูลย้อนหลัง 10 ปี (พ.ศ. 2557 - 2566) จากอุทยานแห่งชาติ 146 แห่ง ด้วยการประยุกต์ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำ ได้แก่ RNN, LSTM และ GRU และเปรียบเทียบความสามารถของแบบจำลองในการคาดการณ์ข้อมูลอนุกรมเวลา โดยการทดลองแบ่งออกเป็น 3 แนวทางหลัก ได้แก่ แบบจำลองรายอุทยาน (IPM), แบบจำลองรวม (CM) และแบบจำลองแบบจัดกลุ่ม (CRM, CUM) โดยมีการทดสอบค่าพารามิเตอร์ Look-back และจำนวน Hidden Units เพื่อเลือกค่าที่เหมาะสมในแต่ละแนวทาง นอกจากนี้ ได้มีการแบ่งการทดลองเป็น 2 ระยะ โดยระยะที่ 1 ใช้เพียงข้อมูลจำนวนนักท่องเที่ยวเป็นตัวแปรเดียว ส่วนระยะที่ 2 ได้เพิ่มตัวแปรเพิ่มเติม ได้แก่ เดือน อุณหภูมิ ปริมาณรถยนต์ และข้อมูลเชิงพื้นที่ ประกอบด้วย ภูมิภาค และลักษณะเด่นของอุทยาน เพื่อประเมินผลกระทบของปัจจัยที่หลากหลายต่อคุณภาพการพยากรณ์ ทั้งนี้ ยังมีการเปรียบเทียบ

กับเทคนิคที่นิยมใช้ในการพยากรณ์ข้อมูลในลักษณะคล้ายกัน และการวิเคราะห์ความสำคัญของกลุ่มตัวแปรต่าง ๆ เพื่อค้นหาข้อมูลที่ส่งผลต่อประสิทธิภาพของแบบจำลองมากที่สุด ซึ่งผลการวิจัยพบว่า แบบจำลองที่ให้ผลลัพธ์ที่ดีที่สุดในงานวิจัยนี้ คือ CM-LSTM ซึ่งเป็นแบบจำลองรวมที่ใช้ LSTM ด้วยการกำหนดค่า Look-back เท่ากับ 9 เดือน และ Hidden Units เท่ากับ 40 หน่วย ให้ค่าตัวชี้วัดที่ดีที่สุดทั้ง MAE (Mean_Test) = 1,955, RMSE (Mean_Test) = 4,269 และ R^2 (Mean_Test) = 0.95 ซึ่งสูงกว่าทั้งแบบจำลองจากระยะที่ 1 และเทคนิคพยากรณ์อื่น ๆ แม้ว่าการเพิ่มตัวแปรอื่น ๆ ในระยะที่ 2 จะไม่ได้ทำให้ค่า R^2 สูงขึ้นมากนัก แต่ส่งผลให้ค่า MAE และ RMSE ลดลงอย่างชัดเจน ซึ่งแสดงถึงการเปลี่ยนแปลงในทางที่ดีขึ้น ในส่วนของการวิเคราะห์ความสำคัญของตัวแปร พบว่า กลุ่มที่มีข้อมูลจำนวนนักท่องเที่ยวเป็นพื้นฐาน และเสริมด้วยข้อมูลเดือน หรือ ข้อมูลอุณหภูมิ มักให้ค่าตัวชี้วัดที่ดีต่อเนื่อง ในขณะที่ตัวแปรเชิงพื้นที่และข้อมูลปริมาณรถยนต์ให้ผลที่หลากหลาย แต่ยังไม่สามารถช่วยให้ผลลัพธ์ดีขึ้นเท่าที่ควร อย่างไรก็ตาม การรวมข้อมูลหลายกลุ่มหรือทุกตัวแปรเข้าด้วยกันมีแนวโน้มช่วยให้แบบจำลองเข้าใจรูปแบบและความสัมพันธ์ของข้อมูลได้ดีและครอบคลุมมากยิ่งขึ้น

(มีจำนวนทั้งสิ้น 115 หน้า)

คำสำคัญ : การพยากรณ์จำนวนนักท่องเที่ยว, อนุกรมเวลา, โครงข่ายประสาทเทียมแบบเรียกซ้ำ, อาร์เอ็นเอ็น, แอลเอสทีเอ็ม, จีอาร์ยู

อาจารย์ที่ปรึกษาการค้นคว้าอิสระหลัก

Name : Miss Natchanan Lai-ngoen-manee
Independent Study Title : The Application of Recurrent Neural Networks for
Forecasting the Number of Tourists in Thailand's
National Parks
Major Field : Computer Science
King Mongkut's University of Technology North
Bangkok
Independent Study Advisor : Assistant Professor Luepol Pipanmekaporn, Ph.D.
Academic Year : 2025

ABSTRACT

Thailand's national parks are significant tourist destinations, serving both natural resource conservation and the promotion of the country's economy. Accurate forecasting data, especially tourist numbers, play a crucial role in personnel planning, resource allocation, and providing services that meet actual demand. In practice, historical tourist count data are commonly used to support decision-making for planning and budget allocation for park management and operations aligned with the agency's mission. However, the forecasting methods currently used typically rely primarily on data from the previous year, which may not reflect trends affected by other factors such as seasonality and weather conditions. This research aims to develop models for forecasting monthly tourist numbers in Thailand's national parks using 10 years of historical data (2014-2023) from 146 national parks, applying recurrent neural networks (RNN, LSTM, and GRU), and to compare the forecasting capabilities of these models on time series data. The experiments were divided into three main approaches: Individual Park Model (IPM), Combined Model (CM), and Clustered Models (CRM, CUM), with parameter tuning of look-back periods and the number of hidden units to determine optimal values for each approach. In addition, the experiments were conducted in two phases: Phase 1 used only tourist count data as a single variable, while Phase 2 included additional variables-month, temperature, vehicle count, and spatial data (region and park characteristics)-to assess the impact of diverse factors on

forecasting quality. The study also compared popular forecasting techniques for similar data and analyzed the importance of different groups of variables to identify those that most influence model performance. The results show that the best-performing model was the CM-LSTM, a combined LSTM model with a look-back period of 9 months and 40 hidden units, achieving the best evaluation metrics with Mean_Test_MAE = 1,955, Mean_Test_RMSE = 4,269, and Mean_Test_R² = 0.95, outperforming both Phase 1 models and other forecasting techniques. Although the inclusion of additional variables in Phase 2 did not substantially increase the R² value, it led to a clear reduction in MAE and RMSE, indicating an overall improvement. The variable importance analysis revealed that groups based on tourist count, supplemented by month or temperature, consistently yielded strong metrics, whereas spatial variables and vehicle count produced more varied results and did not significantly enhance performance. Nonetheless, combining multiple groups or all variables tended to help the model better capture data patterns and relationships more comprehensively.

(Total 115 Pages)

Keywords: Tourist Forecasting, Time Series, Recurrent Neural Networks, RNN, LSTM, GRU

Advisor

กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จลุล่วงตามวัตถุประสงค์ด้วยคำแนะนำและความช่วยเหลือเป็นอย่างดีจากผู้ช่วยศาสตราจารย์ ดร.ลือพล พิพานเมฆาภรณ์ อาจารย์ที่ปรึกษางานวิจัย ซึ่งท่านได้สละเวลาให้คำปรึกษา แนะนำ ตรวจสอบ และดูแล ตลอดระยะเวลาการดำเนินงาน ทั้งในด้านเนื้อหา แนวทางการวิจัย และการปรับปรุงงานวิจัยให้มีความสมบูรณ์ในทุก ๆ ขั้นตอน จนกระทั่งงานวิจัยฉบับนี้สำเร็จลุล่วงและได้จัดทำเป็นรูปเล่มที่สมบูรณ์ ผู้วิจัยจึงขอขอบพระคุณ อาจารย์เป็นอย่างสูง

ขอขอบพระคุณคณาจารย์ สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์ และสารสนเทศ ที่ได้ถ่ายทอดความรู้ทางด้านวิทยาการคอมพิวเตอร์และเสริมสร้างแนวคิด ทำให้เกิดการพัฒนาความรู้ความเข้าใจและนำไปประยุกต์ใช้ในการศึกษาค้นคว้าจนสำเร็จลุล่วง

ขอขอบคุณเจ้าหน้าที่ที่ดูแล ช่วยเหลือ และประสานงานด้านต่าง ๆ ตลอดหลักสูตร การศึกษา รวมถึงเพื่อน ๆ ที่ได้ให้คำปรึกษา แนะนำ ซึ่งเป็นส่วนหนึ่งที่ทำให้งานวิจัยนี้สำเร็จลุล่วงไปด้วยดี

ณัฐชานันท์ ลายเงินมณี

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	ฉ
กิตติกรรมประกาศ	ช
สารบัญ.....	ฌ
สารบัญตาราง (ถ้ามี).....	ฎ
สารบัญรูปภาพ (ถ้ามี).....	ฐ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของการวิจัย	2
1.3 ขอบเขตของการวิจัย	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	4
2.1 การพยากรณ์ข้อมูลเวลา (Time Series Forecasting)	4
2.2 โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Networks: RNNs)	5
2.3 หน่วยความจำระยะยาวและระยะสั้น (Long Short-Term Memory: LSTM)	6
2.4 หน่วยระยะเวลาล้วน (Gated Recurrent Unit: GRU).....	8
2.5 ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย Mean Absolute Error (MAE)	10
2.6 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรากที่สอง Root Mean Squared Error (RMSE)	11
2.7 ค่าสัมประสิทธิ์การตัดสินใจ Coefficient of Determination (R^2).....	12

2.8 งานวิจัยที่เกี่ยวข้อง	13
บทที่ 3 วิธีการดำเนินการวิจัย	16
3.1 ภาพรวมขั้นตอนวิธีดำเนินการวิจัย	16
3.2 กระบวนการรวบรวมและเตรียมข้อมูล	17
3.3 เครื่องมือที่ใช้ดำเนินการ	18
3.4 การออกแบบการทดลอง	18
3.5 การกำหนดพารามิเตอร์ที่สำคัญ	19
3.6 การประเมินผล	19
3.7 แนวทางการสรุปและวิเคราะห์ผลการทดลอง	20
บทที่ 4 ผลการวิจัย	22
4.1 ผลการทดลอง ระยะที่ 1	22
4.2 ผลการทดลอง ระยะที่ 2	53
4.3 สรุปผลการทดลองระหว่าง ระยะที่ 1 และ ระยะที่ 2	83
4.4 การเปรียบเทียบผลการทดลองกับเทคนิคอื่น ๆ	89
4.5 การวิเคราะห์ผลกระทบของพีเจอร์	92
4.6 การพัฒนาเว็บแอปพลิเคชันเพื่อการใช้งาน	94
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	96
5.1 สรุปผลการวิจัย	96
5.2 ข้อเสนอแนะสำหรับงานวิจัย	96
บรรณานุกรม	98
ภาคผนวก ก	102
ภาคผนวก ข	106
ประวัติผู้เขียน	115

สารบัญตาราง (ถ้ามี)

หน้า

ตารางที่ 4 - 1 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 1).....	23
ตารางที่ 4 - 2 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1)	25
ตารางที่ 4 - 3 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1)	28
ตารางที่ 4 - 4 ผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 1).....	30
ตารางที่ 4 - 5 ผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 1).....	33
ตารางที่ 4 - 6 ผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 1)	36
ตารางที่ 4 - 7 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 1).....	38
ตารางที่ 4 - 8 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 1)....	41
ตารางที่ 4 - 9 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 1).....	43
ตารางที่ 4 - 10 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 1).....	45
ตารางที่ 4 - 11 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 1).....	48
ตารางที่ 4 - 12 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 1).....	51
ตารางที่ 4 - 13 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 2)	53
ตารางที่ 4 - 14 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 2)	56
ตารางที่ 4 - 15 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2)	58
ตารางที่ 4 - 16 ผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 2).....	60
ตารางที่ 4 - 17 ผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 2).....	63
ตารางที่ 4 - 18 ผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 2).....	65
ตารางที่ 4 - 19 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 2)....	68
ตารางที่ 4 - 20 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 2) .	71
ตารางที่ 4 - 21 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 2)....	73
ตารางที่ 4 - 22 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 2).....	75
ตารางที่ 4 - 23 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 2).....	78

ตารางที่ 4 - 24 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 2).....	81
ตารางที่ 4 - 25 สรุปผลการทดลองของระยะที่ 1.....	83
ตารางที่ 4 - 26 สรุปผลการทดลองของระยะที่ 2.....	85
ตารางที่ 4 - 27 ผลการเปรียบเทียบแบบจำลองที่คัดเลือกจากระยะที่ 1 และระยะที่ 2	88
ตารางที่ 4 - 28 การเปรียบเทียบประสิทธิภาพของแบบจำลองที่พัฒนากับเทคนิคทั่วไป	89
ตารางที่ 4 - 29 รายละเอียดและความหมายของพีเจอร์.....	92
ตารางที่ 4 - 30 ผลการวิเคราะห์ผลกระทบของพีเจอร์ต่อประสิทธิภาพของแบบจำลอง.....	92
ตาราง ข - 1 ข้อมูลรายชื่ออุทยานแห่งชาติ.....	106
ตาราง ข - 2 คำอธิบายชุดข้อมูลจำนวนนักท่องเที่ยว	111
ตาราง ข - 3 คำอธิบายชุดข้อมูลยานพาหนะ	112
ตาราง ข - 4 คำอธิบายชุดข้อมูลสภาพอากาศ.....	112
ตาราง ข - 5 คำอธิบายชุดข้อมูลเชิงพื้นที่ (ภูมิภาคและลักษณะของอุทยาน)	114

สารบัญรูปภาพ (ถ้ามี)

หน้า

ภาพที่ 2 - 1 โครงสร้างภายในของแบบจำลอง Recurrent Neural Network (RNN) แบบพื้นฐาน (Ibrahim et al., 2023).....	6
ภาพที่ 2 - 2 โครงสร้างภายในของแบบจำลอง Long Short-Term Memory (LSTM) (Ibrahim et al., 2023).....	8
ภาพที่ 2 - 3 โครงสร้างภายในของแบบจำลอง Gated Recurrent Unit (GRU) (Ibrahim et al., 2023).....	10
ภาพที่ 3 - 1 ภาพรวมขั้นตอนวิธีดำเนินการวิจัย.....	16
ภาพที่ 4 - 1 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 1).....	24
ภาพที่ 4 - 2 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 1).....	24
ภาพที่ 4 - 3 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 1).....	25
ภาพที่ 4 - 4 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1).....	26
ภาพที่ 4 - 5 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1).....	27
ภาพที่ 4 - 6 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1).....	27
ภาพที่ 4 - 7 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1).....	29
ภาพที่ 4 - 8 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1).....	29
ภาพที่ 4 - 9 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1).....	30
ภาพที่ 4 - 10 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 1).....	32

ภาพที่ 4 - 11 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 1).....	32
ภาพที่ 4 - 12 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 1).....	32
ภาพที่ 4 - 13 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 1).....	34
ภาพที่ 4 - 14 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 1).....	35
ภาพที่ 4 - 15 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 1).....	35
ภาพที่ 4 - 16 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 1).....	37
ภาพที่ 4 - 17 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 1).....	37
ภาพที่ 4 - 18 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 1).....	37
ภาพที่ 4 - 19 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ RNN (Phase 1).....	39
ภาพที่ 4 - 20 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ RNN (Phase 1).....	40
ภาพที่ 4 - 21 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ RNN (Phase 1).....	40
ภาพที่ 4 - 22 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ LSTM (Phase 1).....	42
ภาพที่ 4 - 23 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ LSTM (Phase 1).....	42
ภาพที่ 4 - 24 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ LSTM (Phase 1).....	42
ภาพที่ 4 - 25 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ GRU (Phase 1).....	44

ภาพที่ 4 - 41 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 2).....	57
ภาพที่ 4 - 42 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 2).....	58
ภาพที่ 4 - 43 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2).....	59
ภาพที่ 4 - 44 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2).....	60
ภาพที่ 4 - 45 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2).....	60
ภาพที่ 4 - 46 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 2).....	62
ภาพที่ 4 - 47 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 2).....	62
ภาพที่ 4 - 48 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 2).....	62
ภาพที่ 4 - 49 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 2).....	64
ภาพที่ 4 - 50 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 2).....	64
ภาพที่ 4 - 51 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 2).....	65
ภาพที่ 4 - 52 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 2).....	67
ภาพที่ 4 - 53 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 2).....	67
ภาพที่ 4 - 54 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 2).....	67
ภาพที่ 4 - 55 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตามภูมิภาค โดยใช้ RNN (Phase 2).....	69

[illegible]

ภาพที่ 4 - 71 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตาม ลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 2).....	82
ภาพที่ 4 - 72 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่ม ตาม ลักษณะเด่นของแต่ละอุทยาน โดยใช้แบบจำลอง GRU (Phase 2)	83
ภาพที่ 4 - 73 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ MAE จากการสรุปผลการทดลองของระยะที่ 1	84
ภาพที่ 4 - 74 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ RMSE จากการสรุปผลการทดลองของระยะที่ 1	84
ภาพที่ 4 - 75 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ R^2 จากการสรุปผลการทดลองของระยะที่ 1	84
ภาพที่ 4 - 76 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ MAE จากการสรุปผลการทดลองของระยะที่ 2	86
ภาพที่ 4 - 77 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ RMSE จากการสรุปผลการทดลองของระยะที่ 2	86
ภาพที่ 4 - 78 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ R^2 จากการสรุปผลการทดลองของระยะที่ 2	86
ภาพที่ 4 - 79 กราฟเปรียบเทียบค่า MAE ระหว่างแบบจำลองที่พัฒนากับเทคนิคทั่วไป	90
ภาพที่ 4 - 80 กราฟเปรียบเทียบค่า RMSE ระหว่างแบบจำลองที่พัฒนากับเทคนิคทั่วไป	90
ภาพที่ 4 - 81 กราฟเปรียบเทียบค่า R^2 ระหว่างแบบจำลองที่พัฒนากับเทคนิคทั่วไป	91
ภาพที่ 4 - 82 ตัวอย่างการแสดงกราฟแนวโน้มและตารางผลลัพธ์ของการพยากรณ์จำนวน นักท่องเที่ยวผ่านเว็บแอปพลิเคชัน.....	95
ภาพ ก - 1 ส่วนประกอบหลักของหน้าเว็บแอปพลิเคชัน	102
ภาพ ก - 2 กำหนดพารามิเตอร์การพยากรณ์	103
ภาพ ก - 3 เริ่มกระบวนการพยากรณ์.....	103
ภาพ ก - 4 กราฟภาพรวมเฉลี่ยทุกอุทยาน	104
ภาพ ก - 5 ตารางข้อมูลค่าเฉลี่ย.....	104
ภาพ ก - 6 ตารางพยากรณ์ (รายอุทยาน).....	104
ภาพ ก - 7 ดาวน์โหลดข้อมูล	105

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

อุทยานแห่งชาติของประเทศไทยเป็นแหล่งท่องเที่ยวที่มีความสำคัญทั้งในด้านการอนุรักษ์ทรัพยากรธรรมชาติและการส่งเสริมเศรษฐกิจของประเทศ ในแต่ละปีมีนักท่องเที่ยวจำนวนมากเดินทางเข้ามาที่อุทยานแห่งชาติ ทั้งชาวไทยและชาวต่างชาติ การบริหารจัดการอุทยานแห่งชาติจึงต้องอาศัยข้อมูลประกอบต่าง ๆ เพื่อสนับสนุนการวางแผนการบริหารจัดการ โดยเฉพาะข้อมูลเกี่ยวกับปริมาณนักท่องเที่ยวที่คาดว่าจะเดินทางมาในแต่ละช่วงเวลา การคาดการณ์จำนวนนักท่องเที่ยวเป็นหนึ่งในข้อมูลสำคัญที่ช่วยสนับสนุนการบริหารจัดการอุทยานแห่งชาติให้มีประสิทธิภาพมากขึ้น ข้อมูลนี้สามารถนำไปสนับสนุนในการวางแผนด้านบุคลากร การจัดสรรทรัพยากร และการให้บริการ เพื่อให้สามารถรองรับนักท่องเที่ยวได้อย่างเหมาะสม หากไม่มีการคาดการณ์ที่ดีและใกล้เคียง อาจส่งผลให้เกิดปัญหาทางด้านการบริหารจัดการ เช่น การขาดแคลนเจ้าหน้าที่ดูแลอุทยานในช่วงที่มีนักท่องเที่ยวจำนวนมาก หรือการใช้ทรัพยากรเกินความจำเป็นในช่วงที่มีนักท่องเที่ยวน้อยลง ซึ่งอาจส่งผลต่อประสิทธิภาพในการให้บริการและการดูแลรักษาสิ่งแวดล้อมของอุทยานได้ นอกจากนี้ ข้อมูลจำนวนนักท่องเที่ยวช่วงเวลาที่ผ่านมาเป็นหนึ่งในปัจจัยที่ใช้พิจารณาประกอบการประมาณการด้านงบประมาณ โดยอุทยานที่มีนักท่องเที่ยวจำนวนมากอาจได้รับการจัดสรรงบประมาณมากขึ้นเพื่อรองรับการดำเนินงาน ขณะที่อุทยานที่มีนักท่องเที่ยวจำนวนน้อยอาจได้รับงบประมาณที่จำกัดลง ดังนั้น การมีข้อมูลพยากรณ์ที่ใกล้เคียงกับความเป็นจริงสามารถเป็นเครื่องมือหนึ่งในการช่วยสนับสนุนการประมาณการงบประมาณมีความเหมาะสมกับสภาพการณ์จริง ซึ่งวิธีการคาดการณ์จำนวนนักท่องเที่ยวมักอาศัยข้อมูลนักท่องเที่ยวจากปีก่อนเป็นเกณฑ์สำหรับการคาดการณ์ในปีถัดไปซึ่งเป็นแนวทางที่ใช้กันทั่วไปในเชิงสถิติ อย่างไรก็ตาม วิธีการดังกล่าวอาจไม่สามารถสะท้อนถึงแนวโน้มที่แท้จริงเท่าที่ควร เนื่องจากจำนวนนักท่องเที่ยวอาจได้รับอิทธิพลจากปัจจัยอื่น เช่น ฤดูกาล สภาพอากาศ เป็นต้น การใช้ข้อมูลนักท่องเที่ยวจากอดีตเพียงอย่างเดียวอาจทำให้เกิดความคลาดเคลื่อนไปมากในกระบวนการคาดการณ์ การพัฒนาแบบจำลองที่สามารถเรียนรู้และจับแนวโน้มความสัมพันธ์ของข้อมูลเชิงลำดับเวลาได้ดี จึงน่าจะเป็นแนวทางที่มีศักยภาพมากกว่าวิธีแบบเดิม โดยช่วงที่ผ่านมามีการนำเทคนิคทางสถิติ เทคนิคทาง Machine Learning และเทคนิคทาง Deep Learning มาประยุกต์ใช้ในการพยากรณ์ข้อมูลอนุกรมเวลา ซึ่งพบว่าการใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำ (Recurrent Neural Networks : RNN) และแบบจำลองที่พัฒนา

ต่อยอดจาก RNN เช่น LSTM (Long Short-Term Memory) และ GRU (Gated Recurrent Unit) สามารถจัดการกับความซับซ้อนของข้อมูลที่มีลำดับเวลาได้อย่างมีประสิทธิภาพ

งานวิจัยนี้จึงเน้นไปที่การพัฒนาและทดสอบแบบจำลอง RNN, LSTM และ GRU สำหรับการพยากรณ์จำนวนนักท่องเที่ยวรายเดือนในอุทยานแห่งชาติ จำนวน 146 แห่งของประเทศไทย โดยใช้ข้อมูลย้อนหลัง 10 ปี (พ.ศ. 2557 - พ.ศ. 2566) เพื่อตรวจสอบว่าแบบจำลองใดให้ประสิทธิภาพดีที่สุดในการคาดการณ์จำนวนนักท่องเที่ยว ทั้งนี้ ผลลัพธ์จากการพัฒนาและทดสอบแบบจำลอง สามารถแสดงให้เห็นถึงศักยภาพของแบบจำลองและแนวโน้มที่สามารถนำไปประยุกต์ใช้จริง ซึ่งจะเป็นเครื่องมือหนึ่งในการสนับสนุนการบริหาร การจัดสรรทรัพยากร และการให้บริการเพื่อรองรับการท่องเที่ยวของอุทยานแห่งชาติได้

1.2 วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาแบบจำลองที่มีประสิทธิภาพในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติของประเทศไทยได้

1.3 ขอบเขตของการวิจัย

1.3.1 ข้อมูลจำนวนนักท่องเที่ยวในอุทยานแห่งชาติของประเทศไทย จำนวน 146 แห่ง ย้อนหลัง 10 ปี (พ.ศ. 2557 - พ.ศ. 2566)

1.3.2 เปรียบเทียบประสิทธิภาพของแบบจำลองที่ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำ (RNN, LSTM, GRU)

1.3.3 ประเมินผลด้วยตัวชี้วัดประสิทธิภาพ MAE (Mean Absolute Error), RMSE (Root Mean Squared Error) และ R^2 (Coefficient of Determination)

1.3.4 ทดสอบแบบจำลองสำหรับการพยากรณ์จำนวนนักท่องเที่ยว ด้วย 3 แนวทาง ได้แก่ แบบจำลองรายอุทยาน แบบจำลองรวม และแบบจำลองแบบจัดกลุ่ม (กลุ่มตามภูมิภาค และกลุ่มตามลักษณะเด่นของอุทยาน)

1.3.5 ทดสอบการใช้ชุดข้อมูล (Feature) ที่หลากหลายเพื่อประเมินผลต่อประสิทธิภาพของแบบจำลอง ได้แก่ ชุดข้อมูลจำนวนนักท่องเที่ยว (TouristCount), ชุดข้อมูลเดือน (Month), ชุดข้อมูลอุณหภูมิ (Temperature), ชุดข้อมูลปริมาณรถยนต์ (VehicleCount) และชุดข้อมูลเชิงพื้นที่ของอุทยาน (AreaFeatures) แบ่งออกเป็น 2 กลุ่ม (1) กลุ่มตามภูมิภาค (ภาคเหนือ, ภาคตะวันออกเฉียงเหนือ, ภาคกลาง, ภาคตะวันออก, ภาคตะวันตก, ภาคใต้) และ (2) กลุ่มตามลักษณะเด่น (กลุ่มภูเขา, กลุ่มน้ำตก, กลุ่มวัฒนธรรม, กลุ่มทะเล)

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1.4.1 สามารถพัฒนาแบบจำลองที่สามารถพยากรณ์จำนวนนักท่องเที่ยวได้

1.4.2 สามารถเป็นเครื่องมือหนึ่งในการสนับสนุนการบริหาร การจัดสรรทรัพยากร และการให้บริการ เพื่อรองรับการท่องเที่ยวของอุทยานแห่งชาติได้



บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองที่มีประสิทธิภาพในการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติของประเทศไทย และสามารถเป็นเครื่องมือสนับสนุนการตัดสินใจของผู้บริหารในด้านการวางแผนและการบริหารจัดการ เพื่อรองรับการท่องเที่ยวในแต่ละอุทยานแห่งชาติได้ โดยในแต่ละขั้นตอนประกอบด้วยทฤษฎีและงานวิจัยที่เกี่ยวข้อง ดังนี้

- 2.1 การพยากรณ์ข้อมูลเวลา (Time Series Forecasting)
- 2.2 โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Networks: RNNs)
- 2.3 หน่วยความจำระยะยาวและระยะสั้น (Long Short-Term Memory: LSTM)
- 2.4 หน่วยระยะเวลาดั้ง (Gated Recurrent Unit: GRU)
- 2.5 ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย Mean Absolute Error (MAE)
- 2.6 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรากที่สอง Root Mean Squared Error (RMSE)
- 2.7 ค่าสัมประสิทธิ์ความกำหนด Coefficient of Determination (R^2)
- 2.8 งานวิจัยที่เกี่ยวข้อง

2.1 การพยากรณ์ข้อมูลเวลา (Time Series Forecasting)

การพยากรณ์ข้อมูลเวลา (Time Series Forecasting) คือ กระบวนการวิเคราะห์และคาดการณ์ค่าของข้อมูลในอนาคต โดยอาศัยข้อมูลที่จัดเก็บต่อเนื่องตามลำดับเวลา เช่น รายวัน รายเดือน หรือรายปี ซึ่งมีการเปลี่ยนแปลงและเชื่อมโยงกันในเชิงเวลา (Box & Jenkins, 1976) ข้อมูลประเภทนี้มักพบรูปแบบ (Pattern) ที่หลากหลาย ไม่ว่าจะเป็น แนวโน้มระยะยาว (Trend), ฤดูกาล (Seasonality) หรือ ความผันผวนอันเกิดจากปัจจัยเฉพาะ (Irregular Variation) ทำให้การพยากรณ์ข้อมูลเวลาแตกต่างจากการคาดการณ์ข้อมูลแบบตาราง (Tabular Data) ที่ไม่ได้เรียงต่อกันตามลำดับเวลา (Wichern et al., 1979)

เทคนิคในการพยากรณ์ข้อมูลเวลามีพัฒนาการอย่างต่อเนื่อง ตั้งแต่วิธีการทางสถิติแบบดั้งเดิม เช่น การปรับเรียบเอ็กซ์โพเนนเชียล (Exponential Smoothing) และ แบบจำลองบ็อกซ์-เจนกินส์ (Box-Jenkins Model) ซึ่งรวมถึง ARIMA (AutoRegressive Integrated Moving Average) ที่เน้นโครงสร้างสมการทางคณิตศาสตร์ในการจับสัญญาณแนวโน้มและฤดูกาล ไปจนถึงวิธีการเรียนรู้ของเครื่อง (Machine Learning) และการเรียนรู้เชิงลึก (Deep Learning) ที่สามารถค้นหารูปแบบที่ซับซ้อนในข้อมูลได้อย่างมีประสิทธิภาพ (Goodfellow Ian, Bengio Yoshua, 2016; Wichern et al., 1979)

ในหลายสาขาที่ต้องพึ่งพาข้อมูลเวลาเป็นหลัก การเลือกใช้แบบจำลองขึ้นอยู่กับลักษณะและความผันผวนของข้อมูลเป็นสำคัญ หากข้อมูลมีความต่อเนื่องหรือมีรูปแบบที่เด่นชัด แบบจำลองสถิติแบบง่ายอาจเพียงพอ แต่หากข้อมูลมีความผันผวนสูง หรือได้รับอิทธิพลจากปัจจัยภายนอกหลายมิติ เช่น สภาพเศรษฐกิจ ฤดูกาล หรือสภาพภูมิอากาศ อาจต้องอาศัยแบบจำลองที่มีความสามารถในการเรียนรู้เชิงลึก เช่น โครงข่ายประสาทเทียมแบบ Recurrent (RNN) หรืออนุกรมย่อยอย่าง LSTM และ GRU ซึ่งได้รับการออกแบบให้จดจำความเชื่อมโยงระหว่างค่าข้อมูลในอดีตและค่าข้อมูลในอนาคตได้ดีขึ้น (Goodfellow et al., Bengio Yoshua, 2016)

ในบริบทของอุตสาหกรรมการท่องเที่ยว งานวิจัยหลายชิ้นระบุว่าข้อมูลเวลาอาจสะท้อนทั้งฤดูกาลท่องเที่ยว (High/Low Season) และเทศกาลสำคัญ ซึ่งอาจส่งผลต่อพฤติกรรมการเดินทางได้อย่างชัดเจน ตัวอย่างเช่น การเพิ่มขึ้นของนักท่องเที่ยวในช่วงวันหยุดยาวหรือเทศกาลประจำปี (Kulendran & Witt, 2001) ด้วยเหตุนี้ การพยากรณ์ข้อมูลเวลาในภาคการท่องเที่ยวจึงมีบทบาทสำคัญต่อการวางแผนจัดสรรทรัพยากรและการบริหารจัดการ ทั้งในระดับองค์กรและระดับนโยบาย

การพยากรณ์ข้อมูลเวลาเป็นแนวทางสำคัญในการวิเคราะห์และตัดสินใจสำหรับข้อมูลที่มีลักษณะเชิงเวลาอย่างต่อเนื่อง ความเข้าใจพื้นฐานเกี่ยวกับรูปแบบข้อมูล เวลา (Time Dimension) และวิธีการพยากรณ์หลากหลายรูปแบบ จึงเป็นพื้นฐานที่นักวิจัยส่วนใหญ่มักเลือกใช้ในสาขาต่าง ๆ จำเป็นต้องศึกษาเพื่อเลือกแบบจำลองให้สอดคล้องกับธรรมชาติของข้อมูลและจุดประสงค์ของงานแต่ละด้าน

2.2 โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Networks: RNNs)

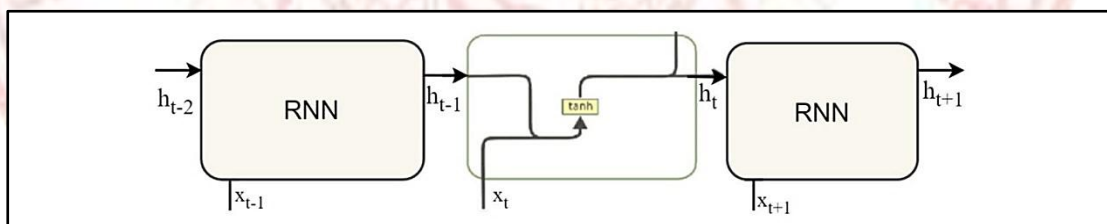
โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Networks หรือ RNNs) เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่ถูกออกแบบมาโดยเฉพาะสำหรับการประมวลผลข้อมูลในรูปแบบลำดับ (Sequential Data) เช่น ข้อมูลอนุกรมเวลา (Time Series Data) ข้อมูลข้อความ หรือข้อมูลเสียง ที่มีลำดับก่อนหลัง โดย RNN นั้นมีโครงสร้างพิเศษที่สามารถจดจำข้อมูลในอดีตเพื่อส่งผ่านข้อมูลเหล่านั้นไปยังขั้นตอนถัดไป ทำให้มีประสิทธิภาพในการจัดการกับข้อมูลที่มีความสัมพันธ์เชิงเวลาอย่างซับซ้อน ในขณะที่โครงข่ายประสาทเทียมแบบ Feedforward (Feedforward Neural Networks) จะประมวลผลข้อมูลในทิศทางเดียว จากชั้นอินพุต (input layer) ไปยังชั้นเอาต์พุต (output layer) โดยไม่มีการวนซ้ำของข้อมูลกลับมาใช้อีก แต่โครงข่ายประสาทเทียมแบบ RNN จะมีการป้อนผลลัพธ์จากชั้นก่อนหน้าย้อนกลับมาใช้เป็นข้อมูลนำเข้าในขั้นตอนถัดไป (Feedback loop) หรือเรียกว่า recurrent connections ซึ่งทำให้แบบจำลองสามารถเก็บข้อมูลเกี่ยวกับสถานะก่อนหน้าไว้ได้ (previous states) การทำงานของ RNN โดยทั่วไปจะสามารถอธิบายได้ด้วยสมการต่อไปนี้ (Goodfellow et al., Bengio Yoshua, 2016)

$$\begin{aligned}
 a^{(t)} &= b + Wh^{(t-1)} + Ux^{(t)} \\
 h^{(t)} &= \tanh(a^{(t)}) \\
 \hat{y}^{(t)} &= c + Vh^{(t)}
 \end{aligned}
 \tag{2-1}$$

โดยที่

$x^{(t)}$	คือ	ข้อมูลนำเข้าของแบบจำลอง ณ เวลาปัจจุบัน (input at time t)
$h^{(t)}$	คือ	hidden state หรือสถานะแฝงของโครงข่ายในขั้นตอนเวลา t
$a^{(t)}$	คือ	การคำนวณผลลัพธ์ก่อนเข้าสู่ activation function
$\hat{y}^{(t)}$	คือ	ผลลัพธ์หรือการทำนาย ณ เวลา t
U, W, V	คือ	เมทริกซ์น้ำหนัก (Weight matrices) ที่ใช้ในการฝึกแบบจำลอง

ในการฝึกแบบจำลอง RNN วิธีการเรียนรู้ที่นิยมใช้กันทั่วไปคือ Backpropagation Through Time (BPTT) ซึ่งเป็นการขยายแนวคิดการเรียนรู้แบบย้อนกลับ (Backpropagation) ที่ใช้ในแบบจำลองทั่วไป โดย BPTT จะคำนวณ gradient ย้อนเวลากลับไปหลายๆ ขั้นตอนเพื่ออัปเดตพารามิเตอร์ให้เหมาะสมที่สุด อย่างไรก็ตาม RNN แบบพื้นฐานมีข้อจำกัดที่สำคัญคือ ปัญหาการสูญเสียวัดหรือระเบิดของค่า Gradient (Vanishing and Exploding Gradient Problem) ซึ่งเป็นปัญหาที่พบได้บ่อยในข้อมูลลำดับยาว ทำให้เกิดการพัฒนารูปแบบโครงข่ายที่ต่อยอดจาก RNN เช่น Long Short-Term Memory (LSTM) และ Gated Recurrent Unit (GRU) เพื่อแก้ไขปัญหานี้โดยเฉพาะ (Goodfellow lan, Bengio Yoshua, 2016)



ภาพที่ 2 - 1 โครงสร้างภายในของแบบจำลอง Recurrent Neural Network (RNN) แบบพื้นฐาน
(Ibrahim et al., 2023)

2.3 หน่วยความจำระยะยาวและระยะสั้น (Long Short-Term Memory: LSTM)

หน่วยความจำระยะยาวและระยะสั้น หรือที่นิยมเรียกกันโดยทั่วไปว่า LSTM เป็นหนึ่งในโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Networks) ที่ถูกพัฒนาขึ้นเพื่อแก้ไขปัญหาหลักที่พบในแบบจำลอง RNN ดั้งเดิม นั่นคือปัญหาค่าความชันหายหรือค่าความชันระเบิด (Vanishing and Exploding Gradient Problem) (Goodfellow lan, Bengio Yoshua, 2016)

LSTM ถูกนำเสนอครั้งแรกในปี ค.ศ. 1997 โดย Hochreiter และ Schmidhuber โดยมีแนวคิดหลักในการเพิ่มกลไกที่เรียกว่า "ประตู (Gate)" ซึ่งทำหน้าที่ควบคุมการไหลของข้อมูลและสามารถเลือกเก็บหรือทิ้งข้อมูลบางส่วนในแต่ละช่วงเวลา (time step) (Hochreiter & Schmidhuber, 1997) จึงทำให้ LSTM สามารถเก็บรักษาความสัมพันธ์ในระยะยาวได้อย่างมีประสิทธิภาพมากขึ้นเมื่อเทียบกับแบบจำลอง RNN ทั่วไป โครงสร้างและการทำงานของ LSTM จะประกอบไปด้วยเซลล์ (memory cells) ซึ่งแต่ละเซลล์จะประกอบด้วยประตูหลัก 3 ประเภท ได้แก่ ประตูอินพุต (Input Gate), ประตูหลงลืม (Forget Gate), และประตูเอาต์พุต (Output Gate) แต่ละประตูมีหน้าที่ควบคุมการไหลของข้อมูลภายในหน่วยความจำ (Goodfellow et al., Bengio Yoshua, 2016) ขั้นตอนการคำนวณภายใน LSTM สามารถแสดงสมการได้ ดังนี้

ประตูลืม (Forget Gate: $f^{(t)}$)

ทำหน้าที่ตัดสินใจว่าจะลบหรือเก็บข้อมูลจากช่วงเวลาก่อนหน้าที่เก็บอยู่ในเซลล์

$$f^{(t)} = \sigma(W_f x^{(t)} + U_f h^{(t-1)} + b_f) \quad (2-2)$$

ประตูอินพุต (Input Gate) และหน่วยความจำชั่วคราว (Candidate state)

เป็นส่วนที่ใช้ตัดสินใจว่าจะรับข้อมูลใหม่เข้าสู่หน่วยความจำในปัจจุบันหรือไม่

$$\begin{aligned} i^{(t)} &= \sigma(W_i x^{(t)} + U_i h^{(t-1)} + b_i) \\ \tilde{C}^{(t)} &= \tanh(W_c x^{(t)} + U_c h^{(t-1)} + b_c) \end{aligned} \quad (2-3)$$

จากนั้นจึงอัปเดตสถานะเซลล์ (cell state) ด้วยสมการ:

$$C^{(t)} = f^{(t)} \odot C^{(t-1)} + i^{(t)} \tilde{C}^{(t)} \quad (2-4)$$

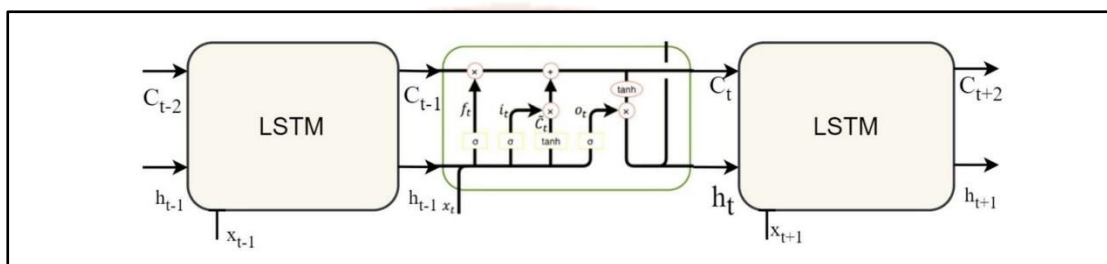
ประตูเอาต์พุต ใช้ในการตัดสินใจว่าจะนำข้อมูลใดจากสถานะเซลล์ออกมาแสดงผล

$$\begin{aligned} o^{(t)} &= \sigma(W_o x^{(t)} + U_o h^{(t-1)} + b_o) \\ h^{(t)} &= o^{(t)} \tanh(C^{(t)}) \end{aligned} \quad (2-5)$$

โดยที่

$x^{(t)}$	คือ	ข้อมูลอินพุต ณ เวลา t
$h^{(t)}$	คือ	ผลลัพธ์หรือ hidden state ณ ช่วงเวลา t
$C^{(t)}$	คือ	สถานะภายในของเซลล์หน่วยความจำที่เก็บข้อมูลระยะยาว
W, U	คือ	พารามิเตอร์น้ำหนัก (Weight matrices)
b	คือ	ค่า bias ของแต่ละประตู
σ	คือ	ฟังก์ชัน sigmoid activation และ $\tanh$ คือฟังก์ชัน activation แบบ hyperbolic tangent

กลไกนี้ช่วยให้ LSTM สามารถรักษาข้อมูลในระยะเวลานานได้โดยไม่สูญเสียไป แม้ว่าจะเป็นข้อมูลที่ผ่านช่วงเวลาไปนานแล้ว ทำให้เหมาะสมอย่างมากกับการใช้งานในการพยากรณ์ข้อมูลที่มีลักษณะความสัมพันธ์เชิงเวลาที่ซับซ้อน เช่น ข้อมูลอนุกรมเวลาหรือข้อมูลด้านภาษาศาสตร์ (Greff et al., 2017)



ภาพที่ 2 - 2 โครงสร้างภายในของแบบจำลอง Long Short-Term Memory (LSTM) (Ibrahim et al., 2023)

2.4 หน่วยระยะเวลานสั้น (Gated Recurrent Unit: GRU)

หน่วยความจำระยะสั้นแบบมีเกต หรือ Gated Recurrent Unit (GRU) เป็นหนึ่งในสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network: RNN) ที่ได้รับการพัฒนาขึ้นเพื่อแก้ไขปัญหาการเรียนรู้ความสัมพันธ์ระยะยาวในข้อมูลลำดับเวลา (time series data) ซึ่งมีความคล้ายคลึงกับ Long Short-Term Memory (LSTM) แต่ได้รับการออกแบบให้มีโครงสร้างที่เรียบง่ายกว่า และสามารถลดระยะเวลาในการคำนวณและการฝึกสอนแบบจำลองได้ดีขึ้น (Cho et al., 2014; Goodfellow Ian, Bengio Yoshua, 2016) โครงสร้างของ GRU ถูกเสนอครั้งแรกโดย Cho และคณะ ในปี ค.ศ. 2014 โดยมีวัตถุประสงค์เพื่อปรับปรุงประสิทธิภาพของ RNN ให้สามารถจัดการกับปัญหา Vanishing Gradient และ Exploding Gradient ที่เกิดขึ้นจากการคำนวณอนุกรมที่ยาวนาน [8] โครงสร้างและหลักการทำงานของ GRU ประกอบด้วยเกตหลัก 2 ตัว ได้แก่ Reset Gate และ Update Gate ซึ่งทำหน้าที่ควบคุมการไหลของข้อมูลในการเก็บข้อมูลสำคัญและลืมข้อมูลที่ไม่สำคัญไปได้อย่างมีประสิทธิภาพ โดยแต่ละเกตมีรายละเอียดดังนี้ (Goodfellow Ian, Bengio Yoshua, 2016)

เกตรีเซ็ต (Reset Gate: r_t)

มีหน้าที่กำหนดว่าข้อมูลในอดีตจะถูกนำมาใช้ในการคำนวณสถานะปัจจุบันมากน้อยเพียงใด
สามารถคำนวณได้จากสมการ

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (2-6)$$

เกตอัปเดต (Update Gate: z_t)

ทำหน้าที่ควบคุมการผสมผสานระหว่างสถานะข้อมูลในอดีตกับข้อมูลใหม่ที่ได้รับเข้ามาในแต่ละช่วงเวลา
ซึ่งจะเป็นตัวกำหนดว่าจะรักษาข้อมูลเก่าไว้หรือปรับปรุงใหม่มากน้อยเพียงใด
โดยคำนวณได้ดังสมการ

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (2-7)$$

จากนั้นสถานะใหม่ของ GRU (h_t) จะถูกคำนวณด้วยสมการ

$$\tilde{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \quad (2-8)$$

และอัปเดตสถานะในแต่ละช่วงเวลาจากสมการ

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (2-9)$$

โดยที่

x_t คือ ข้อมูล input ในช่วงเวลา t

h_{t-1} คือ สถานะที่คำนวณได้จากช่วงเวลาก่อนหน้า (previous hidden state)

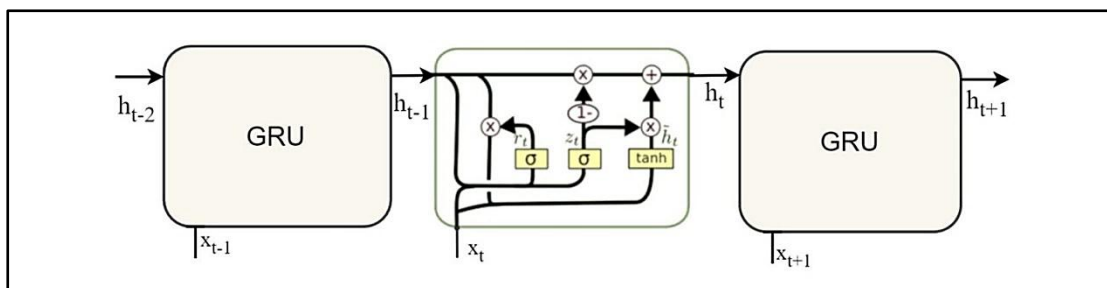
W_z, W_r, W_h, U_z, U_r และ b คือ เมทริกซ์น้ำหนัก (Weight matrices) และ bias ตามลำดับ

σ คือ ฟังก์ชัน sigmoid ซึ่งทำหน้าที่ในการควบคุมให้เกตมีค่าอยู่ระหว่าง 0 และ 1

$\odot$ คือ การคูณแบบ element-wise

$\tilde{h}_t$ คือ สถานะชั่วคราวที่ได้จากการคำนวณภายใน GRU cell

จากสมการที่นำเสนอข้างต้นจะเห็นได้ว่า GRU มีความคล้ายกับ LSTM แต่ถูกออกแบบให้ลดความซับซ้อน โดยรวม Input Gate และ Forget Gate ใน LSTM มาเป็น Update Gate ตัวเดียวใน GRU ซึ่งช่วยลดจำนวนพารามิเตอร์และทำให้การคำนวณเร็วขึ้น (Goodfellow et al., Bengio & Yoshua, 2016) ข้อได้เปรียบสำคัญของ GRU คือมีจำนวนพารามิเตอร์ที่น้อยกว่า LSTM จึงเหมาะสมสำหรับการใช้งานที่มีข้อมูลไม่มาก หรือแบบจำลองที่ต้องการประสิทธิภาพในการคำนวณสูง แต่ในขณะเดียวกันก็มีความสามารถในการจับรูปแบบของข้อมูลลำดับเวลาได้ดีใกล้เคียงกับ LSTM เช่นกัน (Chung et al., 2014) ดังนั้น GRU จึงเป็นทางเลือกที่น่าสนใจในการประยุกต์ใช้สำหรับการคาดการณ์ข้อมูลอนุกรมเวลาที่มีลักษณะเฉพาะหรือมีข้อจำกัดในด้านทรัพยากรการประมวลผล



ภาพที่ 2 - 3 โครงสร้างภายในของแบบจำลอง Gated Recurrent Unit (GRU) (Ibrahim et al., 2023)

2.5 ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย Mean Absolute Error (MAE)

ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) เป็นหนึ่งในตัวชี้วัดทางสถิติที่ใช้กันอย่างแพร่หลายในการประเมินประสิทธิภาพของแบบจำลองในการพยากรณ์ข้อมูลประเภทอนุกรมเวลา (Time Series Forecasting) และการพยากรณ์ทั่วไปในทางสถิติและ Machine Learning โดยค่าดังกล่าวจะแสดงให้เห็นถึงขนาดเฉลี่ยของความผิดพลาดในการทำนายที่เกิดขึ้นระหว่างค่าที่คาดการณ์ (Predicted Values) กับค่าจริงที่สังเกตได้ (Observed Values) (Hyndman & Athanasopoulos, 2021)

สูตรของ MAE สามารถคำนวณได้ตามสมการ ดังนี้ (Chai & Draxler, 2014)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2-10)$$

โดยที่

n คือ จำนวนข้อมูลที่นำมาประเมิน

y_i คือ ค่าจริงที่สังเกตได้ (Actual Value)

$\hat{y}_i$ คือ ค่าที่ได้จากการพยากรณ์ของแบบจำลอง (Predicted Value)

ในการตีความ ค่า MAE ที่ต่ำหมายถึงแบบจำลองมีความแม่นยำสูงในการทำนาย เนื่องจากความคลาดเคลื่อนเฉลี่ยระหว่างค่าจริงและค่าทำนายมีค่าน้อย แต่หากค่า MAE สูง แสดงว่าแบบจำลองมีความผิดพลาดในการคาดการณ์ค่อนข้างมาก (Chai & Draxler, 2014) MAE มีข้อดีในแง่ของการคำนวณที่ง่าย การตีความตรงไปตรงมา เนื่องจากแสดงผลเป็นหน่วยเดียวกับข้อมูลต้นฉบับ (Original Unit) และไม่ให้ความสำคัญกับทิศทางของความผิดพลาดว่าค่าจริงมากกว่าหรือน้อยกว่าค่าที่ทำนาย เพียงแต่เน้นว่าค่าที่ทำนายนั้นห่างจากความจริงโดยเฉลี่ยเท่าไร อย่างไรก็ตาม MAE มีข้อจำกัดในการวัดประสิทธิภาพแบบจำลองคือให้น้ำหนักเท่ากันทุกจุดของความผิดพลาด ทำให้ไม่สะท้อนผลกระทบของค่าผิดพลาดขนาดใหญ่ได้ดีเท่ากับ RMSE ที่มีการให้ความสำคัญกับความ

ผิดพลาดขนาดใหญ่มากกว่า (Willmott & Matsuura, 2005) ด้วยลักษณะดังกล่าว MAE จึงเป็นตัวชี้วัดพื้นฐานที่ถูกนำมาใช้ในการวัดผลลัพธ์ของแบบจำลองการพยากรณ์ข้อมูลอนุกรมเวลา โดยเฉพาะในการพยากรณ์ข้อมูลที่ต้องการเน้นย้ำความแม่นยำของค่าพยากรณ์ในระดับทั่วไปที่เข้าใจง่าย เช่น การคาดการณ์จำนวนนักท่องเที่ยว การคาดการณ์สภาพอากาศ หรือปริมาณยอดขาย เป็นต้น (Hyndman & Athanasopoulos, 2021; Willmott & Matsuura, 2005)

2.6 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรากที่สอง Root Mean Squared Error (RMSE)

ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยรากที่สอง (Root Mean Squared Error: RMSE) เป็นตัวชี้วัดทางสถิติที่ถูกนำมาใช้บ่อยในการประเมินประสิทธิภาพของแบบจำลองพยากรณ์ โดยเฉพาะในงานที่เกี่ยวข้องกับข้อมูลอนุกรมเวลา (Time Series Forecasting) และการพยากรณ์ข้อมูลที่ต้องการเน้นย้ำความสำคัญของความผิดพลาดที่มีขนาดใหญ่ RMSE เป็นค่าที่ได้จากการคำนวณรากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสองระหว่างค่าที่พยากรณ์ (Predicted Values) และค่าจริง (Actual Values) ซึ่งแสดงดังสมการได้ ดังนี้ (Hyndman & Athanasopoulos, 2021)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2-11)$$

โดยที่

n คือ จำนวนข้อมูลที่ใช้ในการคำนวณ

y_i คือ ค่าจริงที่ได้จากการสังเกต (Actual Values)

$\hat{y}_i$ คือ ค่าที่พยากรณ์ได้จากแบบจำลอง (Predicted Values)

ข้อดีของการใช้ RMSE ในการวัดประสิทธิภาพการพยากรณ์คือ RMSE จะให้น้ำหนักกับค่าผิดพลาดขนาดใหญ่ที่โดดเด่นชัดเจนกว่า MAE ดังนั้น RMSE จึงเหมาะสมในสถานการณ์ที่นักวิจัยส่วนใหญ่มักเลือกใช้ต้องการเน้นให้เห็นความสำคัญของข้อผิดพลาดที่มีขนาดใหญ่ และต้องการลดโอกาสเกิดข้อผิดพลาดเหล่านี้ให้มากที่สุด ซึ่งจะส่งผลให้แบบจำลองที่มีค่า RMSE ต่ำกว่ามักได้รับการพิจารณาว่ามีประสิทธิภาพสูงกว่าแบบจำลองที่มี RMSE สูง (Chai & Draxler, 2014) อย่างไรก็ตาม RMSE ก็มีข้อเสียอยู่บ้าง คือการที่มีการยกกำลังสองในสมการ อาจทำให้ค่า RMSE มีความไวต่อข้อมูลที่มีความผิดพลาดรุนแรง (Outliers) ซึ่งอาจมีผลต่อการประเมินประสิทธิภาพแบบจำลองได้ในบางกรณี (Willmott & Matsuura, 2005)

2.7 ค่าสัมประสิทธิ์การตัดสินใจ Coefficient of Determination (R^2)

ค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of Determination หรือเรียกสั้นๆ ว่า R^2) เป็นตัวชี้วัดที่แสดงให้เห็นถึงระดับความแม่นยำในการทำนายหรือพยากรณ์ของแบบจำลอง เมื่อเทียบกับข้อมูลที่สังเกตได้จริง โดยตัวเลข R^2 จะอยู่ระหว่าง 0 ถึง 1 โดยค่าที่เข้าใกล้ 1 บ่งชี้ว่าแบบจำลองมีประสิทธิภาพในการทำนายที่สูง และแสดงให้เห็นถึงความสัมพันธ์หรือความสอดคล้องของค่าที่แบบจำลองพยากรณ์กับค่าที่เกิดขึ้นจริง ขณะที่ค่าใกล้ 0 แสดงว่าแบบจำลองมีประสิทธิภาพต่ำในการอธิบายข้อมูลจริงหรือแทบไม่มีความสัมพันธ์ใดๆ เลย นอกจากนี้ หากค่า R^2 มีค่าติดลบ แสดงว่าการพยากรณ์ด้วยแบบจำลองนั้นด้อยกว่าการใช้ค่าเฉลี่ยอย่างง่ายของข้อมูลที่ผ่านมา (Mean Model) (Miles, 2014) ซึ่งหมายความว่าค่าที่ทำนายโดยแบบจำลองมีทิศทางไม่สอดคล้องกับค่าจริงหรือมีแนวโน้มไปในทิศทางตรงกันข้าม การคำนวณ R^2 สามารถเขียนเป็นสมการทางคณิตศาสตร์ได้ ดังนี้

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2-12)$$

โดยที่

y_i คือ ค่าจริงที่ได้จากการสังเกต (Observed values)

$\hat{y}_i$ คือ ค่าที่แบบจำลองสามารถทำนายได้ (Predicted values)

$\bar{y}$ คือ ค่าเฉลี่ยของข้อมูลจริงทั้งหมด (Mean of observed values)

n คือ จำนวนข้อมูลทั้งหมดที่นำมาพิจารณา

จากสูตรดังกล่าว ส่วนประกอบในสมการแบ่งได้เป็นสองส่วนสำคัญ คือ ส่วนที่เป็นผลรวมของกำลังสองของความคลาดเคลื่อนจากค่าจริงและค่าทำนาย และส่วนที่เป็นผลรวมของการกระจายตัวรอบค่าเฉลี่ย ข้อมูลทั้งสองส่วนนี้สะท้อนถึงระดับความถูกต้องของการทำนาย เมื่อเปรียบเทียบกับความผันผวนของข้อมูลจริงที่มีอยู่ (James et al., 2021) ในแง่ของการแปลความหมาย ค่าที่ได้จาก R^2 สามารถมองได้ว่า เป็นเปอร์เซ็นต์ของข้อมูลจริงที่ถูกอธิบายโดยแบบจำลอง เช่น หาก R^2 เท่ากับ 0.85 สามารถตีความได้ว่า แบบจำลองดังกล่าวสามารถอธิบายข้อมูลได้ประมาณ 85% ของทั้งหมด และอีก 15% เป็นข้อมูลที่ไม่สามารถอธิบายได้จากแบบจำลองนี้ หรือเรียกง่ายๆ ว่าเป็นค่าความแม่นยำหรือประสิทธิภาพของแบบจำลอง (Hair et al., 2019)

2.8 งานวิจัยที่เกี่ยวข้อง

ในการศึกษางานวิจัยที่เกี่ยวข้องกับการพยากรณ์จำนวนนักท่องเที่ยว พบว่ามีการใช้เทคนิคต่าง ๆ ทั้งในส่วนของสถิติพื้นฐาน (Statistical Techniques), การเรียนรู้ของเครื่อง (Machine Learning Techniques), และการเรียนรู้เชิงลึก (Deep Learning Techniques) ซึ่งแต่ละวิธีมีจุดเด่น ข้อจำกัด และบริบทที่แตกต่างกัน โดยสามารถสรุปและอธิบายงานวิจัยที่สำคัญได้ ดังนี้

2.8.1 เทคนิคทางสถิติพื้นฐาน

Boonlha (2019) ศึกษาการพยากรณ์จำนวนนักท่องเที่ยวอุทยานแห่งชาติภูหินร่องกล้า จังหวัดพิษณุโลก โดยใช้ข้อมูลรายเดือนตั้งแต่มกราคม 2555 ถึงเมษายน 2562 เพื่อเปรียบเทียบ 3 เทคนิค ได้แก่ Holt-Winters Multiplicative Method, Box-Jenkins (ARIMA), และ Combined Forecasting Method โดยพบว่า Holt-Winters Multiplicative Method มีประสิทธิภาพดีที่สุด เนื่องจากเทคนิคนี้สามารถจัดการกับข้อมูลที่มีฤดูกาลได้อย่างชัดเจน ส่วนวิธี Combined Forecasting Method และ ARIMA มีค่าความผิดพลาดที่สูงกว่าเล็กน้อย แสดงให้เห็นว่า การเลือกใช้เทคนิคที่เหมาะสมกับลักษณะข้อมูลฤดูกาลมีความสำคัญอย่างมากในการพยากรณ์ข้อมูลนักท่องเที่ยว (Boonlha, 2019)

Khatrı และคณะ (2024) ใช้ข้อมูลรายปีตั้งแต่ปี ค.ศ. 1964-2023 ในการพยากรณ์แนวโน้มนักท่องเที่ยวประเทศเนปาล ด้วยวิธีการ ARIMA โดยพบว่าแบบจำลอง ARIMA(2,1,3) ให้ผลลัพธ์ที่ดีที่สุดมีความแม่นยำในการคาดการณ์สูง และสามารถจัดการกับข้อมูลที่ได้รับผลกระทบจากเหตุการณ์สำคัญ เช่น แผ่นดินไหวและโรคระบาดโควิด-19 อย่างมีประสิทธิภาพ ซึ่งแสดงถึงข้อดีของ ARIMA ในการจัดการข้อมูลที่มีเหตุการณ์ไม่คาดคิดเป็นช่วง ๆ ได้ดี (Khatrı et al., 2024)

Choden และ Unhapipat (2018) ใช้แบบจำลอง ARIMA(0,0,0) $\times$ (1,1,0)₁₂ พยากรณ์นักท่องเที่ยวต่างชาติในเขตบมธัง ประเทศภูฏาน ด้วยข้อมูลรายเดือนในช่วงปี 2012-2016 พบว่าแบบจำลองนี้สามารถพยากรณ์ได้แม่นยำสูง แต่ข้อจำกัดของงานนี้คือช่วงเวลาพยากรณ์ที่สั้น และไม่สามารถพิจารณาปัจจัยภายนอกที่สำคัญอื่น ๆ ได้ (Choden & Unhapipat, 2018)

2.8.2 เทคนิค Machine Learning

Liu และคณะ (2019) ได้ศึกษาการใช้แบบจำลอง XGBoost และ Grey Model (GM) เพื่อพยากรณ์ปริมาณนักท่องเที่ยวเมืองชานย่า ประเทศจีน จากข้อมูลย้อนหลัง 7 ปี ผลการวิจัยพบว่า ทั้งสองแบบจำลองมีประสิทธิภาพใกล้เคียงกัน แต่เมื่อนำผลลัพธ์จากทั้งสองแบบจำลองมาผสมรวมกัน (ensemble) พบว่าความแม่นยำดีขึ้นชัดเจน ซึ่งสะท้อนถึงจุดเด่นของการนำเทคนิค Machine Learning มาผสานกับแบบจำลองสถิติที่เรียบง่าย (Liu et al., 2019)

Brilliandy และคณะ (2022) ใช้แบบจำลองถดถอยหลายประเภท รวมทั้ง Random Forest, Support Vector Regression และแบบจำลองอื่น ๆ เพื่อทำนายจำนวนนักท่องเที่ยวที่เดินทางมายังประเทศอินโดนีเซียในช่วงวิกฤต COVID-19 พบว่า Random Forest มีประสิทธิภาพดีที่สุด โดยแสดงให้เห็นว่าเทคนิคนี้สามารถจัดการข้อมูลที่มีความสัมพันธ์ที่ซับซ้อน เช่น สถานการณ์โรคระบาดที่ส่งผลต่อการเดินทางท่องเที่ยวได้อย่างมีประสิทธิภาพ (Brilliandy et al., 2022)

Limpornchitwilai และคณะ (2022) ได้พัฒนาการพยากรณ์จำนวนนักท่องเที่ยวของพิพิธภัณฑ์ที่ตั้งอยู่บริเวณเดียวกันในประเทศไทย ด้วยเทคนิค Random Forest และ XGBoost โดยการเพิ่มปัจจัยทางสภาพอากาศ เช่น ปริมาณฝนและอุณหภูมิ ทำให้แบบจำลองสามารถพยากรณ์ได้แม่นยำมากขึ้น ซึ่งแสดงให้เห็นว่า การเพิ่มปัจจัยภายนอกที่เกี่ยวข้องช่วยเพิ่มความแม่นยำในการพยากรณ์ได้ (Limpornchitwilai et al., 2022)

2.8.3 เทคนิค Deep Learning

Law และคณะ (2019) ใช้แบบจำลอง Deep Learning แบบ LSTM ร่วมกับ Attention Mechanism เพื่อพยากรณ์จำนวนนักท่องเที่ยวรายเดือนที่เดินทางมายังมาเก๊า ผลการศึกษาชี้ให้เห็นว่า แบบจำลอง Deep Learning ที่สามารถคัดเลือกปัจจัยสำคัญจากข้อมูลขนาดใหญ่ (เช่น การค้นหาข้อมูลบนอินเทอร์เน็ต) ได้อัตโนมัติ สามารถให้ค่าความผิดพลาดต่ำ ซึ่งดีกว่าแบบจำลองแบบดั้งเดิม เช่น SVR และ ARIMA อย่างชัดเจน (Law et al., 2019)

Lemmel และคณะ (2022) ได้เปรียบเทียบแบบจำลอง Deep Learning (GRU และ LSTM) กับแบบจำลอง ARIMA ในการพยากรณ์จำนวนนักท่องเที่ยวในเมืองซาลซ์บูร์ก ประเทศออสเตรีย ซึ่งมีข้อมูลจำกัด ผลการวิจัยพบว่าแบบจำลอง GRU และ LSTM มีความสามารถในการเรียนรู้ข้อมูลที่มีการเปลี่ยนแปลงซับซ้อนได้ดีกว่า ARIMA โดยเฉพาะเมื่อมีข้อมูลที่มีความละเอียดและมีปัจจัยภายนอก เช่น สภาพอากาศหรือวันหยุดเข้ามาประกอบ อีกทั้งแบบจำลอง Deep Learning ยังให้ผลการพยากรณ์ได้รวดเร็วกว่า (Lemmel et al., 2022)

Kummong และ Homdok (2019) ใช้แบบจำลองโครงข่ายประสาทเทียมแบบ Recurrent Neural Network (RNN) เปรียบเทียบกับแบบจำลอง ANN พื้นฐาน (Back Propagation Neural Network) เพื่อพยากรณ์จำนวนนักท่องเที่ยวต่างชาติในภาคเหนือของประเทศไทย ผลการวิจัยแสดงให้เห็นว่า RNN มีความแม่นยำสูงกว่า ANN พื้นฐาน ซึ่งแสดงให้เห็นถึงประสิทธิภาพของการใช้แบบจำลองที่สามารถเรียนรู้ความสัมพันธ์เชิงเวลาในการพยากรณ์ข้อมูลอนุกรมเวลาได้ (Kummong & Homdok, 2019)

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า เทคนิคที่นำมาใช้ในการพยากรณ์จำนวนนักท่องเที่ยวมีความหลากหลาย ทั้งเทคนิคทางสถิติ เทคนิค Machine Learning และเทคนิค Deep Learning โดยเฉพาะอย่างยิ่งแบบจำลองโครงข่ายประสาทเทียมแบบ RNN, LSTM และ GRU ซึ่งได้รับการออกแบบมาเพื่อจัดการกับข้อมูลอนุกรมเวลา (Time Series) โดยเฉพาะ พบว่ามีความสามารถในการเรียนรู้ลักษณะเชิงเวลา ความสัมพันธ์ระยะยาว และความซับซ้อนของข้อมูลได้ดีกว่าแบบจำลองแบบดั้งเดิมในหลายบริบท งานวิจัยหลายชิ้นชี้ให้เห็นว่าเทคนิค Deep Learning สามารถให้ผลการพยากรณ์ที่มีประสิทธิภาพสูง โดยเฉพาะเมื่อมีข้อมูลจำนวนมากหรือมีปัจจัยแวดล้อมที่หลากหลาย เช่น สภาพอากาศ วันหยุด หรือเหตุการณ์สำคัญต่าง ๆ ทั้งนี้ การเลือกใช้เทคนิคที่เหมาะสมกับลักษณะของข้อมูลและบริบทของการพยากรณ์ยังคงเป็นปัจจัยสำคัญที่ส่งผลต่อประสิทธิภาพของแบบจำลอง

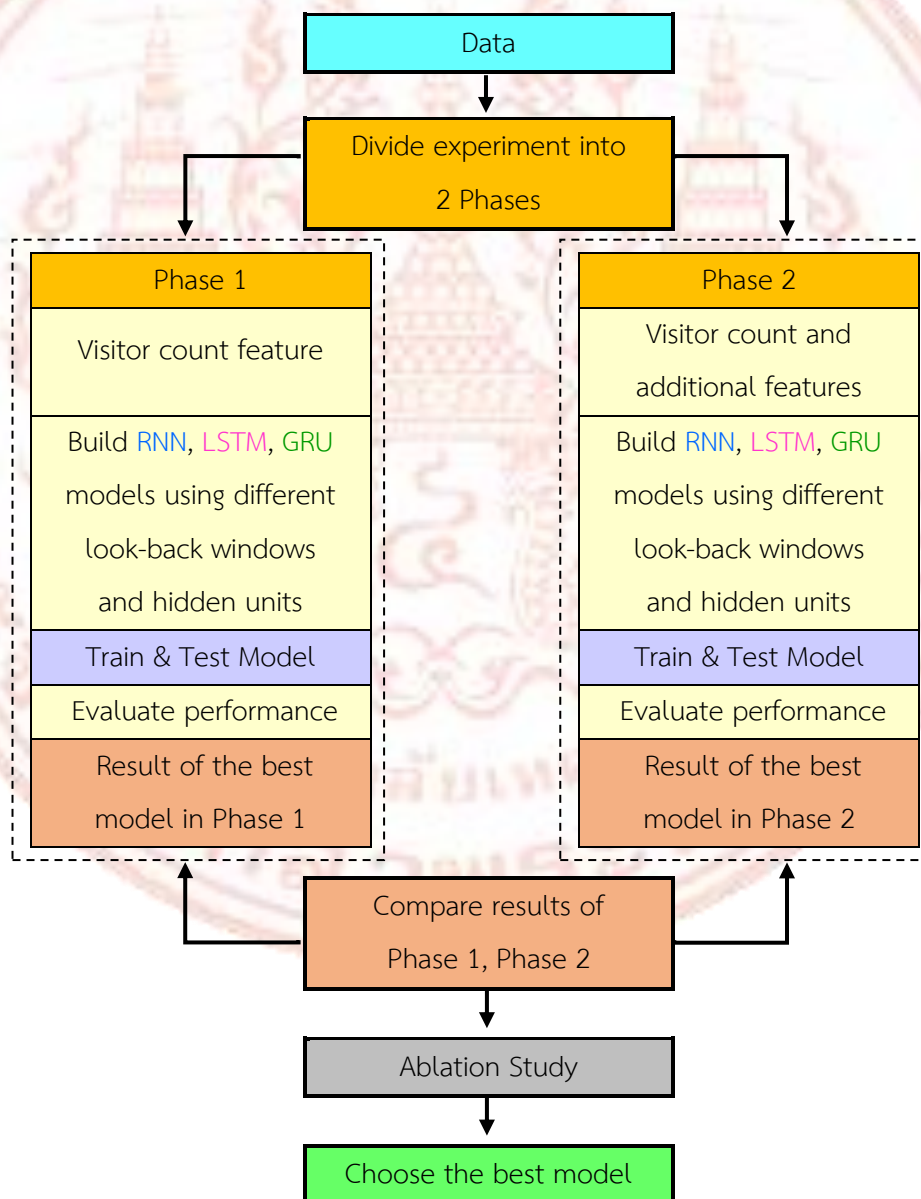


บทที่ 3

วิธีการดำเนินการวิจัย

งานวิจัยนี้เป็นการวิจัยเพื่อพัฒนาแบบจำลองโครงข่ายประสาทเทียมแบบเรียกซ้ำสำหรับพยากรณ์จำนวนนักท่องเที่ยวรายเดือนในอุทยานแห่งชาติของประเทศไทย โดยเปรียบเทียบประสิทธิภาพระหว่าง RNN, LSTM, GRU และประเมินผลด้วยตัวชี้วัดประสิทธิภาพ MAE, RMSE และ R^2

3.1 ภาพรวมขั้นตอนวิธีดำเนินการวิจัย



ภาพที่ 3 - 1 ภาพรวมขั้นตอนวิธีดำเนินการวิจัย

ขั้นตอนการดำเนินการวิจัยแบ่งการทดลองออกเป็น 2 ระยะ (Phase) เนื่องจากการพยากรณ์ข้อมูลอนุกรมเวลานั้นสามารถพัฒนาได้จากหลายแนวทาง ในการศึกษาเบื้องต้นพบว่า ปัจจัยต่าง ๆ ที่นำมาใช้ในแบบจำลองสามารถส่งผลต่อประสิทธิภาพการพยากรณ์ที่แตกต่างกัน การแยกออกเป็น 2 ระยะจึงช่วยให้สามารถเปรียบเทียบและประเมินผลได้ชัดเจนมากขึ้น ว่าการเพิ่มฟีเจอร์อื่นนอกเหนือจากจำนวนนักท่องเที่ยว จะส่งผลดีหรือผลเสียต่อประสิทธิภาพของแบบจำลองเมื่อเปรียบเทียบกับการใช้ข้อมูลนักท่องเที่ยวเพียงอย่างเดียว

Phase 1 เป็นระยะเริ่มต้นที่ใช้เฉพาะข้อมูลจำนวนนักท่องเที่ยวเพียงฟีเจอร์เดียว เพื่อแสดงให้เห็นความสามารถพื้นฐานของเทคนิคต่างๆ ที่ถูกนำมาใช้ ได้แก่ RNN, LSTM และ GRU ว่าหากใช้ข้อมูลพื้นฐานเพียงฟีเจอร์เดียวแล้ว แบบจำลองแต่ละประเภทสามารถให้ประสิทธิภาพในการทำนายได้ในระดับใด

Phase 2 ระยะที่สองมีการเพิ่มข้อมูลฟีเจอร์ที่เกี่ยวข้องอื่นๆ ได้แก่ ข้อมูลลักษณะภูมิอากาศ ข้อมูลเดือน ข้อมูลปริมาณยานพาหนะที่เข้า-ออกพื้นที่อุทยานแต่ละแห่ง และข้อมูลเชิงพื้นที่ ได้แก่ ภูมิภาค และลักษณะเด่นของอุทยานแต่ละแห่ง ที่อาจมีผลต่อการท่องเที่ยวในอุทยานแห่งชาติ จุดประสงค์ของ Phase 2 คือการศึกษาและทดสอบว่าการเพิ่มข้อมูลที่มีความซับซ้อนเข้าไป จะทำให้แบบจำลองที่พัฒนาขึ้นมีประสิทธิภาพในการทำนายดีขึ้นหรือไม่เมื่อเปรียบเทียบกับ Phase 1

ในส่วนถัดไปเป็นการทดสอบหาระดับความสำคัญของฟีเจอร์ (Ablation Study) โดยศึกษาว่าฟีเจอร์แต่ละรายการส่งผลต่อประสิทธิภาพของแบบจำลองในการพยากรณ์มากน้อยเพียงใด การทดลองดำเนินการโดยการปรับเปลี่ยนชุดฟีเจอร์ที่ใช้ในการฝึกแบบจำลอง เพื่อพิจารณาว่าการตัดฟีเจอร์บางรายการออกจะทำให้ผลการพยากรณ์ดีขึ้น แย่ลง หรือเปลี่ยนแปลงในลักษณะใด การวิเคราะห์นี้ช่วยสะท้อนบทบาทของฟีเจอร์ที่จำเป็นต่อแบบจำลอง และสนับสนุนการปรับปรุงโครงสร้างแบบจำลองให้มีประสิทธิภาพยิ่งขึ้น

3.2 กระบวนการรวบรวมและเตรียมข้อมูล

3.2.1 การรวบรวมข้อมูลปริมาณนักท่องเที่ยวรายเดือนของอุทยานแห่งชาติในประเทศไทยจำนวน 146 แห่ง ในช่วงเวลา 10 ปีที่ผ่านมา ตั้งแต่เดือนมกราคม พ.ศ. 2557 ถึงเดือนธันวาคม พ.ศ. 2566 และข้อมูลปริมาณยานพาหนะที่เข้า-ออกพื้นที่อุทยานแต่ละแห่ง จากกรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช รวมถึงข้อมูลสภาพอากาศจากกรมอุตุนิยมวิทยา และ www.timeanddate.com เพื่อใช้ในการประกอบการวิเคราะห์ โดยข้อมูลที่ใช้ ประกอบด้วย

- 3.2.1.1 ข้อมูลนักท่องเที่ยวรายเดือน
- 3.2.1.2 ข้อมูลอุณหภูมิรายเดือน (สูงสุด, ต่ำสุด, ค่าเฉลี่ย และค่าเบี่ยงเบนมาตรฐาน)
- 3.2.1.3 ข้อมูลปริมาณยานพาหนะเข้า-ออกอุทยานรายเดือน จากกรมอุทยานแห่งชาติ
- 3.2.1.4 ข้อมูลเชิงพื้นที่ของอุทยาน เช่น ภูมิภาคที่ตั้ง และลักษณะเด่นเฉพาะของอุทยาน

3.2.2 กระบวนการเตรียมข้อมูลประกอบด้วย

3.2.2.1 การตรวจสอบข้อมูล (Data Quality Assurance) โดยทำการตรวจสอบความถูกต้อง ครบถ้วน และความสม่ำเสมอของข้อมูลทั้งหมด

3.2.2.2 การแปลงรูปแบบข้อมูล จาก Wide-form เป็น Long-form เพื่อให้เหมาะสมกับการวิเคราะห์อนุกรมเวลา

3.2.2.3 การจัดการข้อมูลสูญหาย (Missing Data) โดยเติมข้อมูลที่สูญหายด้วยค่าเฉลี่ยหรือค่าศูนย์ตามความเหมาะสมของแต่ละฟีเจอร์

3.2.2.4 การแปลงข้อมูลเชิงหมวดหมู่ (Categorical Variables) โดยใช้เทคนิค One-Hot Encoding เพื่อให้สามารถนำไปใช้ในการฝึกแบบจำลองได้อย่างมีประสิทธิภาพ

3.2.2.5 การปรับสเกลข้อมูล (Normalization) ด้วย StandardScaler เพื่อให้แต่ละฟีเจอร์มีสเกลที่สอดคล้องกัน และช่วยให้การเรียนรู้ของแบบจำลองมีประสิทธิภาพสูงสุด

3.3 เครื่องมือที่ใช้ดำเนินการ

การพัฒนาแบบจำลองและการวิเคราะห์ข้อมูล ได้ใช้เครื่องมือและไลบรารีที่สำคัญ ดังนี้

3.3.1 TensorFlow และ Keras สำหรับสร้าง ฝึก และทดสอบแบบจำลอง RNN, LSTM และ GRU

3.3.2 NumPy และ Pandas: สำหรับการจัดการข้อมูล การรวมข้อมูล และการจัดโครงสร้างข้อมูลให้อยู่ในรูปแบบที่เหมาะสมสำหรับการวิเคราะห์

3.3.3 Scikit-learn: สำหรับการคำนวณค่าความแม่นยำ (MAE, RMSE, R^2) และการทำ Normalization ข้อมูล

3.3.4 Matplotlib และ Seaborn: สำหรับแสดงผลข้อมูลด้วยกราฟและช่วยในการวิเคราะห์ผลลัพธ์เชิงภาพ

3.4 การออกแบบการทดลอง

การพัฒนาและทดสอบแบบจำลองแบ่งออกเป็น 3 แนวทาง ได้แก่

3.4.1 แบบจำลองรายอุทยาน (Individual Park Models : IPM) คือ การพัฒนาแบบจำลองแยกเฉพาะแต่ละอุทยาน

3.4.2 แบบจำลองรวม (Combined Models : CM) คือ การรวมข้อมูลจากทุกอุทยานเพื่อสร้างแบบจำลองร่วมกัน

3.4.3 แบบจำลองแบบจัดกลุ่ม คือ การสร้างแบบจำลองแยกกลุ่มตามข้อมูลเชิงพื้นที่ ได้แก่ แบ่งตามภูมิภาค (Clustered Regional Models: CRM) และแบ่งตามลักษณะเฉพาะของอุทยาน (Clustered Unique Models: CUM)

ทั้งนี้ เพื่อศึกษาว่าแนวทางใดเหมาะสมที่สุดในการนำไปประยุกต์ใช้งานจริง โดยประเมินผ่านตัวชี้วัด MAE, RMSE และ R^2 และทำการทดลองแต่ละแนวทางด้วยเทคนิค RNN, LSTM และ GRU และปรับพารามิเตอร์ Look-back และ Hidden Units ดังที่กล่าวไว้ข้างต้น

3.5 การกำหนดพารามิเตอร์ที่สำคัญ

3.5.1 กำหนดพารามิเตอร์ของจำนวนข้อมูลย้อนหลัง (Look-back) ใช้ช่วงเวลา 3, 6 และ 9 เดือน เพื่อทดสอบประสิทธิภาพในการพยากรณ์

3.5.2 กำหนดจำนวนหน่วยซ่อน (Hidden Units) โดยทดลองจำนวนหน่วยซ่อนตั้งแต่ 10 ถึง 100 (ทีละ 10 หน่วย) เพื่อให้แบบจำลองได้ทดลองกับหลายหน่วยซ่อนเพื่อหาช่วงที่เหมาะสม

3.6 การประเมินผล

การประเมินประสิทธิภาพของแบบจำลองในงานวิจัยนี้ใช้ตัวชี้วัด 3 รายการ ได้แก่ Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) และ Coefficient of Determination (R^2) ซึ่งเป็นตัวชี้วัดมาตรฐานที่ใช้กันทั่วไปในการพยากรณ์ข้อมูลเชิงเวลา โดยใช้เพื่อเปรียบเทียบผลลัพธ์ในแต่ละชุดการทดลอง โดยแนวทางการพิจารณาผลลัพธ์ของตัวชี้วัด มีดังนี้

(ก) ค่า MAE และ RMSE ต่ำ

แบบจำลองที่มีค่า MAE และ RMSE ต่ำแสดงว่ามีความคลาดเคลื่อนในการพยากรณ์ค่อนข้างน้อย สามารถทำนายได้ใกล้เคียงค่าจริง

(ข) ค่า R^2 สูง

หากแบบจำลองใดมีค่า R^2 สูง (ใกล้ 1) แสดงว่าแบบจำลองสามารถจับความสัมพันธ์ของข้อมูลได้ดี ส่งผลให้แบบจำลองมีประสิทธิภาพโดยรวม

(ค) การเปรียบเทียบในภาพรวม

เมื่อพิจารณาทั้ง MAE, RMSE และ R^2 ร่วมกัน หากแบบจำลองใดมีค่า MAE และ RMSE อยู่ในระดับต่ำ พร้อมทั้งให้ค่า R^2 สูง จะถือว่าเป็นแบบจำลองที่มีประสิทธิภาพการพยากรณ์ที่ดี อย่างไรก็ตาม ในบางกรณีที่ค่า MAE หรือ RMSE อาจไม่ได้ต่ำที่สุด แต่หาก R^2 อยู่ในเกณฑ์ที่สูงก็แสดงว่าแบบจำลองยังคงสามารถทำนายได้ใกล้เคียงกับข้อมูลจริงในภาพรวม ในทางกลับกัน หากค่า R^2 ต่ำมาก แม้ว่า MAE หรือ RMSE จะต่ำ ก็อาจสะท้อนถึงข้อจำกัดของแบบจำลองในการจับความสัมพันธ์โดยรวมของข้อมูล

3.7 แนวทางการสรุปและวิเคราะห์ผลการทดลอง

3.7.1 ดำเนินการสรุปผลลัพธ์ของแบบจำลองในแต่ละแนวทาง

หลังจากการพัฒนาแบบจำลองในแต่ละชุดการทดลองเรียบร้อยแล้ว ผู้วิจัยจะดำเนินการคัดเลือกแบบจำลองที่มีประสิทธิภาพสูงสุดเพื่อนำมาเปรียบเทียบกับแต่ละชุดการทดลอง โดยมีรายละเอียดขั้นตอนของการเปรียบเทียบมี ดังนี้

(ก) แบบจำลองรายอุทยาน (Individual Park Models) ผู้วิจัยสร้างแบบจำลองของแต่ละอุทยานแห่งชาติ จำนวน 146 แห่งด้วยพารามิเตอร์ที่กำหนด เมื่อฝึก (Train) และทดสอบ (Test) เสร็จสิ้น จึงรวบรวมผลของค่าประสิทธิภาพ (MAE, RMSE, R^2) ทั้งหมดและคำนวณหาค่าเฉลี่ยโดยรวม

(ข) แบบจำลองรวม (Combined Model) ผู้วิจัยรวมข้อมูลนักท่องเที่ยวจากอุทยานทั้ง 146 แห่งเข้าด้วยกันเป็น “ชุดข้อมูลเดียว” เสมือนเป็นแบบจำลองเดียวที่เรียนรู้ข้อมูลจากทุกอุทยานพร้อมกัน และทำการทดสอบซ้ำ 10 รอบ เพื่อให้ผลลัพธ์มีความเสถียรและเป็นกลางเมื่อนำไปเปรียบเทียบ แล้วนำค่าประสิทธิภาพ (MAE, RMSE, R^2) มาคำนวณหาค่าเฉลี่ยโดยรวม

(ค) แบบจำลองแบบจัดกลุ่ม (Clustering Models) ผู้วิจัยแบ่งอุทยานออกเป็น 2 แบบ ได้แก่ (1) แบ่งตามภูมิภาค (ภาคเหนือ, ภาคตะวันออกเฉียงเหนือ, ภาคกลาง, ภาคตะวันออก, ภาคตะวันตก, ภาคใต้) และ (2) แบ่งตามลักษณะเด่น (กลุ่มภูเขา, กลุ่มน้ำตก, กลุ่มวัฒนธรรม, กลุ่มทะเล) แต่ละกลุ่มจะสร้างแบบจำลองของตนเอง จากนั้นนำค่าประสิทธิภาพของทุกกลุ่มมาคำนวณหาค่าเฉลี่ยโดยรวม

3.7.2 ดำเนินการเปรียบเทียบผลลัพธ์ระหว่างระยะที่ 1 และ ระยะที่ 2

เพื่อให้เห็นผลของการเพิ่มฟีเจอร์เพิ่มเติมในระยะที่ 2 ผู้วิจัยจะดำเนินการเปรียบเทียบแบบจำลองที่มีประสิทธิภาพสูงสุดจากแต่ละระยะ โดยใช้ชุดข้อมูลทดสอบชุดเดียวกัน และประเมินด้วยตัวชี้วัดเดียวกัน ผลการเปรียบเทียบนี้ช่วยให้สามารถประเมินได้ว่า การใช้ข้อมูลจำนวนนักท่องเที่ยวเพียงอย่างเดียว (Phase 1) กับ การเพิ่มข้อมูลฟีเจอร์สนับสนุน (Phase 2) นั้น มีผลต่อประสิทธิภาพของแบบจำลองในลักษณะใดบ้าง

3.7.3 ดำเนินการเปรียบเทียบกับเทคนิคที่นิยมใช้ทั่วไป

เพื่อประเมินศักยภาพของแบบจำลองที่พัฒนาขึ้น ผู้วิจัยได้ดำเนินการเปรียบเทียบผลลัพธ์กับแบบจำลองพื้นฐาน (Baseline Model) ซึ่งใช้ข้อมูลจำนวนนักท่องเที่ยวของเดือนเดียวกันในปีที่ผ่านมาเป็นค่าพยากรณ์ (Previous Year Data) รวมถึงเทคนิคการพยากรณ์ที่นิยมใช้ทั่วไป ได้แก่ ARIMA, Moving Average, Linear Regression, Random Forest และ XGBoost โดยใช้ข้อมูลชุดเดียวกันในการทดลอง และนำค่าผลลัพธ์ที่ดีที่สุดจากแต่ละเทคนิคมาเปรียบเทียบกับแบบจำลองที่ดีที่สุดในงานวิจัยนี้ การเปรียบเทียบนี้มีวัตถุประสงค์เพื่อแสดงให้เห็นว่าแบบจำลองที่พัฒนาขึ้นมีแนวโน้มให้ผลลัพธ์ดีกว่าเทคนิคทั่วไปหรือไม่ และแตกต่างหรือใกล้เคียงกับเทคนิคส่วนใหญ่อย่างน้อยเพียงใด

3.7.4 ดำเนินการการวิเคราะห์ผลกระทบของฟิเจอร์ (Ablation Study)

ผู้วิจัยจะดำเนินการวิเคราะห์ความสำคัญของฟิเจอร์ โดยใช้แบบจำลองที่มีประสิทธิภาพดีที่สุดของ Phase 2 เป็นฐาน แล้วทำการสร้างแบบจำลองย่อยเพิ่มเติม โดยแต่ละชุดการทดลองจะมีการเพิ่มหรือลบฟิเจอร์บางตัว เช่น เดือน (Month), อุณหภูมิ (Temperature), ปริมาณรถยนต์ (Vehicle Count) และข้อมูลเชิงพื้นที่ (AreaFeatures) เพื่อประเมินว่าแต่ละฟิเจอร์มีผลต่อประสิทธิภาพของแบบจำลองมากน้อยเพียงใด ผลการวิเคราะห์นี้จะช่วยในการจัดลำดับความสำคัญของฟิเจอร์สำหรับการนำแบบจำลองไปใช้งาน และยังสะท้อนให้เห็นว่า ฟิเจอร์ใดเป็นองค์ประกอบสำคัญที่ทำให้แบบจำลองสามารถพยากรณ์ได้มีประสิทธิภาพที่สุด



บทที่ 4

ผลการวิจัย

จากการดำเนินงานตามขั้นตอนและวิธีการวิจัยที่ได้ระบุไว้ในบทที่ 3 งานวิจัยนี้ได้พัฒนาแบบจำลองโครงข่ายประสาทเทียมแบบเรียกซ้ำ (Recurrent Neural Networks: RNN) รวมถึงแบบจำลอง LSTM และ GRU โดยมีการทดสอบพารามิเตอร์สำคัญ ได้แก่ ช่วงข้อมูลย้อนหลัง (Look-back) และจำนวนหน่วยซ่อน (Hidden Units) รวมทั้งประเมินประสิทธิภาพของแบบจำลองด้วยค่าชี้วัด MAE, RMSE และ R^2 ซึ่งการทดลองแบ่งออกเป็น 2 ระยะ (Phase 1 และ Phase 2) เพื่อศึกษาผลของการเพิ่มพารามิเตอร์ที่แตกต่างกันต่อความสามารถในการพยากรณ์ โดยเปรียบเทียบผลลัพธ์ใน 3 แนวทางของแบบจำลอง ได้แก่ แบบจำลองรายอุทยาน (Individual Park Models : IPM) แบบจำลองรวม (Combined Models : CM) และแบบจำลองแบบจัดกลุ่ม ซึ่งแบ่งตามภูมิภาค (Clustered Regional Models: CRM) และลักษณะเฉพาะของอุทยาน (Clustered Unique Models: CUM) นอกจากนี้ ยังมีการเปรียบเทียบผลการทดลองกับเทคนิคการพยากรณ์อื่น ๆ และดำเนินการวิเคราะห์ผลกระทบของพารามิเตอร์ (Ablation Study) เพื่อศึกษาปัจจัยที่มีผลต่อประสิทธิภาพของแบบจำลองโดยรวม ผลการทดลองทั้งหมดมีดังนี้

4.1 ผลการทดลอง ระยะที่ 1

การทดลองระยะที่ 1 (Phase 1) เป็นการทดลองเบื้องต้นที่ใช้ข้อมูลจำนวนนักท่องเที่ยวเพียงพารามิเตอร์เดียว โดยทำการทดสอบด้วยค่าพารามิเตอร์ Look-back ที่ 3, 6 และ 9 เดือน และจำนวน Hidden Units ที่เพิ่มขึ้นทีละ 10 จนถึง 100 หน่วย เพื่อหาค่าที่ให้ผลลัพธ์ที่ดีที่สุดสำหรับแต่ละแบบจำลอง ผลลัพธ์ในระยะนี้เป็นการประเมินประสิทธิภาพเบื้องต้นของแต่ละแบบจำลองโดยไม่อาศัยปัจจัยเสริมอื่น โดยผลการทดลองทั้งหมดของระยะที่ 1 มีดังนี้

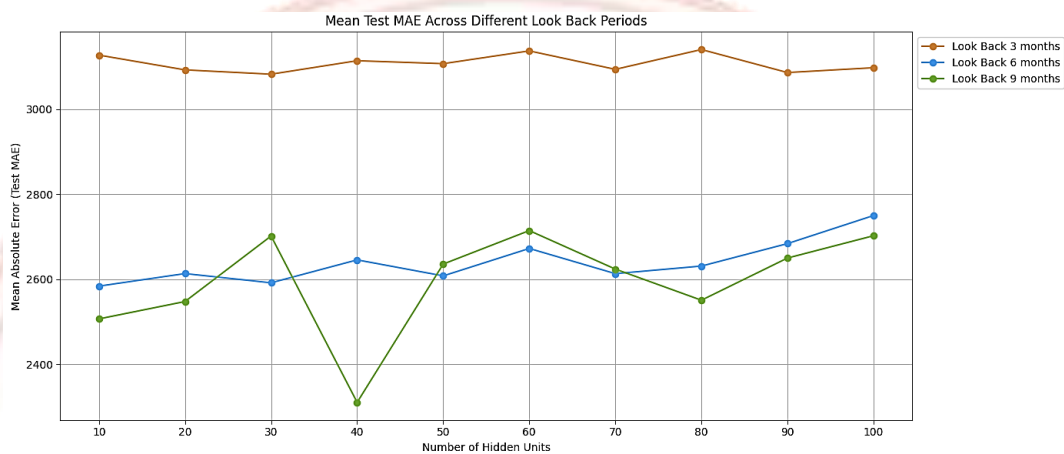
4.1.1 ผลการทดลองแบบจำลองรายอุทยานระยะที่ 1 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 20 หน่วย ให้ผลลัพธ์ดีที่สุด โดยมี Mean_Test_MAE = 3,093, Mean_Test_RMSE = 3,395 และ Mean_Test_ R^2 = 0.03 ซึ่งแม้ R^2 จะยังไม่สูงมาก แต่เมื่อเปรียบเทียบกับรูปแบบอื่น ๆ แล้ว รูปแบบนี้ยังมีแนวโน้มการพยากรณ์ที่ดีกว่าโดยรวม ขณะที่ Look-back ที่ยาวขึ้น เช่น 6 และ 9 เดือน ทำให้ค่า R^2 ลดลง โดยเฉพาะ Look-back = 9 เดือน ที่ค่า R^2 ตีลบในทุกชุดพารามิเตอร์ แสดงให้เห็นว่าแบบจำลองไม่สามารถจับแนวโน้มเชิงเวลาได้ดีในช่วงเวลาที่ยาวเกินไป โดยมีรายละเอียดของผลการทดลอง ดังนี้

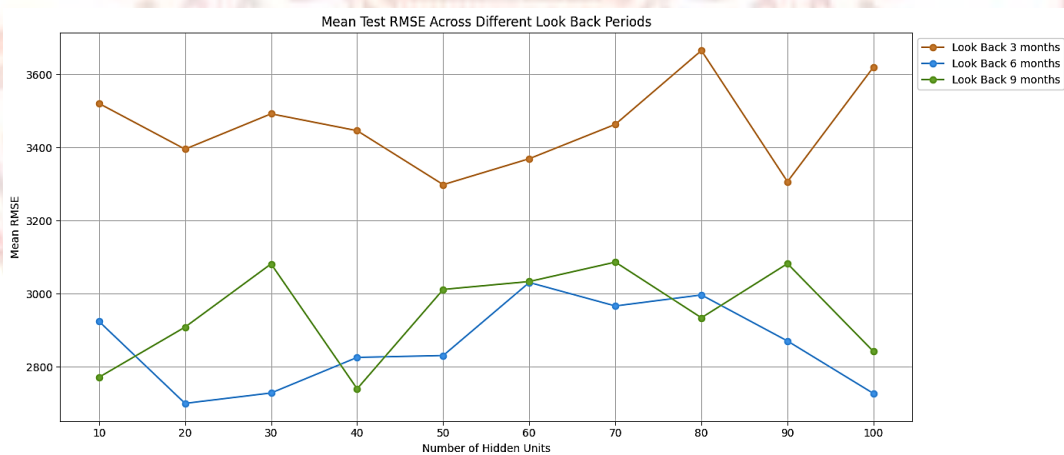
ตารางที่ 4 - 1 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	3,127	-0.01	3,520
3	20	3,093	0.03	3,395
3	30	3,082	0.00	3,492
3	40	3,114	-0.01	3,446
3	50	3,107	0.01	3,298
3	60	3,137	0.00	3,369
3	70	3,094	-0.02	3,463
3	80	3,141	-0.02	3,665
3	90	3,086	0.01	3,306
3	100	3,098	0.00	3,620
6	10	2,584	-0.41	2,925
6	20	2,613	-0.44	2,700
6	30	2,591	-0.52	2,729
6	40	2,645	-0.56	2,826
6	50	2,608	-0.50	2,831
6	60	2,672	-0.79	3,031
6	70	2,613	-1.39	2,967
6	80	2,631	-0.69	2,997
6	90	2,684	-0.59	2,871
6	100	2,750	-0.70	2,727
9	10	2,506	-2.36	2,772
9	20	2,547	-2.23	2,909
9	30	2,701	-2.53	3,082
9	40	2,310	-2.21	2,740
9	50	2,635	-2.41	3,012
9	60	2,714	-2.78	3,033
9	70	2,624	-3.35	3,086

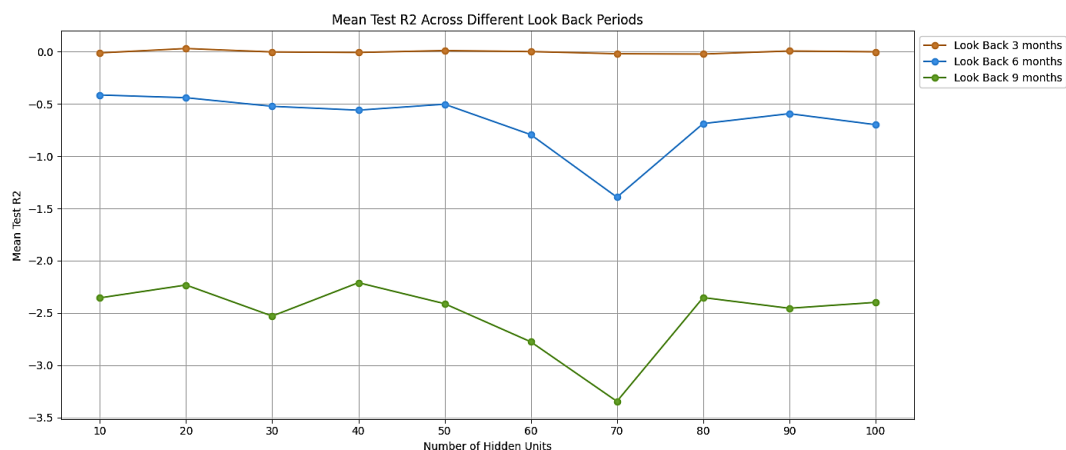
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	80	2,550	-2.35	2,934
9	90	2,650	-2.46	3,082
9	100	2,702	-2.40	2,842



ภาพที่ 4 - 1 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 2 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 3 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 1)

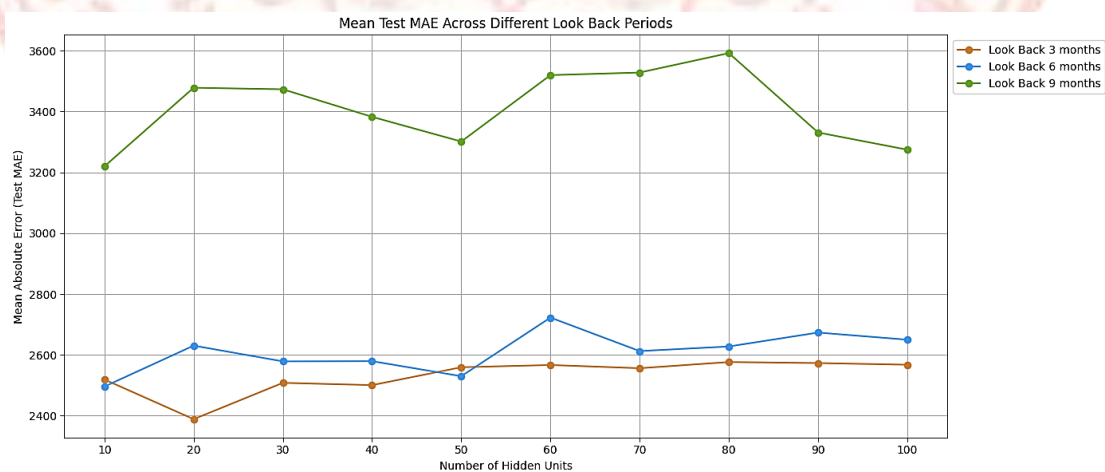
4.1.2 ผลการทดลองแบบจำลองรายอุทยานระยะที่ 1 โดยใช้แบบจำลอง LSTM

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 20 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,389, Mean_Test_RMSE = 3,275 และ Mean_Test_R² = 0.00 ซึ่งแม้ค่า R² จะไม่ได้สูงมาก แต่เมื่อเปรียบเทียบกับช่วง Look-back ที่นานขึ้น โดยเฉพาะ Look-back = 9 เดือน ที่ให้ค่า R² ต่ำสุดถึง -7.26 พบว่าแบบจำลองในรูปแบบนี้ยังคงสามารถจับแนวโน้มของข้อมูลได้ดีกว่าแบบอื่นภายใต้เงื่อนไขเดียวกัน โดยมีรายละเอียดของผลการทดลอง ดังนี้

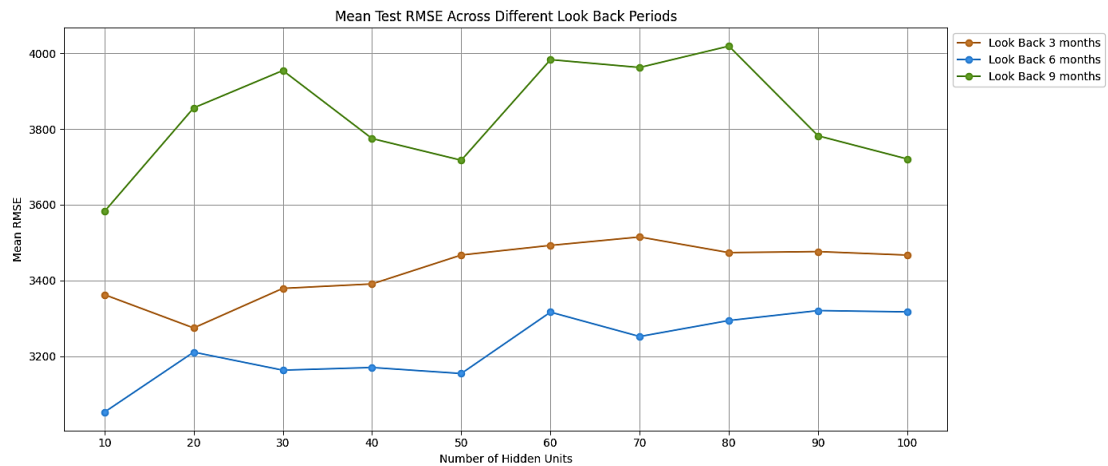
ตารางที่ 4 - 2 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	2,519	-0.06	3,363
3	20	2,389	0.00	3,275
3	30	2,508	-0.09	3,379
3	40	2,501	-0.08	3,391
3	50	2,559	-0.04	3,467
3	60	2,567	-0.11	3,493
3	70	2,556	-0.19	3,515
3	80	2,576	-0.12	3,474
3	90	2,573	-0.06	3,476
3	100	2,568	-0.06	3,467
6	10	2,496	-0.79	3,053
6	20	2,630	-1.04	3,211

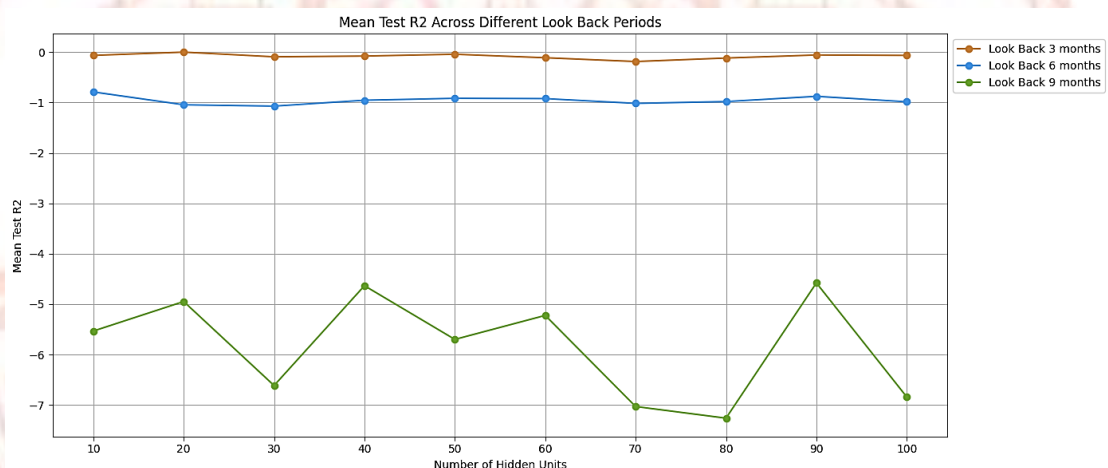
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
6	30	2,579	-1.07	3,163
6	40	2,579	-0.95	3,170
6	50	2,530	-0.92	3,155
6	60	2,723	-0.92	3,317
6	70	2,612	-1.02	3,252
6	80	2,628	-0.98	3,294
6	90	2,673	-0.88	3,321
6	100	2,650	-0.98	3,317
9	10	3,220	-5.53	3,582
9	20	3,478	-4.95	3,856
9	30	3,473	-6.62	3,954
9	40	3,382	-4.64	3,775
9	50	3,301	-5.70	3,718
9	60	3,519	-5.22	3,983
9	70	3,528	-7.03	3,962
9	80	3,592	-7.26	4,019
9	90	3,331	-4.58	3,782
9	100	3,274	-6.84	3,721



ภาพที่ 4 - 4 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 5 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 6 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 1)

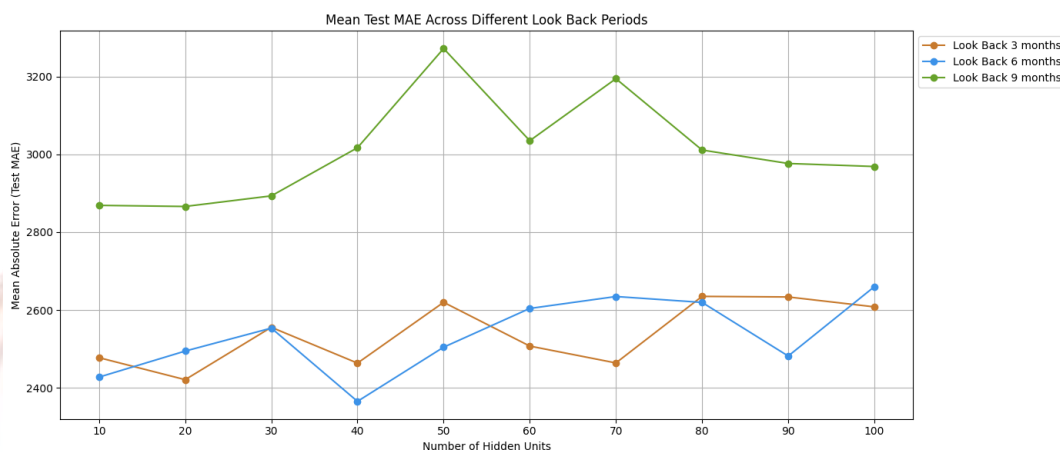
4.1.3 ผลการทดลองแบบจำลองรายอุทยานระยะที่ 1 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 20 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,421, Mean_Test_RMSE = 3,290 และ Mean_Test_R² = 0.06 ซึ่งเป็นค่าที่สูงที่สุดเมื่อเทียบกับรูปแบบอื่นในกลุ่มเดียวกัน ขณะที่แบบจำลองที่มี Look-back ยาวขึ้น เช่น 6 และ 9 เดือน พบว่าค่า R² มีแนวโน้มลดลงอย่างต่อเนื่อง และในช่วง Look-back = 9 เดือน ค่า R² ติดลบในทุกกรณี โดยมีค่าต่ำที่สุดถึง -6.92 พร้อมกับ RMSE ที่สูงขึ้น แสดงให้เห็นถึงความสามารถที่ลดลงในการจับแนวโน้มของข้อมูลเมื่อระยะเวลาที่ย้อนกลับยาวเกินไป โดยมีรายละเอียดของผลการทดลอง ดังนี้

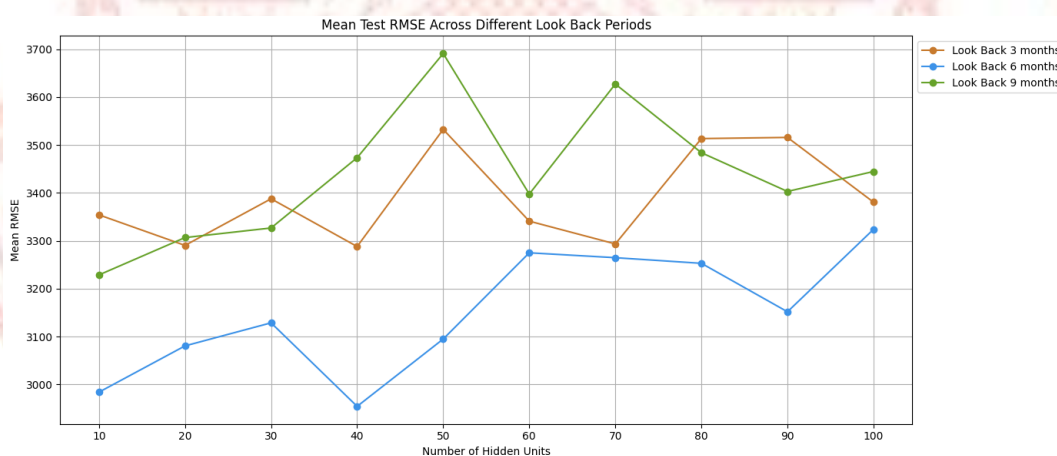
ตารางที่ 4 - 3 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	2,478	-0.12	3,354
3	20	2,421	0.06	3,290
3	30	2,556	-0.05	3,388
3	40	2,464	-0.02	3,288
3	50	2,620	-0.07	3,533
3	60	2,508	-0.25	3,341
3	70	2,464	-0.04	3,293
3	80	2,635	-0.08	3,513
3	90	2,634	-0.01	3,516
3	100	2,608	-0.11	3,381
6	10	2,428	-0.70	2,984
6	20	2,495	-0.80	3,081
6	30	2,554	-0.83	3,129
6	40	2,366	-0.69	2,954
6	50	2,505	-0.87	3,095
6	60	2,604	-1.11	3,275
6	70	2,635	-0.99	3,265
6	80	2,620	-1.20	3,253
6	90	2,482	-1.08	3,152
6	100	2,660	-1.42	3,324
9	10	2,869	-5.46	3,229
9	20	2,866	-3.85	3,307
9	30	2,893	-4.57	3,327
9	40	3,017	-5.49	3,473
9	50	3,272	-5.56	3,691
9	60	3,035	-5.69	3,398
9	70	3,194	-6.92	3,627
9	80	3,011	-6.80	3,484

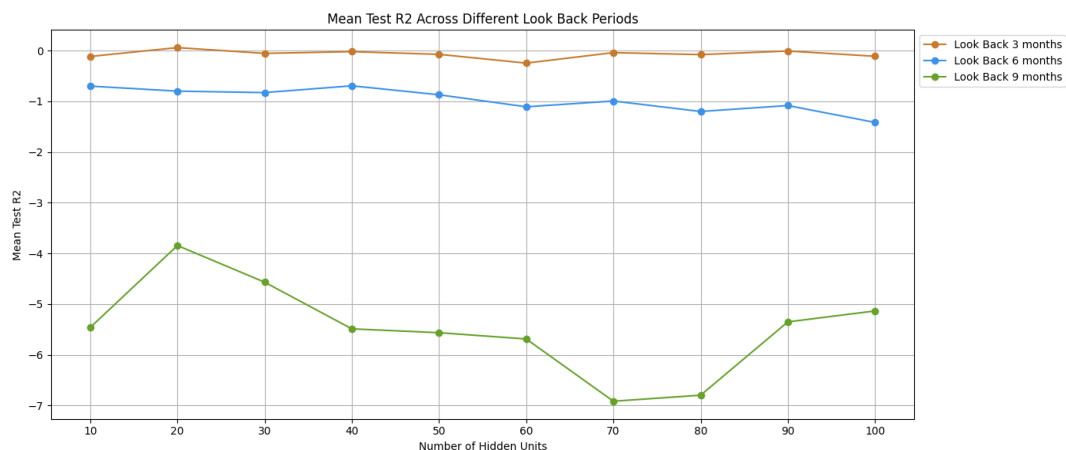
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	90	2,976	-5.35	3,403
9	100	2,969	-5.14	3,445



ภาพที่ 4 - 7 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 8 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 9 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 1)

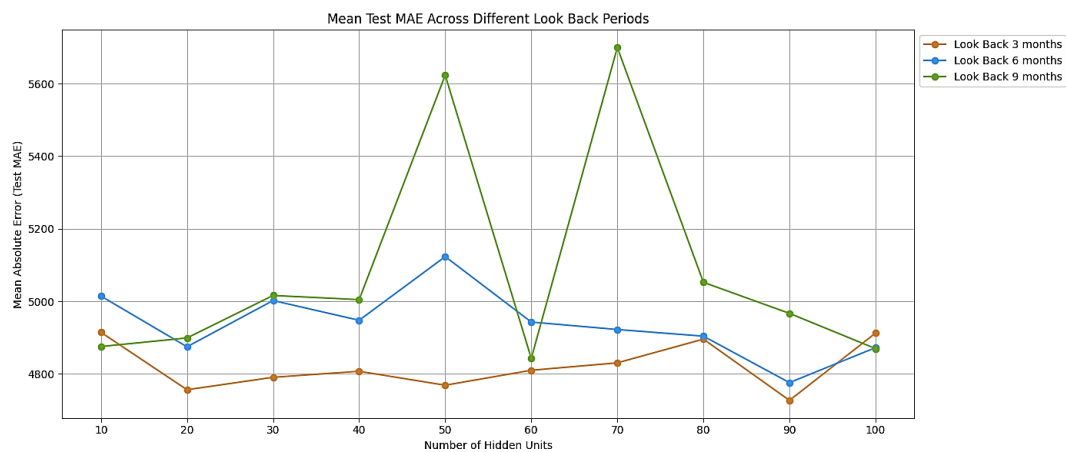
4.1.4 ผลการทดลองแบบจำลองรวมระยะที่ 1 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 100 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,912, Mean_Test_RMSE = 5,617 และ Mean_Test_ R^2 = 0.93 ซึ่งแม้ค่า MAE จะไม่ต่ำที่สุด แต่เมื่อพิจารณาร่วมกับค่า RMSE ที่ต่ำที่สุดในกลุ่ม และค่า R^2 ที่สูงที่สุด แสดงว่าแบบจำลองนี้ให้สมดุลที่ดีระหว่างความแม่นยำเชิงตัวเลขและการสะท้อนแนวโน้มข้อมูลจริงได้ชัดเจนที่สุด ขณะที่แบบจำลองที่มี Look-back = 6 และ 9 เดือน แม้จะมีค่า R^2 อยู่ในระดับใกล้เคียงกัน (ประมาณ 0.90-0.92) แต่ไม่สามารถให้ค่า RMSE ที่ต่ำได้เท่าแบบจำลองที่เลือกนี้ โดยมีรายละเอียดของผลการทดลอง ดังนี้

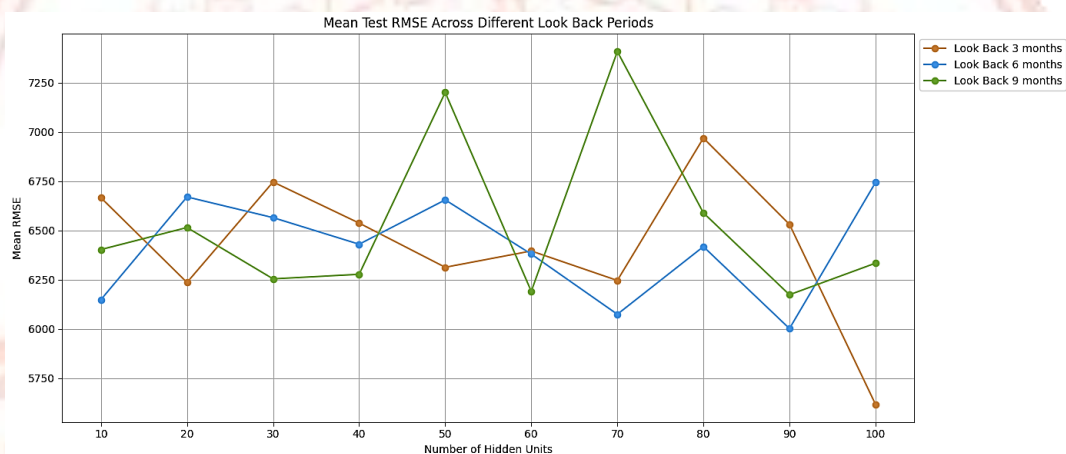
ตารางที่ 4 - 4 ผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_ R^2	Mean_Test_RMSE
3	10	4,914	0.90	6,665
3	20	4,756	0.91	6,236
3	30	4,791	0.90	6,746
3	40	4,807	0.91	6,537
3	50	4,769	0.91	6,313
3	60	4,810	0.91	6,395
3	70	4,830	0.91	6,246
3	80	4,896	0.89	6,969
3	90	4,727	0.91	6,530

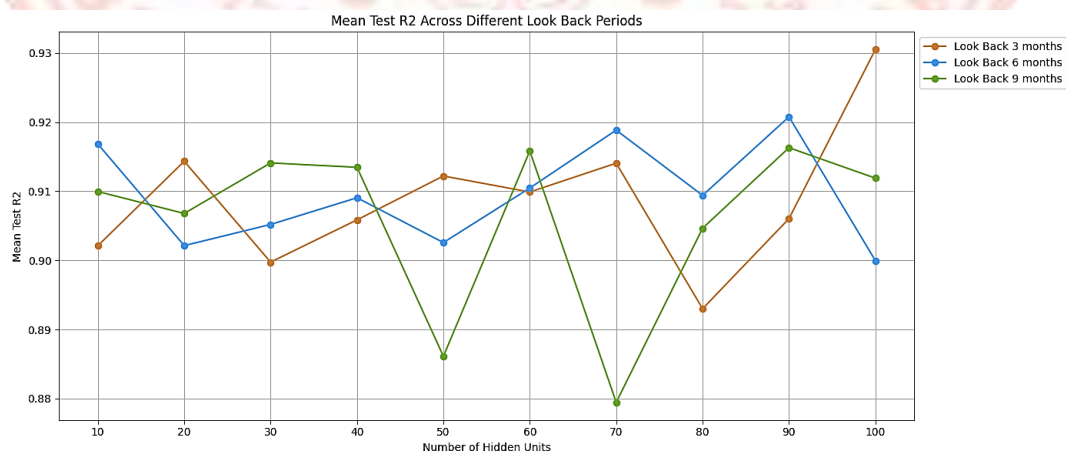
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	100	4,912	0.93	5,617
6	10	5,014	0.92	6,148
6	20	4,875	0.90	6,670
6	30	5,002	0.91	6,565
6	40	4,948	0.91	6,430
6	50	5,123	0.90	6,655
6	60	4,943	0.91	6,380
6	70	4,922	0.92	6,074
6	80	4,904	0.91	6,417
6	90	4,776	0.92	6,001
6	100	4,872	0.90	6,743
9	10	4,875	0.91	6,403
9	20	4,899	0.91	6,514
9	30	5,016	0.91	6,254
9	40	5,004	0.91	6,277
9	50	5,623	0.89	7,200
9	60	4,843	0.92	6,190
9	70	5,699	0.88	7,408
9	80	5,052	0.90	6,589
9	90	4,967	0.92	6,173
9	100	4,870	0.91	6,333



ภาพที่ 4 - 10 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม
โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 11 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม
โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 12 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรวม
โดยใช้ RNN (Phase 1)

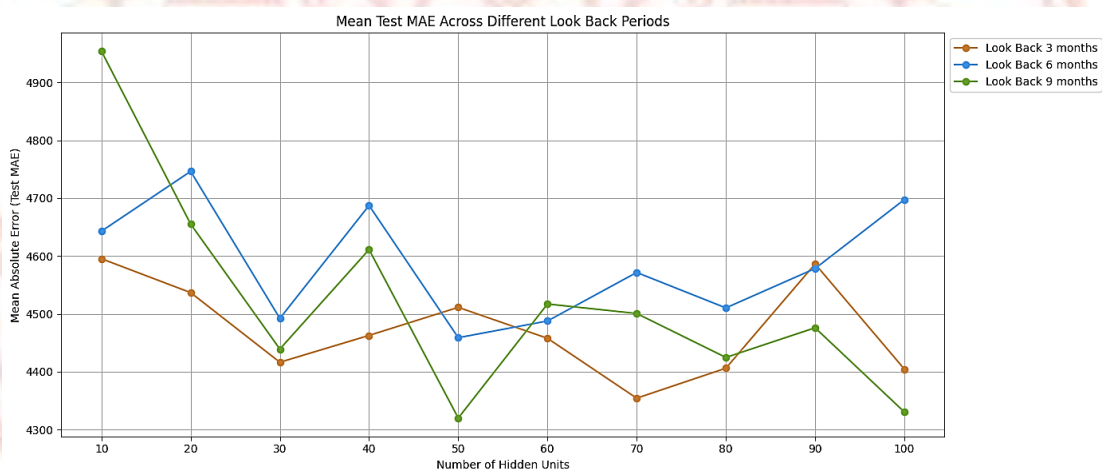
4.1.5 ผลการทดลองแบบจำลองรวมระยะที่ 1 โดยใช้แบบจำลอง LSTM

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 70 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,501, Mean_Test_RMSE = 5,498 และ Mean_Test_R² = 0.93 ซึ่งแม้ค่า MAE จะไม่ต่ำที่สุดในกลุ่ม แต่ค่า R² และ RMSE ที่โดดเด่นแสดงถึงความสามารถในการพยากรณ์แนวโน้มโดยรวมของข้อมูลได้ดีและเสถียรกว่ารูปแบบอื่น นอกจากนี้ แบบจำลองที่ใช้ Look-back = 6 เดือน และ Hidden Units = 90 หน่วย ก็ให้ค่า R² เท่ากันที่ 0.93 โดยมีค่า RMSE ใกล้เคียงกันที่ 5,731 สะท้อนว่า LSTM สามารถเรียนรู้ลำดับข้อมูลในระยะยาวได้ดีเมื่อใช้พารามิเตอร์ที่เหมาะสม โดยมีรายละเอียดของผลการทดลอง ดังนี้

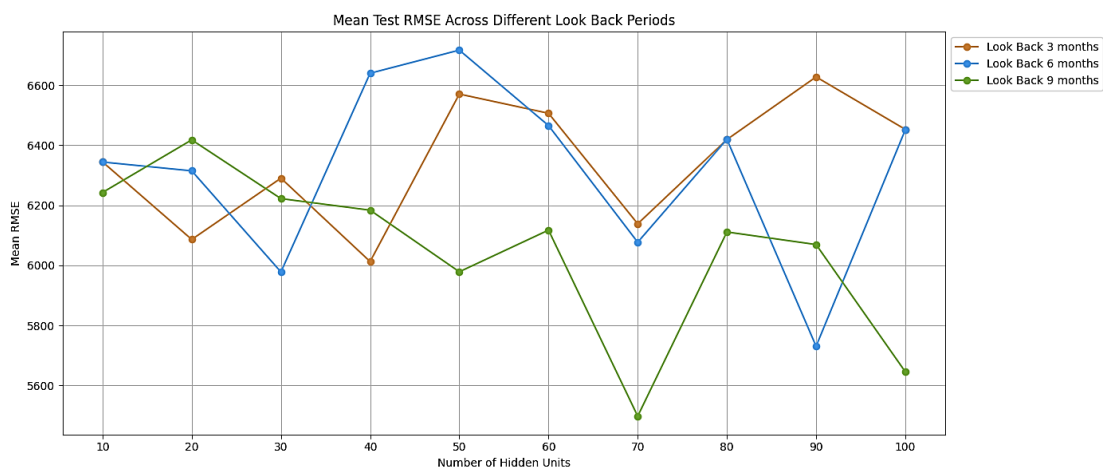
ตารางที่ 4 - 5 ผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,595	0.91	6,345
3	20	4,537	0.92	6,086
3	30	4,416	0.91	6,290
3	40	4,463	0.92	6,013
3	50	4,511	0.90	6,571
3	60	4,458	0.91	6,507
3	70	4,354	0.92	6,139
3	80	4,406	0.91	6,419
3	90	4,586	0.90	6,628
3	100	4,405	0.91	6,453
6	10	4,643	0.91	6,344
6	20	4,746	0.91	6,315
6	30	4,492	0.92	5,978
6	40	4,687	0.90	6,640
6	50	4,459	0.90	6,717
6	60	4,488	0.91	6,466
6	70	4,572	0.92	6,077
6	80	4,510	0.91	6,420
6	90	4,578	0.93	5,731
6	100	4,697	0.91	6,451

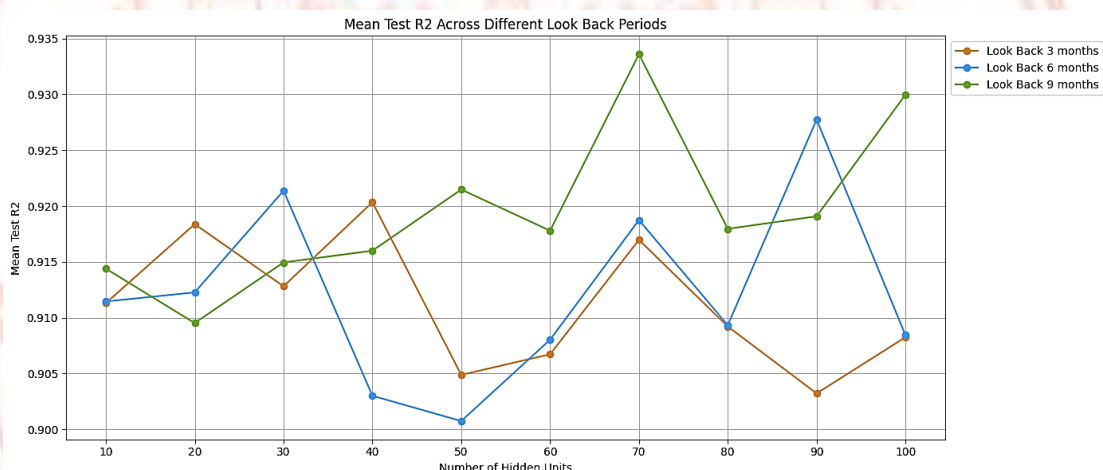
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	10	4,954	0.91	6,243
9	20	4,655	0.91	6,418
9	30	4,439	0.91	6,222
9	40	4,611	0.92	6,184
9	50	4,320	0.92	5,979
9	60	4,517	0.92	6,118
9	70	4,501	0.93	5,498
9	80	4,425	0.92	6,111
9	90	4,476	0.92	6,069
9	100	4,330	0.93	5,646



ภาพที่ 4 - 13 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม
โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 14 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม
โดยใช้ LSTM (Phase 1)



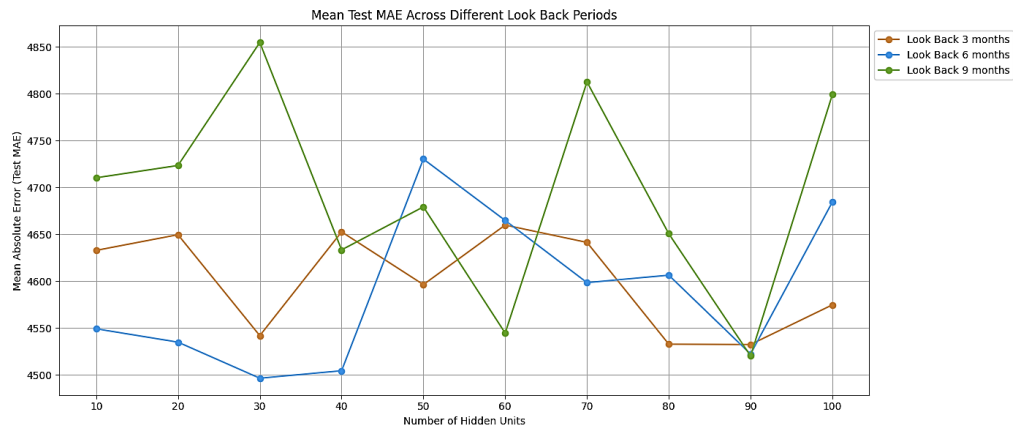
ภาพที่ 4 - 15 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรวม
โดยใช้ LSTM (Phase 1)

4.1.6 ผลการทดลองแบบจำลองรวมระยะที่ 1 โดยใช้แบบจำลอง GRU

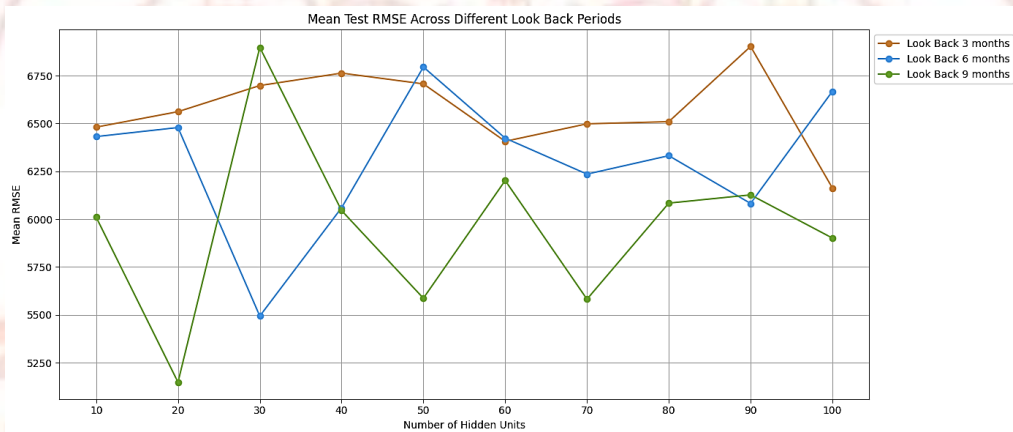
แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 20 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,723, Mean_Test_RMSE = 5,148 และ Mean_Test_R² = 0.94 ซึ่งเป็นค่า R² สูงที่สุดในชุดการทดลองนี้ และมีค่า RMSE ต่ำที่สุดเช่นกัน แม้ว่า MAE จะไม่ต่ำที่สุด แต่ค่าประสิทธิภาพโดยรวมถือว่าดีกว่ารูปแบบอื่นของ GRU โดยเฉพาะเมื่อใช้ Look-back ระยะยาว ร่วมกับ Hidden Units ที่เหมาะสม นอกจากนี้ แบบจำลองอื่นที่ใช้ Look-back เดียวกัน เช่น Hidden Units = 50 และ 70 ก็ให้ค่า R² ใกล้เคียงกัน (0.93) และ RMSE ต่ำกว่า 5,600 แสดงถึงความเสถียรในการพยากรณ์ของ GRU เมื่อใช้พารามิเตอร์ที่เหมาะสมในโครงสร้างแบบรวม โดยมีรายละเอียดของผลการทดลอง ดังนี้

ตารางที่ 4 - 6 ผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 1)

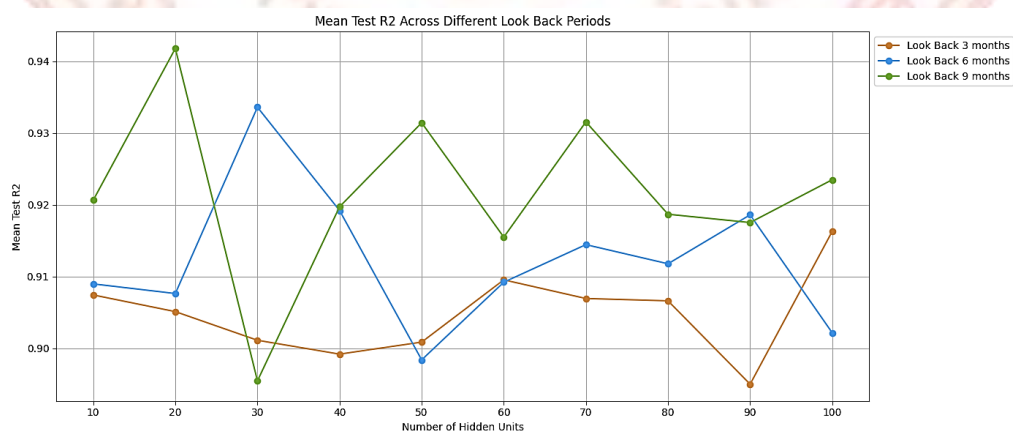
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,633	0.91	6,481
3	20	4,650	0.91	6,562
3	30	4,542	0.90	6,698
3	40	4,653	0.90	6,763
3	50	4,596	0.90	6,707
3	60	4,660	0.91	6,407
3	70	4,641	0.91	6,498
3	80	4,533	0.91	6,509
3	90	4,532	0.90	6,902
3	100	4,575	0.92	6,163
6	10	4,549	0.91	6,431
6	20	4,535	0.91	6,479
6	30	4,496	0.93	5,494
6	40	4,504	0.92	6,060
6	50	4,730	0.90	6,795
6	60	4,665	0.91	6,423
6	70	4,598	0.91	6,235
6	80	4,606	0.91	6,332
6	90	4,522	0.92	6,082
6	100	4,684	0.90	6,668
9	10	4,710	0.92	6,010
9	20	4,723	0.94	5,148
9	30	4,855	0.90	6,898
9	40	4,633	0.92	6,045
9	50	4,679	0.93	5,587
9	60	4,544	0.92	6,202
9	70	4,812	0.93	5,582
9	80	4,651	0.92	6,084
9	90	4,520	0.92	6,127
9	100	4,799	0.92	5,901



ภาพที่ 4 - 16 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม
โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 17 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม
โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 18 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรวม
โดยใช้ GRU (Phase 1)

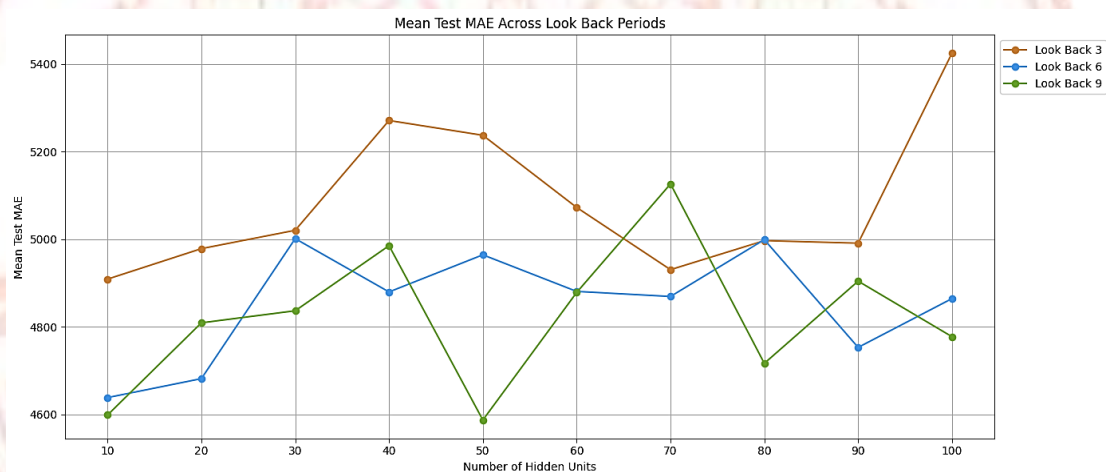
4.1.7 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามภูมิภาค ระยะที่ 1 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 60 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 5,072, Mean_Test_RMSE = 5,398 และ Mean_Test_R² = 0.91 ซึ่งเป็นค่าที่ดีที่สุดในการทดลองนี้ สะท้อนถึงความสามารถของแบบจำลองในการอธิบายแนวโน้มของข้อมูลในระดับกลุ่มภูมิภาคได้ดี แม้บางรูปแบบที่ใช้ Look-back = 9 เดือน เช่น Hidden Units = 50 จะให้ค่า MAE และ RMSE ต่ำกว่าเล็กน้อย (Mean_Test_MAE = 4,587, Mean_Test_RMSE = 5,238) แต่ค่า R² กลับต่ำกว่าที่ 0.85 โดยมีรายละเอียดของผลการทดลอง ดังนี้

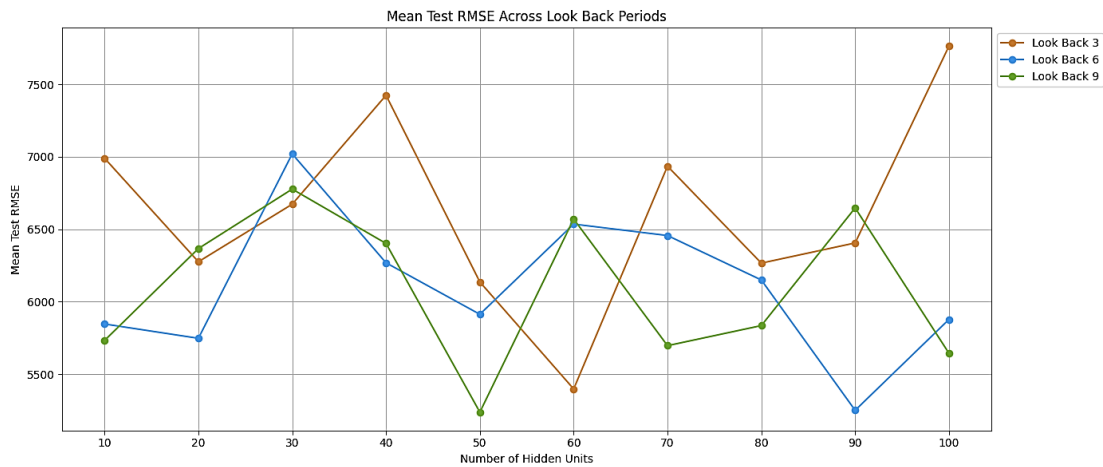
ตารางที่ 4 - 7 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,909	0.86	6,990
3	20	4,978	0.86	6,276
3	30	5,020	0.86	6,673
3	40	5,271	0.84	7,424
3	50	5,237	0.85	6,136
3	60	5,072	0.91	5,398
3	70	4,930	0.86	6,935
3	80	4,997	0.87	6,266
3	90	4,991	0.87	6,405
3	100	5,424	0.76	7,764
6	10	4,639	0.82	5,846
6	20	4,682	0.81	5,747
6	30	5,001	0.73	7,020
6	40	4,880	0.78	6,268
6	50	4,964	0.78	5,913
6	60	4,881	0.78	6,535
6	70	4,869	0.78	6,456
6	80	4,999	0.75	6,150
6	90	4,753	0.83	5,250
6	100	4,864	0.77	5,877
9	10	4,599	0.81	5,732

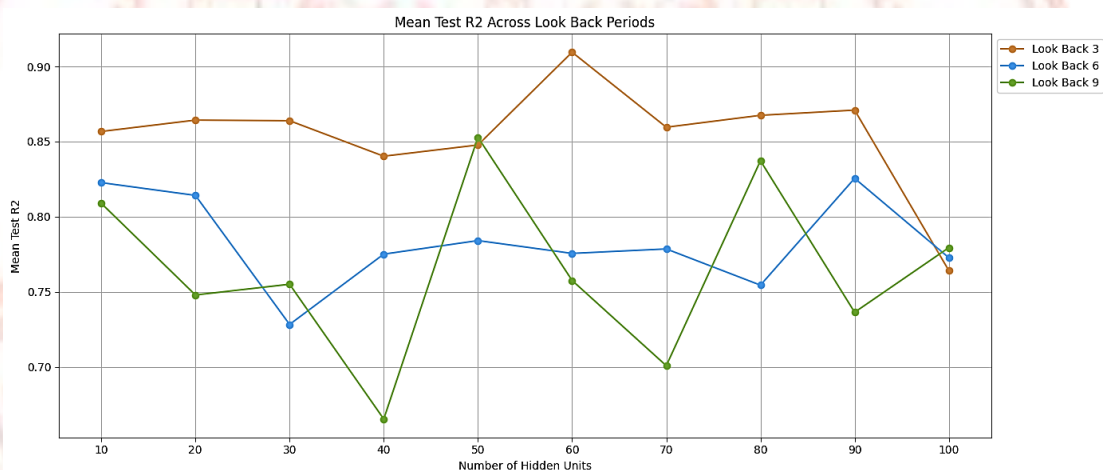
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	20	4,809	0.75	6,367
9	30	4,837	0.76	6,777
9	40	4,985	0.67	6,402
9	50	4,587	0.85	5,238
9	60	4,878	0.76	6,571
9	70	5,126	0.70	5,696
9	80	4,717	0.84	5,835
9	90	4,904	0.74	6,647
9	100	4,778	0.78	5,644



ภาพที่ 4 - 19 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 20 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 1)



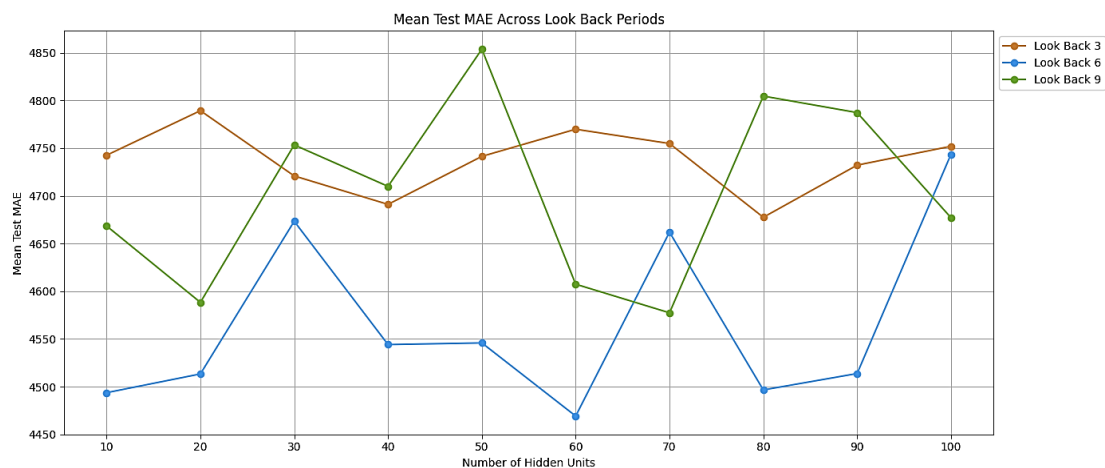
ภาพที่ 4 - 21 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 1)

4.1.8 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามภูมิภาค ระยะที่ 1 โดยใช้แบบจำลอง LSTM

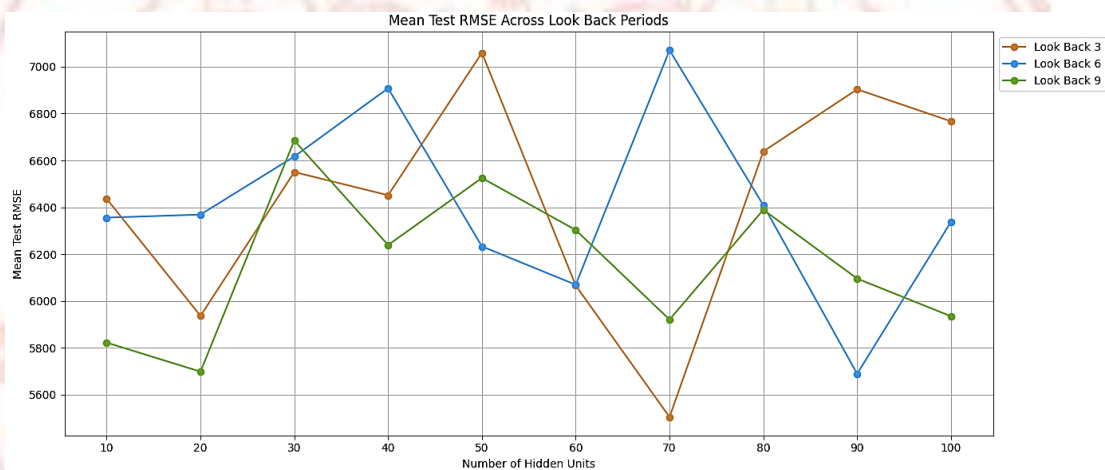
แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 70 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,755, Mean_Test_RMSE = 5,505 และ Mean_Test_R² = 0.88 ซึ่งแม้จะมีบางค่าที่ใช้ Look-back ที่ยาวขึ้น เช่น 9 เดือน และ Hidden Units = 70 แล้วให้ค่า R² อยู่ที่ 0.82 และ MAE ต่ำกว่าที่ 4,577 แต่ค่าความคลาดเคลื่อน RMSE ยังคงสูงกว่า แสดงให้เห็นว่าการใช้ Look-back ที่สั้นกว่ายังคงให้แนวโน้มการพยากรณ์ที่ดีกว่าโดยรวมในการทดลองนี้ โดยมีรายละเอียดของผลการทดลอง ดังนี้

ตารางที่ 4 - 8 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 1)

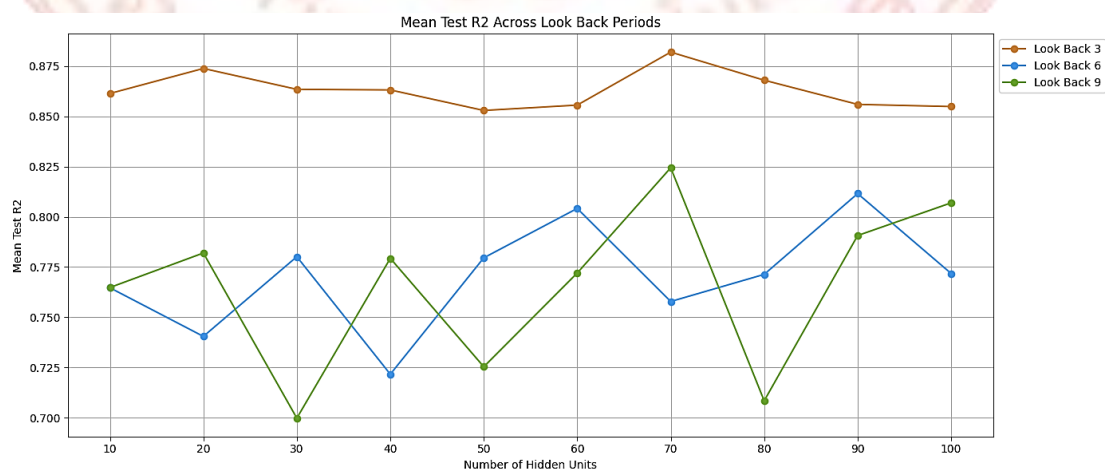
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,743	0.86	6,436
3	20	4,789	0.87	5,936
3	30	4,721	0.86	6,550
3	40	4,691	0.86	6,452
3	50	4,741	0.85	7,059
3	60	4,770	0.86	6,066
3	70	4,755	0.88	5,505
3	80	4,678	0.87	6,639
3	90	4,732	0.86	6,904
3	100	4,752	0.85	6,767
6	10	4,494	0.76	6,356
6	20	4,513	0.74	6,369
6	30	4,673	0.78	6,617
6	40	4,544	0.72	6,908
6	50	4,546	0.78	6,233
6	60	4,469	0.80	6,069
6	70	4,662	0.76	7,071
6	80	4,497	0.77	6,409
6	90	4,514	0.81	5,689
6	100	4,744	0.77	6,338
9	10	4,668	0.76	5,822
9	20	4,588	0.78	5,698
9	30	4,753	0.70	6,685
9	40	4,710	0.78	6,239
9	50	4,854	0.73	6,525
9	60	4,607	0.77	6,303
9	70	4,577	0.82	5,921
9	80	4,805	0.71	6,389
9	90	4,787	0.79	6,096
9	100	4,677	0.81	5,935



ภาพที่ 4 - 22 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 23 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 24 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 1)

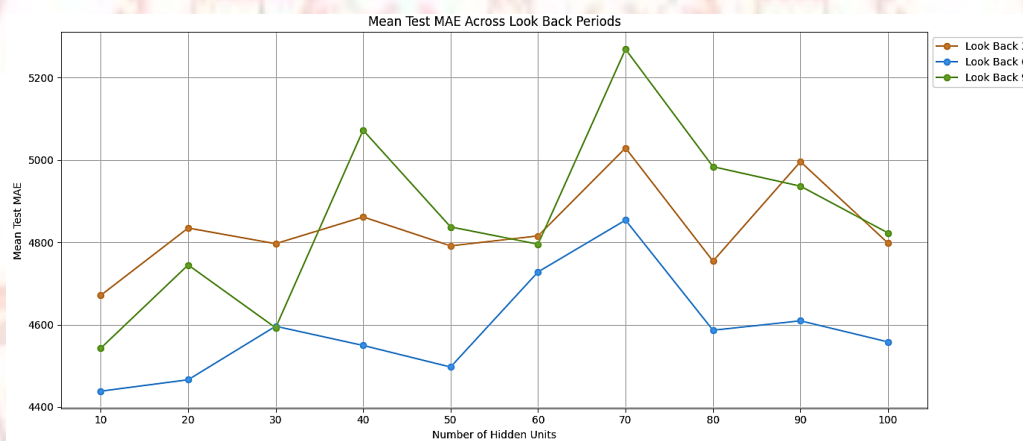
4.1.9 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามภูมิภาค ระยะที่ 1 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 40 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,862, Mean_Test_RMSE = 6,202 และ Mean_Test_R² = 0.87 ซึ่งแสดงว่าแบบจำลองสามารถจับความสัมพันธ์และแนวโน้มของค่าจริงได้ดีและสอดคล้องกันมากที่สุดในกลุ่มนี้ แม้แบบจำลองที่ใช้ Hidden Units = 100 จะให้ค่า R² เท่ากันที่ 0.87 แต่ RMSE กลับสูงกว่าชัดเจน แสดงถึงความคลาดเคลื่อนที่มากกว่า นอกจากนี้ Look-back ที่ยาวขึ้น เช่น 6 และ 9 เดือน ยังส่งผลให้ค่า R² ลดลง โดยบางค่าต่ำกว่า 0.70 แสดงถึงแนวโน้มการพยากรณ์ที่ด้อยลงเมื่อใช้ระยะเวลาย้อนหลังนานเกินไป โดยมีรายละเอียดของผลการทดลอง ดังนี้

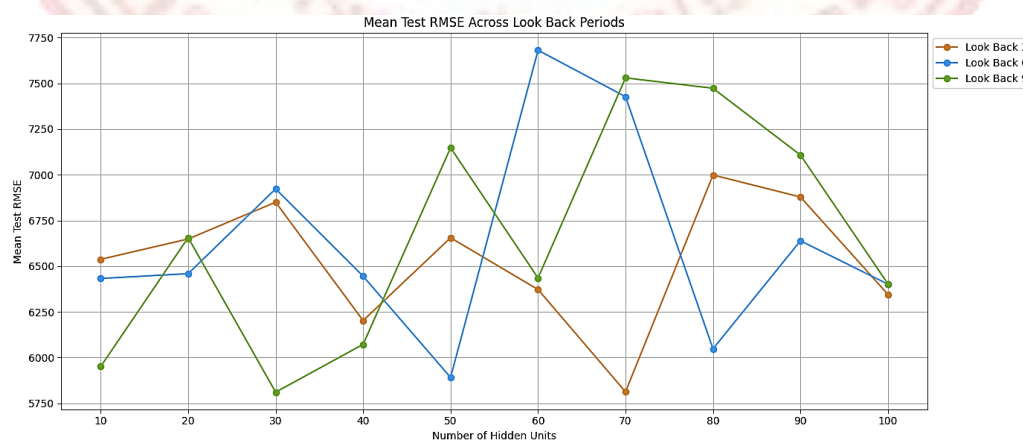
ตารางที่ 4 - 9 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,671	0.87	6,537
3	20	4,835	0.86	6,649
3	30	4,797	0.85	6,850
3	40	4,862	0.87	6,202
3	50	4,791	0.86	6,655
3	60	4,816	0.87	6,372
3	70	5,029	0.85	5,812
3	80	4,754	0.85	6,998
3	90	4,996	0.86	6,878
3	100	4,798	0.87	6,345
6	10	4,438	0.79	6,432
6	20	4,466	0.76	6,459
6	30	4,596	0.67	6,923
6	40	4,549	0.78	6,445
6	50	4,497	0.76	5,891
6	60	4,728	0.64	7,681
6	70	4,854	0.63	7,427
6	80	4,586	0.79	6,048
6	90	4,609	0.80	6,639
6	100	4,558	0.77	6,401

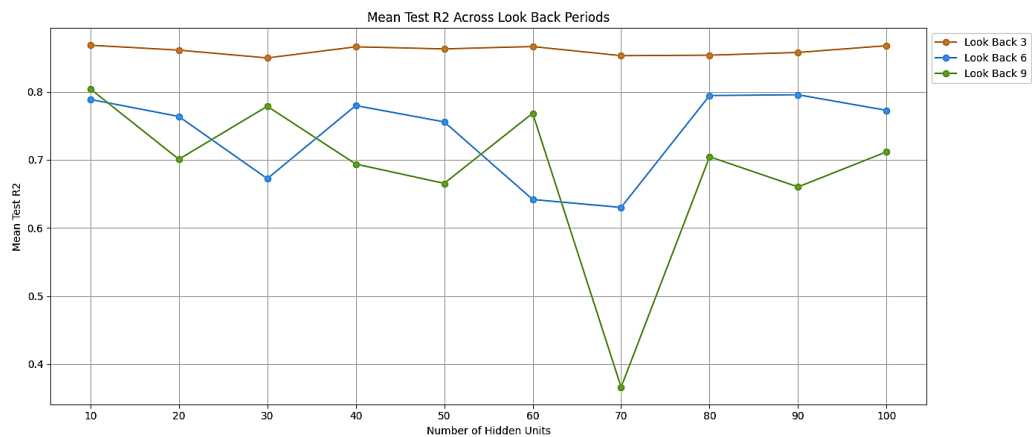
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	10	4,543	0.80	5,951
9	20	4,745	0.70	6,656
9	30	4,592	0.78	5,810
9	40	5,072	0.69	6,072
9	50	4,837	0.67	7,147
9	60	4,795	0.77	6,435
9	70	5,269	0.37	7,530
9	80	4,984	0.70	7,473
9	90	4,936	0.66	7,108
9	100	4,823	0.71	6,403



ภาพที่ 4 - 25 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 26 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 27 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 1)

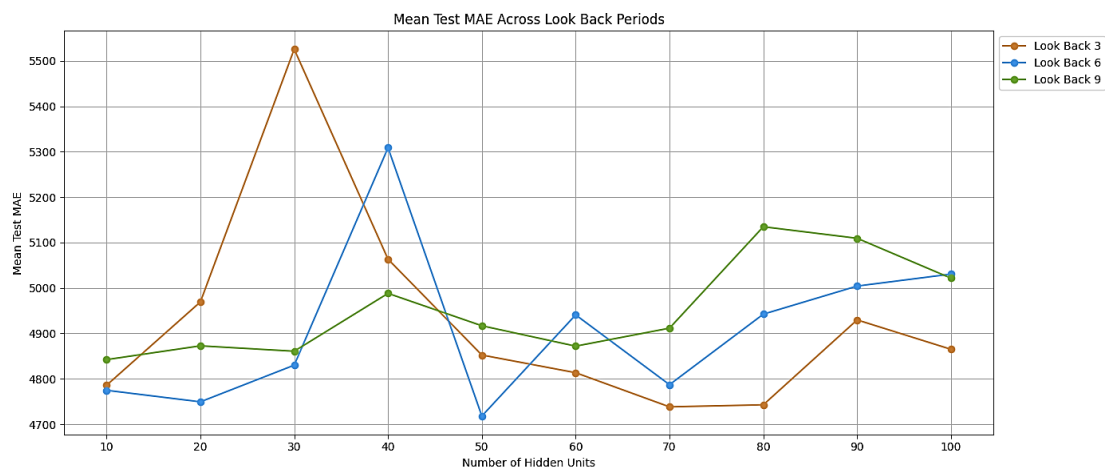
4.1.10 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน ระยะที่ 1 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 80 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,743, Mean_Test_RMSE = 6,254 และ Mean_Test_R² = 0.72 ซึ่งแสดงถึงความสามารถของแบบจำลองในการพยากรณ์ข้อมูลได้อย่างสอดคล้องกับค่าจริง โดยเมื่อเปรียบเทียบกับพารามิเตอร์อื่น เช่น Look-back = 9 เดือน และ Hidden Units = 80 ที่ให้ R² ใกล้เคียงกันที่ 0.71 แต่กลับมี RMSE สูงกว่า จึงแสดงว่ารูปแบบที่ใช้ Look-back สั้นกว่าให้ผลพยากรณ์แม่นยำกว่า ทั้งในแง่ความคลาดเคลื่อนและความต่อเนื่องของแนวโน้ม โดยมีรายละเอียดของผลการทดลอง ดังนี้

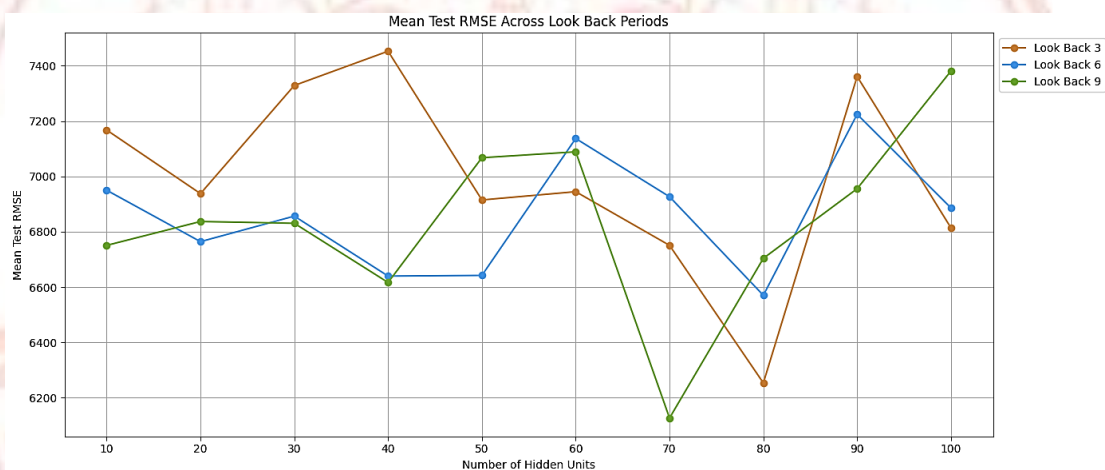
ตารางที่ 4 - 10 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,786	0.66	7,167
3	20	4,969	0.63	6,938
3	30	5,526	0.64	7,328
3	40	5,063	0.55	7,452
3	50	4,852	0.67	6,915
3	60	4,813	0.68	6,945
3	70	4,738	0.71	6,751

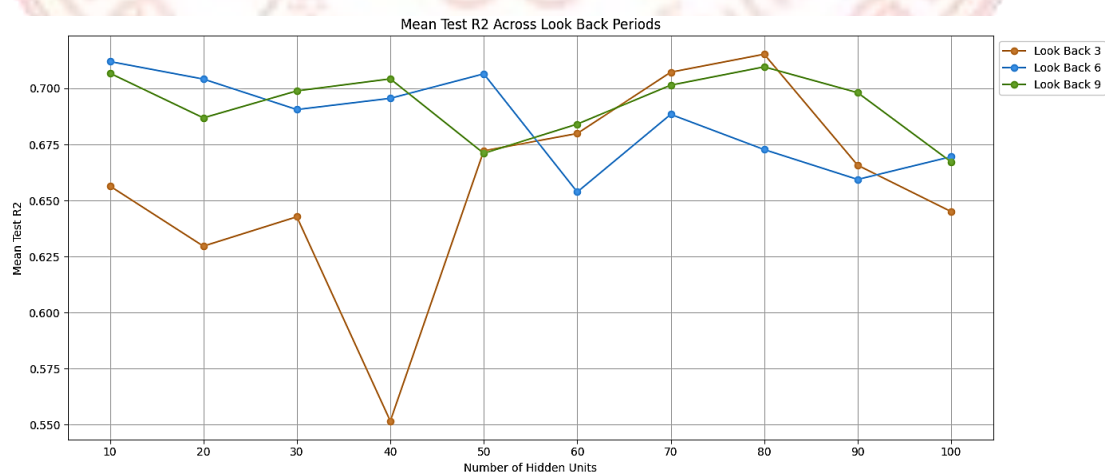
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	80	4,743	0.72	6,254
3	90	4,930	0.67	7,361
3	100	4,865	0.65	6,814
6	10	4,775	0.71	6,950
6	20	4,749	0.70	6,764
6	30	4,830	0.69	6,856
6	40	5,309	0.70	6,640
6	50	4,718	0.71	6,642
6	60	4,941	0.65	7,137
6	70	4,787	0.69	6,926
6	80	4,942	0.67	6,570
6	90	5,004	0.66	7,224
6	100	5,031	0.67	6,886
9	10	4,842	0.71	6,751
9	20	4,873	0.69	6,837
9	30	4,861	0.70	6,830
9	40	4,988	0.70	6,616
9	50	4,917	0.67	7,067
9	60	4,872	0.68	7,089
9	70	4,912	0.70	6,127
9	80	5,135	0.71	6,704
9	90	5,109	0.70	6,955
9	100	5,022	0.67	7,381



ภาพที่ 4 - 28 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 29 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 1)



ภาพที่ 4 - 30 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 1)

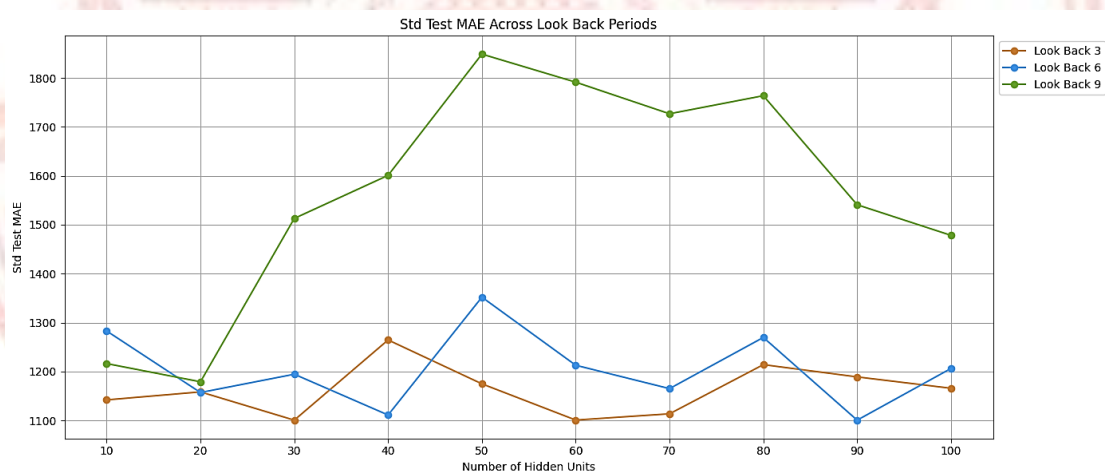
4.1.11 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน ระยะที่ 1 โดยใช้แบบจำลอง LSTM

แบบจำลองที่มี Look-back = 6 เดือน และ Hidden Units = 10 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,562, Mean_Test_RMSE = 6,677 และ Mean_Test_R² = 0.72 ซึ่งถือว่าอยู่ในระดับสูงและมีความสมดุลระหว่างค่าความคลาดเคลื่อนและความสามารถในการอธิบายแนวโน้มของข้อมูลจริง แม้จะมีแบบจำลองที่ใช้ Look-back = 9 เดือน และ Hidden Units = 90 ซึ่งได้ค่า R² สูงถึง 0.80 แต่ค่าความคลาดเคลื่อน Mean_Test_MAE = 4,988 ยังคงสูงกว่า และผลลัพธ์ในกลุ่ม Look-back 9 เดือนมีความผันผวนมากกว่าเมื่อเทียบกับ Look-back 6 เดือน ที่ให้ผลลัพธ์ที่เสถียรและคงที่มากกว่าในหลายชุดพารามิเตอร์ โดยมีรายละเอียดของการทดลอง ดังนี้

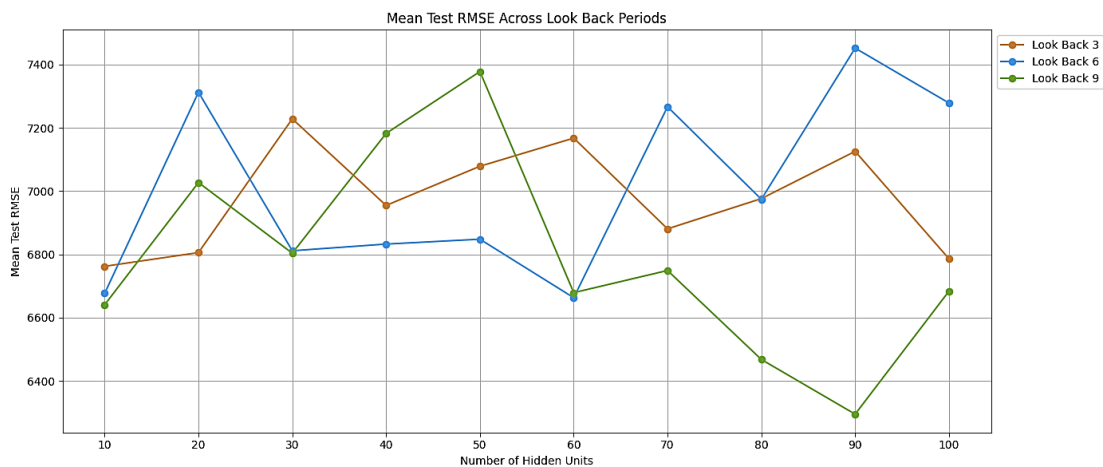
ตารางที่ 4 - 11 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,538	0.70	6,762
3	20	4,477	0.71	6,806
3	30	4,559	0.68	7,228
3	40	4,588	0.70	6,955
3	50	4,543	0.70	7,078
3	60	4,547	0.69	7,167
3	70	4,530	0.71	6,881
3	80	4,654	0.69	6,976
3	90	4,578	0.69	7,125
3	100	4,511	0.70	6,786
6	10	4,562	0.72	6,677
6	20	4,446	0.67	7,312
6	30	4,500	0.69	6,811
6	40	4,469	0.68	6,832
6	50	4,588	0.69	6,848
6	60	4,632	0.69	6,663
6	70	4,670	0.65	7,266
6	80	4,739	0.67	6,974

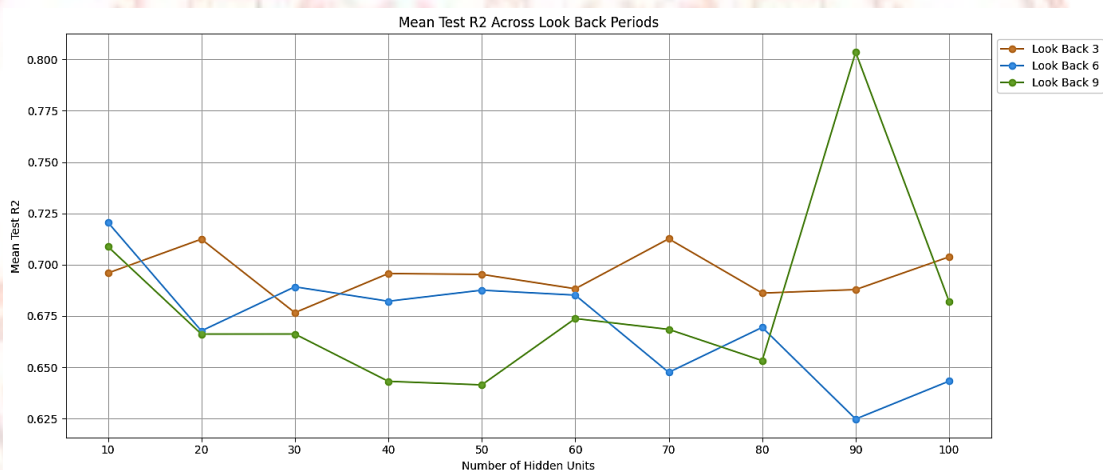
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
6	90	4,786	0.62	7,452
6	100	4,738	0.64	7,278
9	10	4,516	0.71	6,641
9	20	4,812	0.67	7,027
9	30	4,969	0.67	6,804
9	40	4,916	0.64	7,182
9	50	5,081	0.64	7,378
9	60	5,206	0.67	6,679
9	70	4,877	0.67	6,749
9	80	5,130	0.65	6,468
9	90	4,988	0.80	6,295
9	100	4,933	0.68	6,684



ภาพที่ 4 - 31 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 32 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 1)



ภาพที่ 4 - 33 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 1)

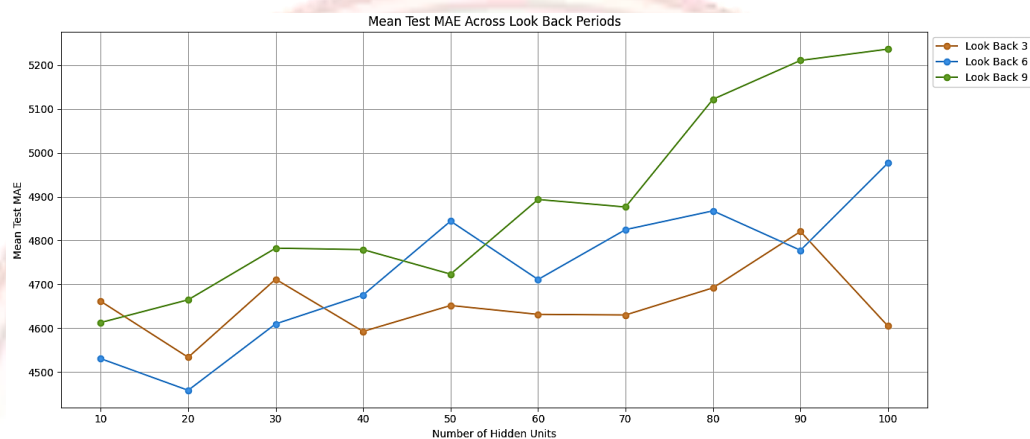
4.1.12 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน ระยะที่ 1 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 70 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 4,630, Mean_Test_RMSE = 6,277 และ Mean_Test_R² = 0.72 ซึ่งเป็นค่าที่ดีที่สุดในกลุ่มนี้ แสดงว่าแบบจำลองสามารถจับแนวโน้มของข้อมูลและให้ผลลัพธ์ที่ดีในเชิงปริมาณและเชิงแนวโน้ม แม้จะมีแบบจำลองบางชุด เช่น Look-back = 9 เดือน และ Hidden Units = 90 ที่ได้ R² สูงถึง 0.78 แต่ค่าความคลาดเคลื่อน MAE และ RMSE ยังสูงกว่าแบบจำลองที่เลือกอยู่มาก โดยมีรายละเอียดของผลการทดลอง ดังนี้

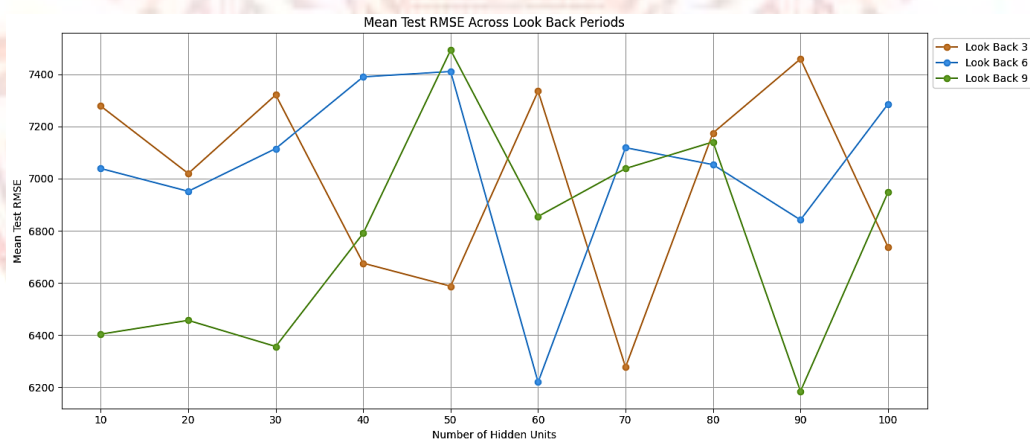
ตารางที่ 4 - 12 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 1)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,661	0.60	7,278
3	20	4,533	0.69	7,020
3	30	4,712	0.63	7,321
3	40	4,592	0.69	6,676
3	50	4,652	0.69	6,587
3	60	4,631	0.63	7,335
3	70	4,630	0.72	6,277
3	80	4,692	0.64	7,175
3	90	4,820	0.69	7,459
3	100	4,605	0.69	6,738
6	10	4,530	0.70	7,039
6	20	4,458	0.69	6,951
6	30	4,610	0.67	7,115
6	40	4,676	0.65	7,390
6	50	4,844	0.63	7,410
6	60	4,711	0.70	6,220
6	70	4,825	0.65	7,119
6	80	4,868	0.64	7,053
6	90	4,778	0.61	6,842
6	100	4,977	0.62	7,285
9	10	4,613	0.70	6,404
9	20	4,665	0.72	6,457
9	30	4,782	0.71	6,356
9	40	4,779	0.68	6,790
9	50	4,723	0.64	7,492
9	60	4,894	0.67	6,855
9	70	4,876	0.65	7,039

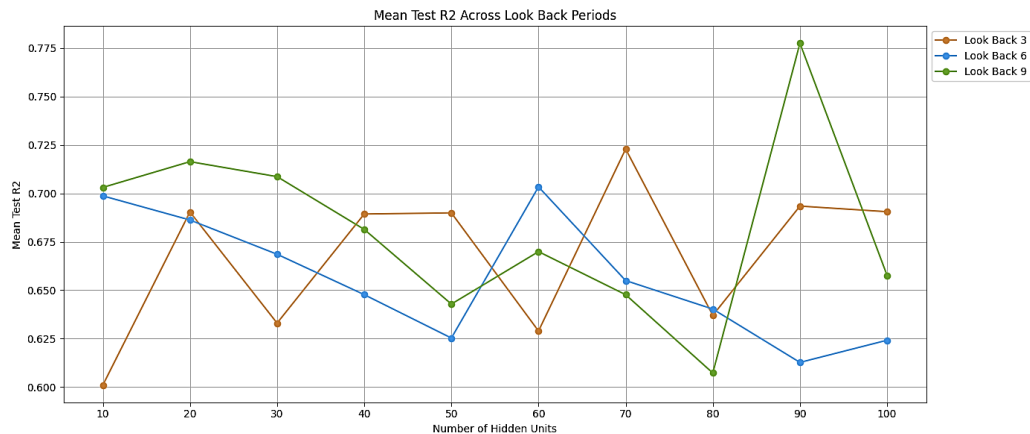
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	80	5,122	0.61	7,141
9	90	5,211	0.78	6,185
9	100	5,236	0.66	6,947



ภาพที่ 4 - 34 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 35 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 1)



ภาพที่ 4 - 36 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 1)

4.2 ผลการทดลอง ระยะที่ 2

การทดลองระยะที่ 2 (Phase 2) เป็นการทดลองต่อยอดจากระยะที่ 1 โดยเพิ่มฟีเจอร์สนับสนุนอื่น ๆ เข้ามาในแบบจำลอง ได้แก่ ข้อมูลเดือน, อุณหภูมิ, ปริมาณยานพาหนะ และข้อมูลเชิงพื้นที่ โดยยังคงทำการทดสอบด้วยค่าพารามิเตอร์ Look-back ที่ 3, 6 และ 9 เดือน และจำนวน Hidden Units ที่เพิ่มขึ้นทีละ 10 จนถึง 100 หน่วย เพื่อหาค่าที่ดีที่สุดสำหรับแต่ละแบบจำลอง ผลลัพธ์ในระยณะนี้เป็นการประเมินว่าการใช้ข้อมูลที่ซับซ้อนขึ้นจะช่วยเพิ่มประสิทธิภาพในการพยากรณ์ของแบบจำลองได้หรือไม่ โดยผลการทดลองทั้งหมดของระยะที่ 2 มีดังนี้

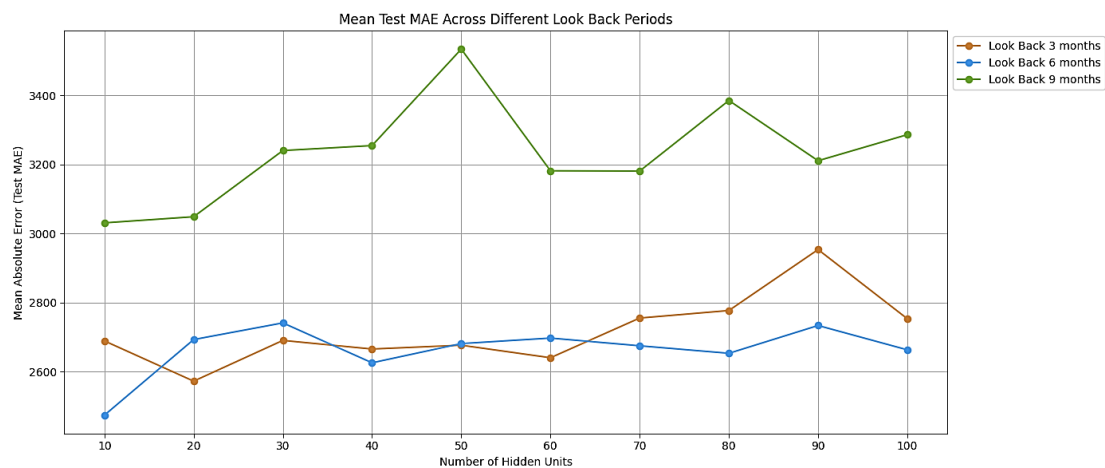
4.2.1 ผลการทดลองแบบจำลองรายอุทยานระยะที่ 2 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 6 เดือน และ Hidden Units = 10 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,474, Mean_Test_RMSE = 3,024 และ Mean_Test_ R^2 = -0.79 ซึ่งแม้ค่า R^2 จะยังไม่ดีนัก แต่เป็นค่าที่ดีที่สุดเมื่อเทียบกับแบบจำลองอื่นในกลุ่มเดียวกัน โดยเฉพาะ Look-back = 9 เดือน ที่ R^2 ติดลบอย่างรุนแรงทุกพารามิเตอร์ เช่น -13.41 ใน Hidden Units = 30 แสดงถึงข้อจำกัดของ RNN ในการจัดการกับข้อมูลลำดับที่ยาวและมีหลายฟีเจอร์ โดยมีรายละเอียดของผลการทดลอง ดังนี้

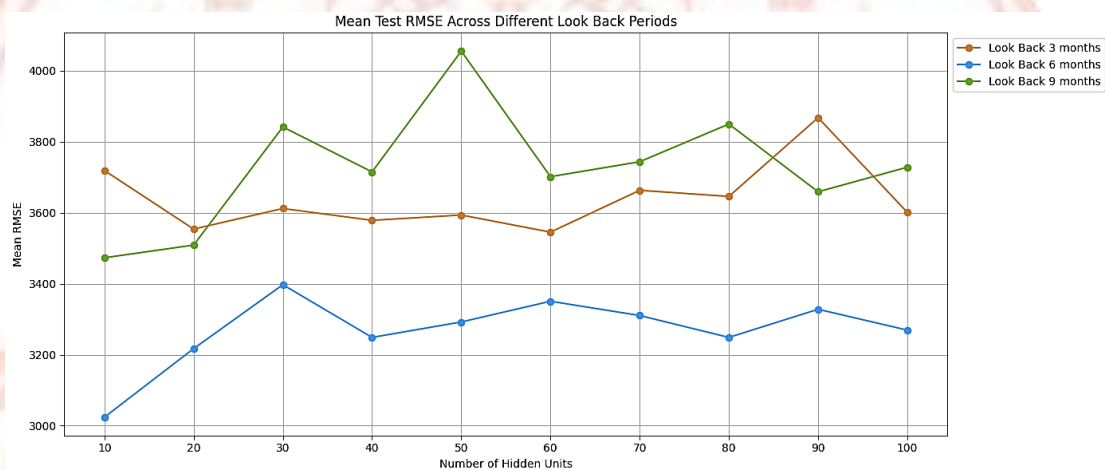
ตารางที่ 4 - 13 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ RNN (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_ R^2	Mean_Test_RMSE
3	10	2,689	-0.08	3,719
3	20	2,572	-0.24	3,554
3	30	2,690	-0.33	3,612
3	40	2,665	-0.20	3,579

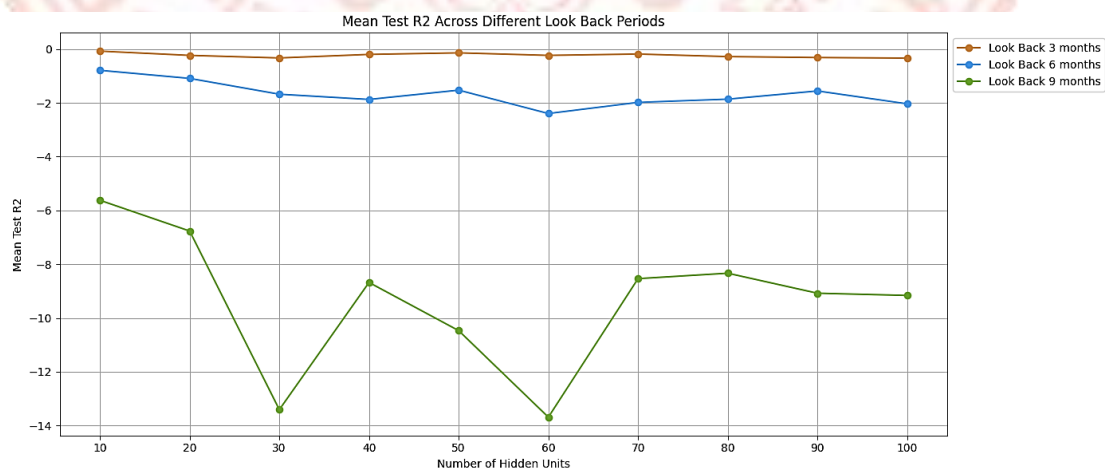
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	50	2,676	-0.14	3,594
3	60	2,640	-0.24	3,545
3	70	2,755	-0.19	3,663
3	80	2,777	-0.28	3,646
3	90	2,954	-0.32	3,868
3	100	2,753	-0.34	3,601
6	10	2,474	-0.79	3,024
6	20	2,693	-1.09	3,217
6	30	2,741	-1.68	3,397
6	40	2,625	-1.87	3,249
6	50	2,681	-1.53	3,292
6	60	2,697	-2.40	3,351
6	70	2,675	-1.98	3,310
6	80	2,653	-1.86	3,249
6	90	2,734	-1.56	3,328
6	100	2,663	-2.04	3,269
9	10	3,031	-5.62	3,473
9	20	3,049	-6.77	3,509
9	30	3,241	-13.41	3,842
9	40	3,255	-8.68	3,715
9	50	3,534	-10.47	4,055
9	60	3,182	-13.69	3,702
9	70	3,181	-8.54	3,744
9	80	3,386	-8.33	3,850
9	90	3,211	-9.08	3,659
9	100	3,287	-9.16	3,728



ภาพที่ 4 - 37 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 38 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 39 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ RNN (Phase 2)

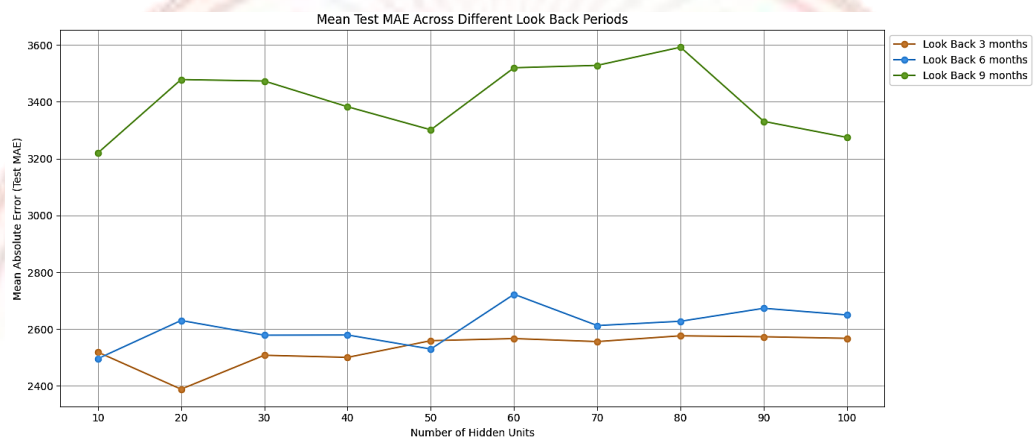
4.2.2 ผลการทดลองแบบจำลองรายอุทยานระยะที่ 2 โดยใช้แบบจำลอง LSTM

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 20 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,389, Mean_Test_RMSE = 3,275 และ Mean_Test_R² = 0.00 ซึ่งแม้ค่า R² จะไม่สูง แต่ดีกว่าทุกค่าที่เกิดจาก Look-back ที่นานขึ้น โดยเฉพาะ Look-back = 9 เดือน ซึ่งให้ค่า R² ต่ำสุดถึง -7.26 แสดงให้เห็นว่าแบบจำลองไม่สามารถสะท้อนแนวโน้มได้ดีเมื่อลำดับข้อมูลยาวเกินไป โดยมีรายละเอียดของผลการทดลอง ดังนี้

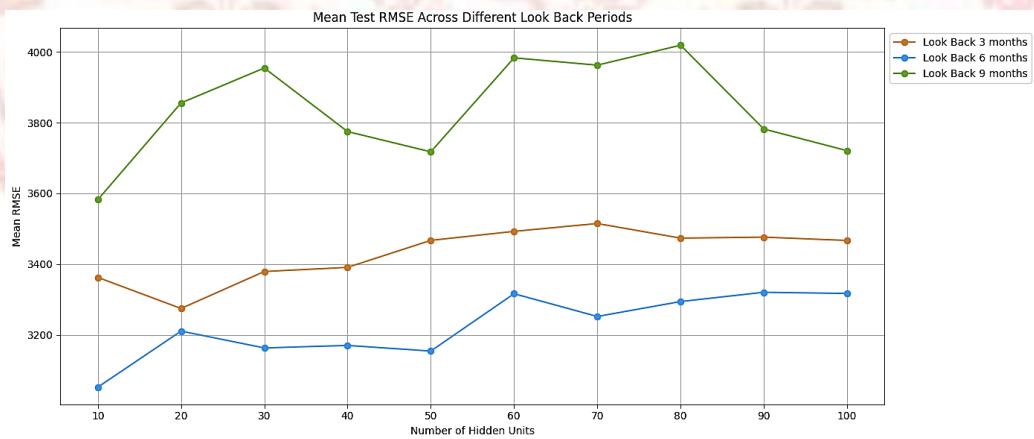
ตารางที่ 4 - 14 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	2,519	-0.06	3,363
3	20	2,389	0.00	3,275
3	30	2,508	-0.09	3,379
3	40	2,501	-0.08	3,391
3	50	2,559	-0.04	3,467
3	60	2,567	-0.11	3,493
3	70	2,556	-0.19	3,515
3	80	2,576	-0.12	3,474
3	90	2,573	-0.06	3,476
3	100	2,568	-0.06	3,467
6	10	2,496	-0.79	3,053
6	20	2,630	-1.04	3,211
6	30	2,579	-1.07	3,163
6	40	2,579	-0.95	3,170
6	50	2,530	-0.92	3,155
6	60	2,723	-0.92	3,317
6	70	2,612	-1.02	3,252
6	80	2,628	-0.98	3,294
6	90	2,673	-0.88	3,321
6	100	2,650	-0.98	3,317
9	10	3,220	-5.53	3,582
9	20	3,478	-4.95	3,856
9	30	3,473	-6.62	3,954
9	40	3,382	-4.64	3,775
9	50	3,301	-5.70	3,718

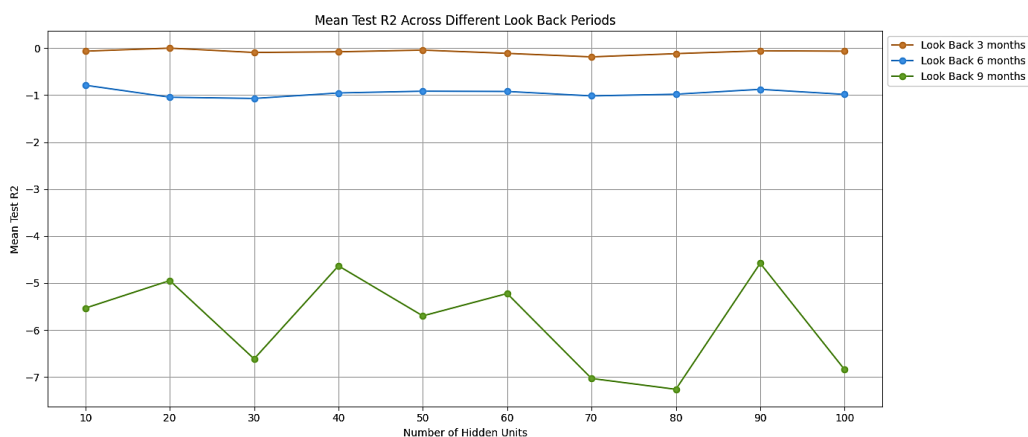
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	60	3,519	-5.22	3,983
9	70	3,528	-7.03	3,962
9	80	3,592	-7.26	4,019
9	90	3,331	-4.58	3,782
9	100	3,274	-6.84	3,721



ภาพที่ 4 - 40 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 41 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 42 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ LSTM (Phase 2)

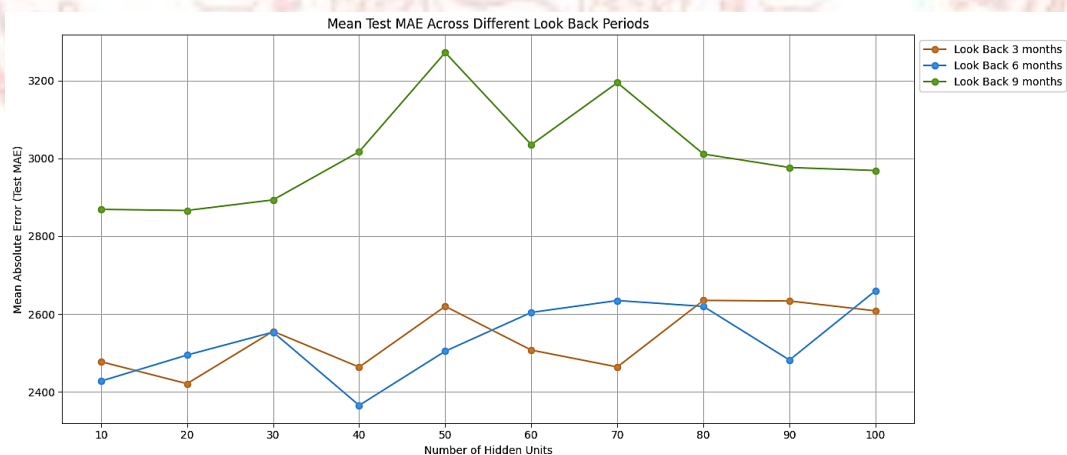
4.2.3 ผลการทดลองแบบจำลองรายอุทยานระยะที่ 2 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 3 เดือน และ Hidden Units = 20 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,421, Mean_Test_RMSE = 3,290 และ Mean_Test_ R^2 = 0.06 ซึ่งเป็นค่าที่สูงที่สุดในกลุ่มนี้ และดีกว่าทุกค่าที่ได้จาก Look-back ที่ยาวกว่า โดยเฉพาะ Look-back = 9 เดือนที่ค่า R^2 ตีลบสูงสุดถึง -6.92 แสดงให้เห็นว่าแบบจำลองไม่สามารถสะท้อนแนวโน้มของข้อมูลได้ดีเมื่อลำดับยาวเกินไป โดยมีรายละเอียดของผลการทดลอง ดังนี้

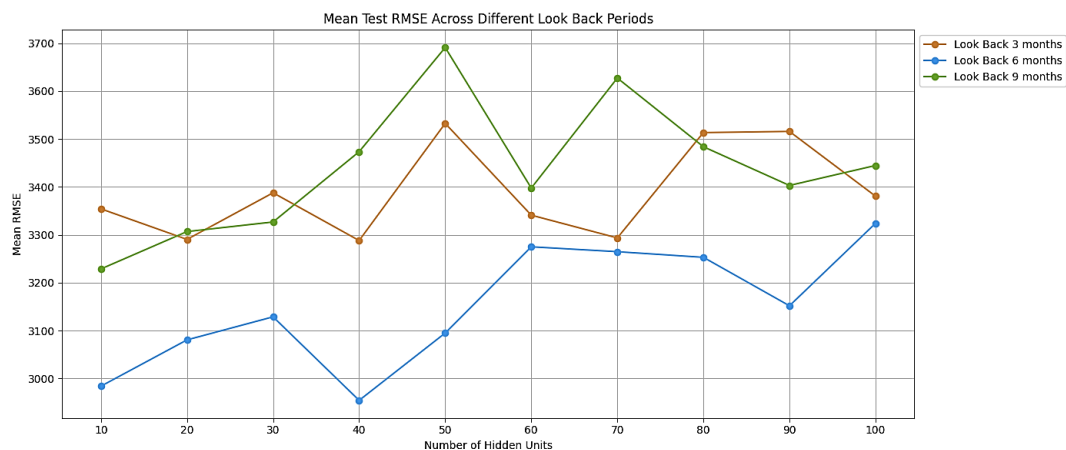
ตารางที่ 4 - 15 ผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_ R^2	Mean_Test_RMSE
3	10	2,478	-0.12	3,354
3	20	2,421	0.06	3,290
3	30	2,556	-0.05	3,388
3	40	2,464	-0.02	3,288
3	50	2,620	-0.07	3,533
3	60	2,508	-0.25	3,341
3	70	2,464	-0.04	3,293
3	80	2,635	-0.08	3,513
3	90	2,634	-0.01	3,516
3	100	2,608	-0.11	3,381
6	10	2,428	-0.70	2,984
6	20	2,495	-0.80	3,081
6	30	2,554	-0.83	3,129

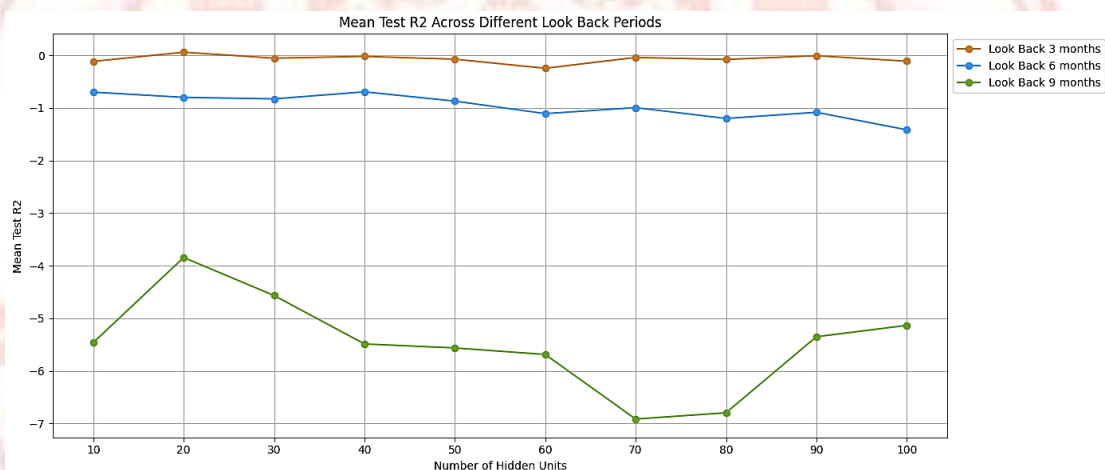
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
6	40	2,366	-0.69	2,954
6	50	2,505	-0.87	3,095
6	60	2,604	-1.11	3,275
6	70	2,635	-0.99	3,265
6	80	2,620	-1.20	3,253
6	90	2,482	-1.08	3,152
6	100	2,660	-1.42	3,324
9	10	2,869	-5.46	3,229
9	20	2,866	-3.85	3,307
9	30	2,893	-4.57	3,327
9	40	3,017	-5.49	3,473
9	50	3,272	-5.56	3,691
9	60	3,035	-5.69	3,398
9	70	3,194	-6.92	3,627
9	80	3,011	-6.80	3,484
9	90	2,976	-5.35	3,403
9	100	2,969	-5.14	3,445



ภาพที่ 4 - 43 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรายอุทยาน
โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 44 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 45 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรายอุทยาน โดยใช้ GRU (Phase 2)

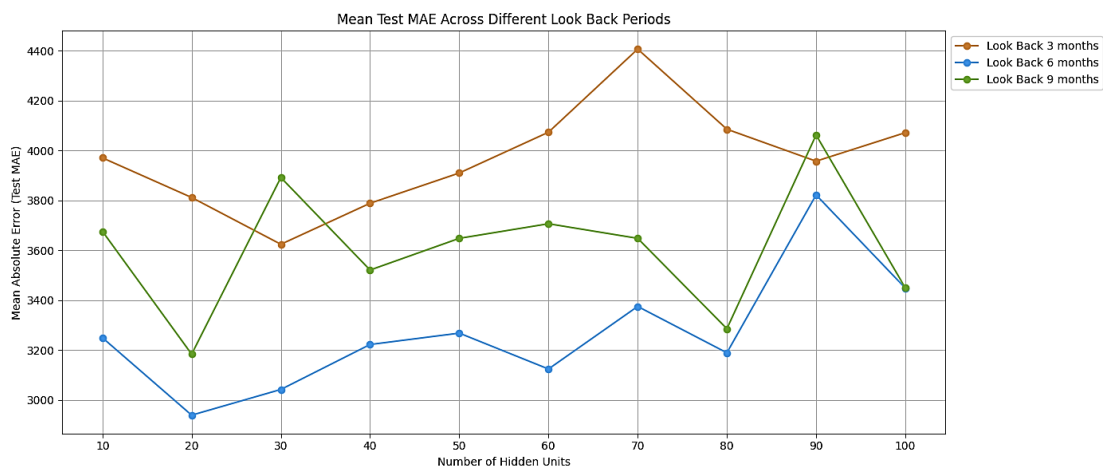
4.2.4 ผลการทดลองแบบจำลองรวมระยะที่ 2 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 6 เดือน และ Hidden Units = 50 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 3,268, Mean_Test_RMSE = 8,257 และ Mean_Test_R² = 0.79 ซึ่งเป็นค่าที่สูงที่สุดในชุดการทดลองนี้ และยังต่ำที่สุดในด้าน MAE และ RMSE เมื่อเทียบกับพารามิเตอร์อื่น ๆ ในกลุ่มเดียวกัน โดยมีรายละเอียดของผลการทดลอง ดังนี้

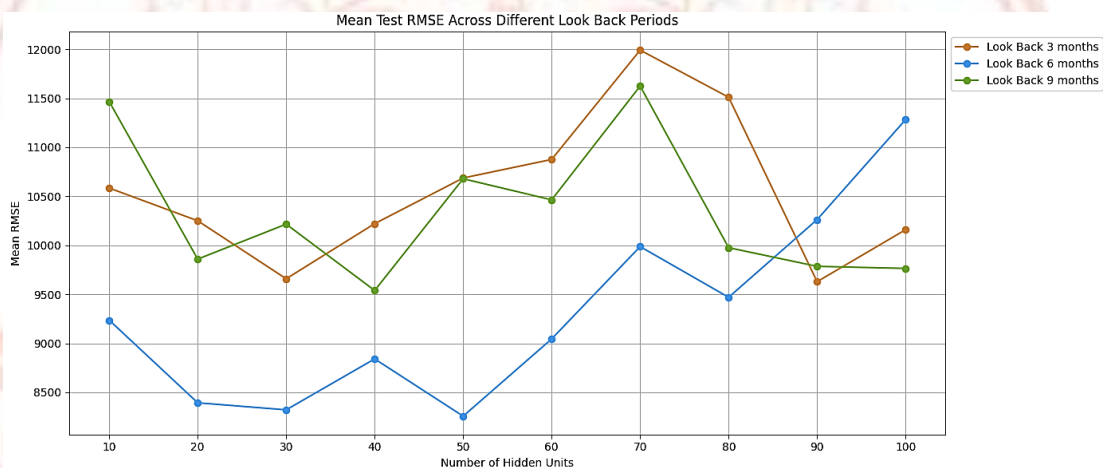
ตารางที่ 4 - 16 ผลการทดลองของแบบจำลองรวม โดยใช้ RNN (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	3,970	0.68	10,584
3	20	3,812	0.70	10,250

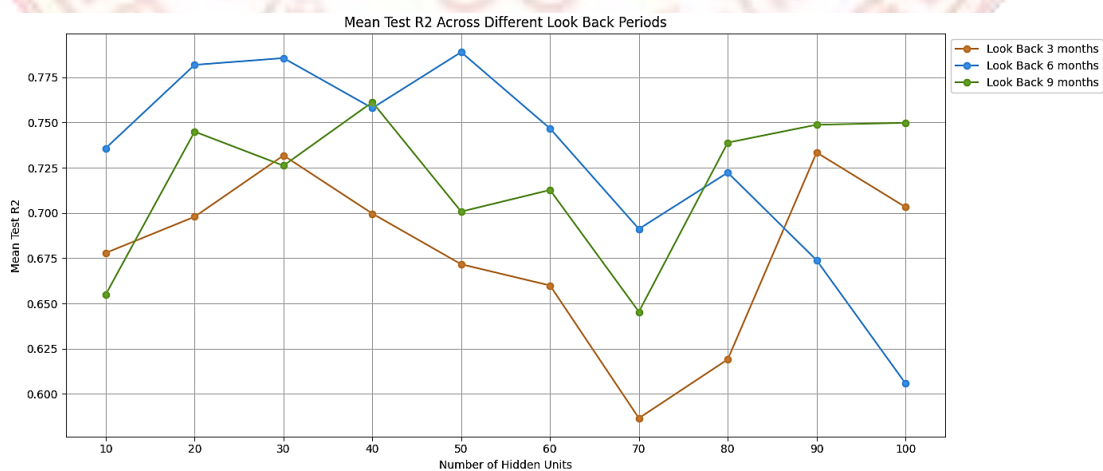
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	30	3,624	0.73	9,658
3	40	3,789	0.70	10,221
3	50	3,910	0.67	10,686
3	60	4,073	0.66	10,875
3	70	4,406	0.59	11,991
3	80	4,085	0.62	11,510
3	90	3,957	0.73	9,629
3	100	4,071	0.70	10,159
6	10	3,250	0.74	9,236
6	20	2,940	0.78	8,393
6	30	3,043	0.79	8,321
6	40	3,223	0.76	8,841
6	50	3,268	0.79	8,257
6	60	3,125	0.75	9,045
6	70	3,375	0.69	9,986
6	80	3,190	0.72	9,469
6	90	3,822	0.67	10,262
6	100	3,449	0.61	11,282
9	10	3,676	0.66	11,466
9	20	3,184	0.74	9,859
9	30	3,892	0.73	10,216
9	40	3,521	0.76	9,539
9	50	3,648	0.70	10,679
9	60	3,706	0.71	10,464
9	70	3,648	0.65	11,624
9	80	3,286	0.74	9,975
9	90	4,062	0.75	9,785
9	100	3,450	0.75	9,764



ภาพที่ 4 - 46 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม
โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 47 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม
โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 48 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรวม
โดยใช้ RNN (Phase 2)

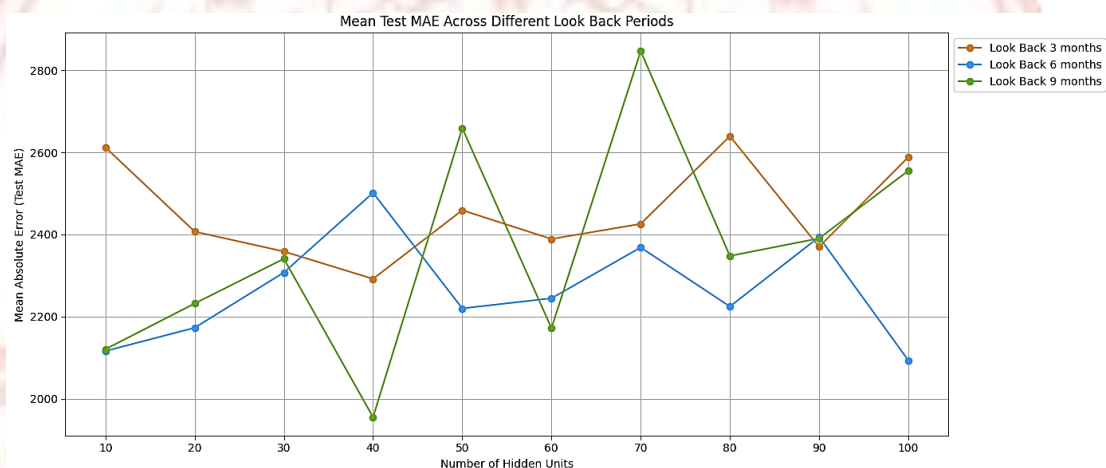
4.2.5 ผลการทดลองแบบจำลองรวมระยะที่ 2 โดยใช้แบบจำลอง LSTM

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 40 หน่วย ให้ผลลัพธ์ดีที่สุดในการทดลองชุดนี้ โดยมี Mean_Test_MAE = 1,955, Mean_Test_RMSE = 4,269 และ Mean_Test_R² = 0.95 ซึ่งเป็นค่าที่อยู่ในระดับดีที่สุดเมื่อเทียบกับพารามิเตอร์อื่นทั้งหมด แสดงว่าแบบจำลองสามารถจับลำดับข้อมูลและความเปลี่ยนแปลงได้ดีเมื่อใช้ข้อมูลย้อนหลังที่ยาวพอสมควร โดยเฉพาะเมื่อรวมกับพีเจอร์เสริม ทำให้เป็นตัวเลือกที่เหมาะสมที่สุดในชุดการทดลองนี้ โดยมีรายละเอียดของผลการทดลอง ดังนี้

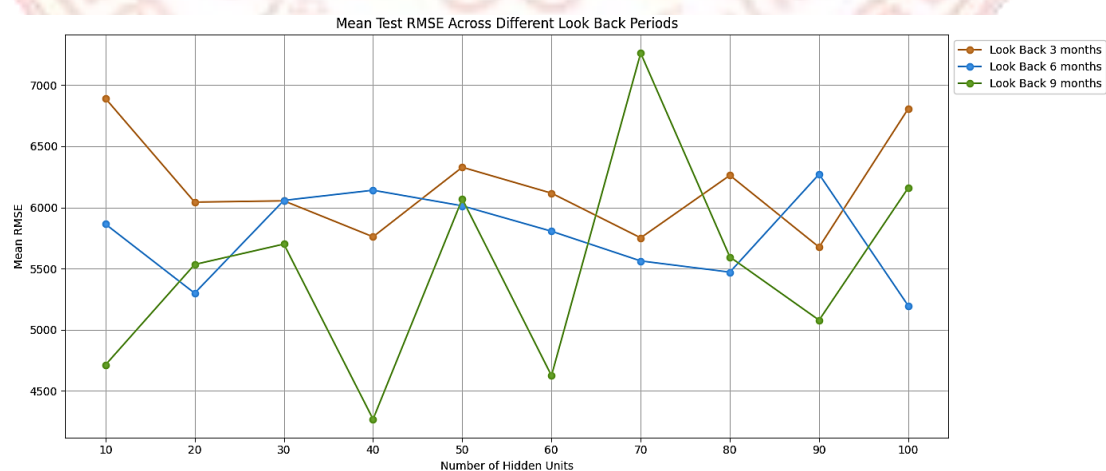
ตารางที่ 4 - 17 ผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	2,612	0.86	6,891
3	20	2,407	0.89	6,044
3	30	2,359	0.89	6,054
3	40	2,292	0.90	5,760
3	50	2,459	0.88	6,329
3	60	2,389	0.89	6,117
3	70	2,426	0.90	5,752
3	80	2,639	0.89	6,262
3	90	2,370	0.91	5,676
3	100	2,588	0.87	6,808
6	10	2,116	0.89	5,864
6	20	2,173	0.91	5,299
6	30	2,307	0.89	6,058
6	40	2,501	0.88	6,141
6	50	2,220	0.89	6,013
6	60	2,244	0.90	5,805
6	70	2,368	0.90	5,563
6	80	2,224	0.91	5,471
6	90	2,394	0.88	6,272
6	100	2,093	0.92	5,196
9	10	2,120	0.94	4,713

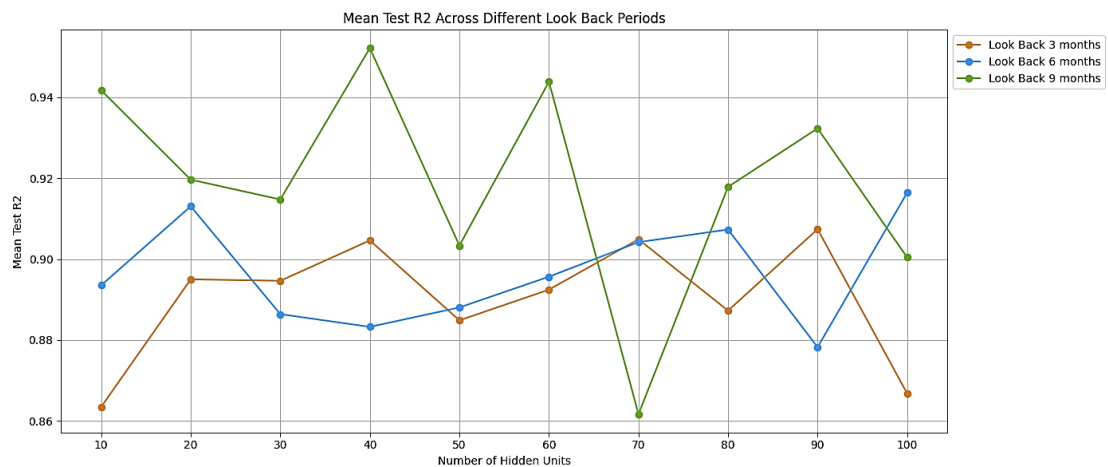
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	20	2,232	0.92	5,534
9	30	2,341	0.91	5,700
9	40	1,955	0.95	4,269
9	50	2,659	0.90	6,069
9	60	2,171	0.94	4,625
9	70	2,847	0.86	7,261
9	80	2,348	0.92	5,595
9	90	2,390	0.93	5,079
9	100	2,555	0.90	6,161



ภาพที่ 4 - 49 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม
โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 50 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม
โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 51 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองรวม โดยใช้ LSTM (Phase 2)

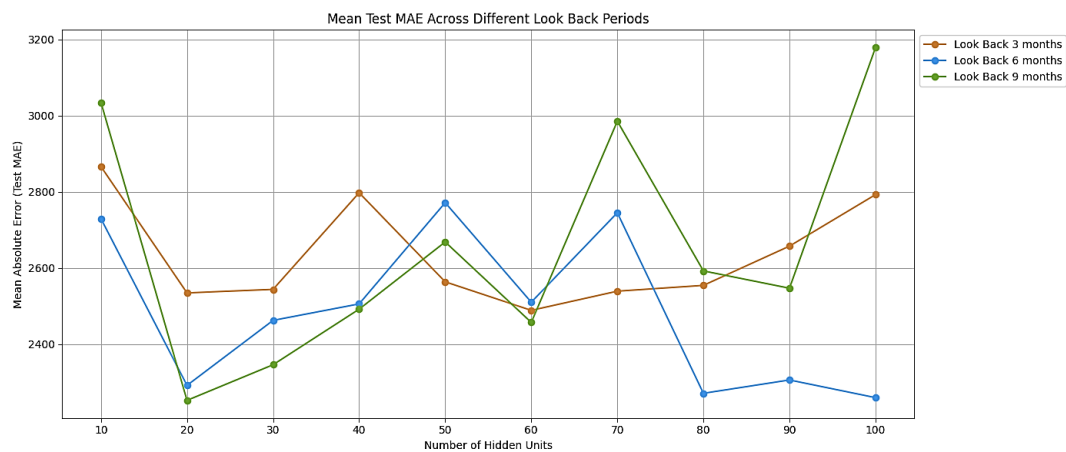
4.2.6 ผลการทดลองแบบจำลองรวมระยะที่ 2 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 60 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,457, Mean_Test_RMSE = 5,125 และ Mean_Test_ R^2 = 0.93 ซึ่งเป็นค่าที่สูงที่สุดในการทดลองชุดนี้ เมื่อเปรียบเทียบกับพารามิเตอร์ชุดอื่นจะเห็นว่าค่า MAE และ RMSE ต่ำกว่าในภาพรวม ขณะที่ Look-back ที่สั้นกว่านี้บางค่ามีค่า MAE ใกล้เคียง แต่ค่า R^2 ยังไม่สูงเท่า และเมื่อใช้ Look-back ที่ยาวเกินไปหรือ Hidden Units มากกว่า 70 หน่วย ผลลัพธ์เริ่มแยกลงในบางตัวชี้วัด เช่น MAE สูงขึ้นหรือ R^2 ลดลง โดยมีรายละเอียดของผลการทดลอง ดังนี้

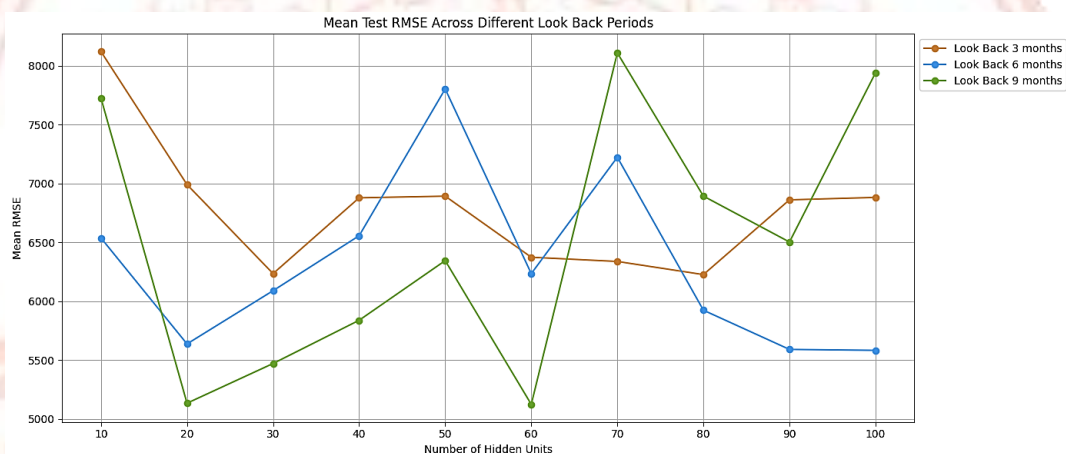
ตารางที่ 4 - 18 ผลการทดลองของแบบจำลองรวม โดยใช้ GRU (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_ R^2	Mean_Test_RMSE
3	10	2,867	0.81	8,119
3	20	2,535	0.86	6,990
3	30	2,544	0.89	6,235
3	40	2,798	0.86	6,878
3	50	2,564	0.86	6,892
3	60	2,489	0.88	6,374
3	70	2,539	0.88	6,338
3	80	2,555	0.89	6,226
3	90	2,658	0.86	6,861
3	100	2,793	0.86	6,881

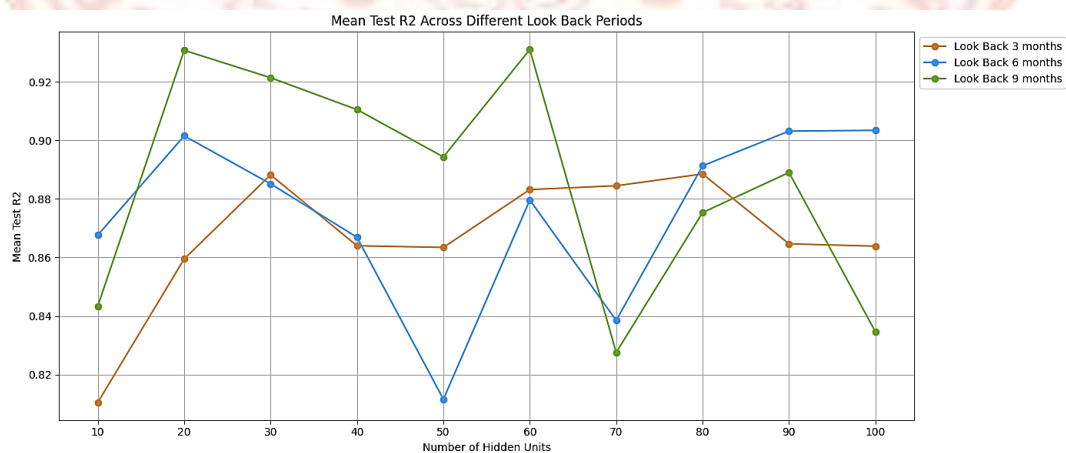
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
6	10	2,730	0.87	6,538
6	20	2,292	0.90	5,638
6	30	2,463	0.89	6,090
6	40	2,506	0.87	6,556
6	50	2,772	0.81	7,802
6	60	2,510	0.88	6,235
6	70	2,745	0.84	7,222
6	80	2,271	0.89	5,924
6	90	2,306	0.90	5,592
6	100	2,260	0.90	5,584
9	10	3,035	0.84	7,725
9	20	2,253	0.93	5,135
9	30	2,346	0.92	5,472
9	40	2,492	0.91	5,839
9	50	2,669	0.89	6,345
9	60	2,457	0.93	5,125
9	70	2,986	0.83	8,108
9	80	2,593	0.88	6,892
9	90	2,547	0.89	6,503
9	100	3,179	0.83	7,939



ภาพที่ 4 - 52 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองรวม
โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 53 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองรวม
โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 54 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองรวม
โดยใช้ GRU (Phase 2)

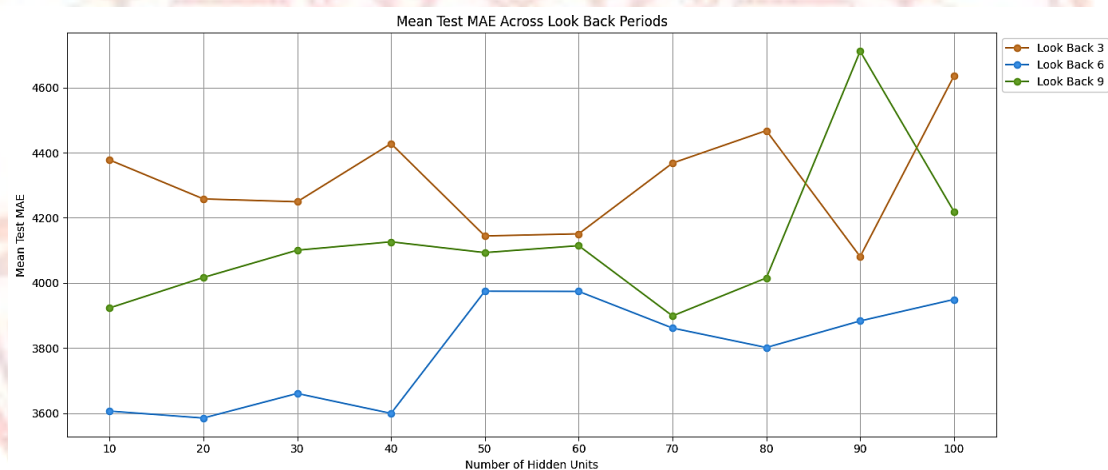
4.2.7 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามภูมิภาค ระยะที่ 2 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 6 เดือน และ Hidden Units = 10 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 3,606, Mean_Test_RMSE = 7,679 และ Mean_Test_R² = 0.76 ซึ่งดีกว่าพารามิเตอร์ชุดอื่นในการทดลองชุดนี้ เมื่อเปรียบเทียบกับ Look-back = 3 และ 9 เดือน จะเห็นว่าค่าชี้วัดโดยรวมในช่วง Look-back = 6 เดือน มีความสม่ำเสมอมากกว่า ทั้งในด้านค่า MAE และ RMSE ที่ไม่สูง และค่า R² ที่อยู่ในระดับดีในหลายค่าพารามิเตอร์ ขณะที่เมื่อ Hidden Units มากกว่า 60 หน่วย ผลลัพธ์เริ่มแย่ลง เช่น MAE เพิ่มขึ้น หรือ R² ลดลง โดยมีรายละเอียดของการทดลอง ดังนี้

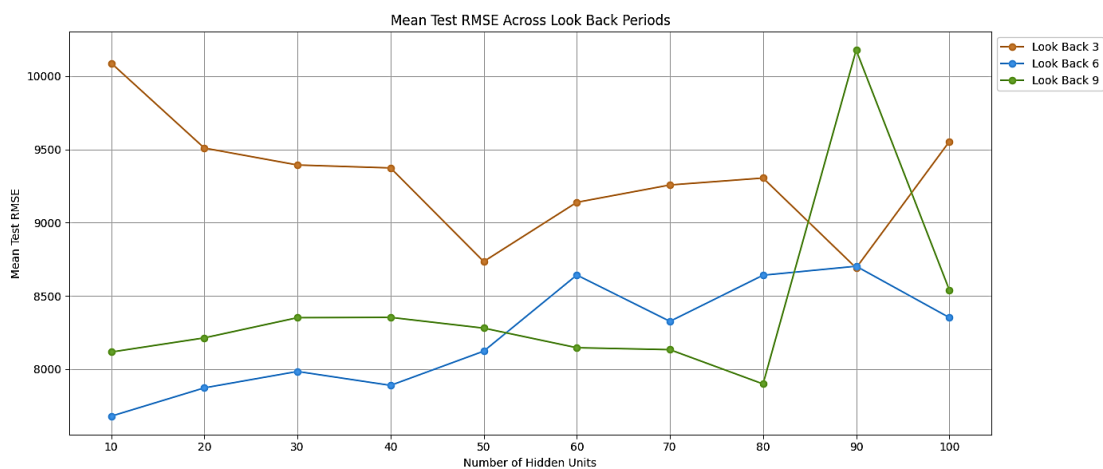
ตารางที่ 4 - 19 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	4,377	0.70	10,090
3	20	4,258	0.73	9,510
3	30	4,249	0.74	9,394
3	40	4,428	0.73	9,374
3	50	4,144	0.78	8,735
3	60	4,151	0.75	9,139
3	70	4,368	0.75	9,258
3	80	4,468	0.74	9,305
3	90	4,081	0.77	8,691
3	100	4,635	0.73	9,553
6	10	3,606	0.76	7,679
6	20	3,585	0.72	7,872
6	30	3,660	0.73	7,984
6	40	3,599	0.73	7,889
6	50	3,975	0.72	8,123
6	60	3,974	0.70	8,643
6	70	3,861	0.72	8,326
6	80	3,802	0.72	8,641
6	90	3,883	0.70	8,703
6	100	3,949	0.70	8,352

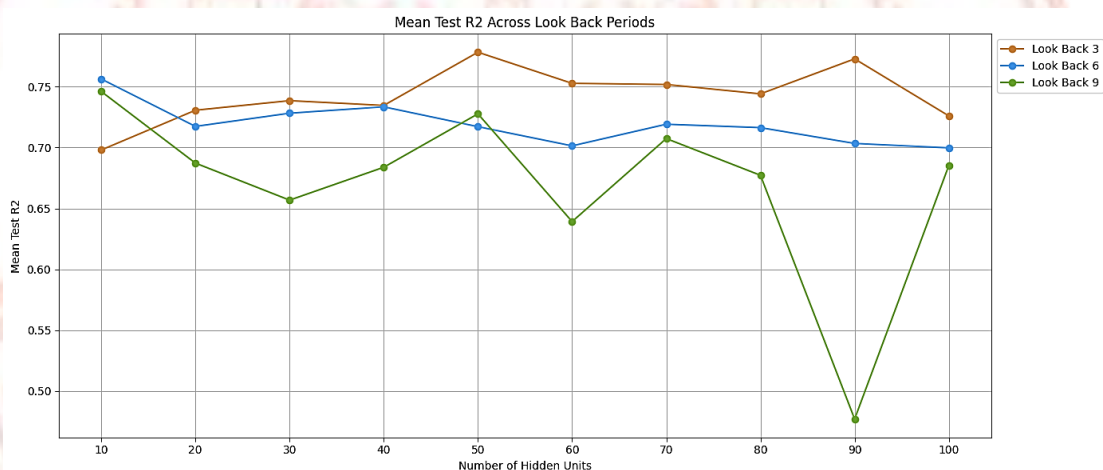
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	10	3,923	0.75	8,117
9	20	4,017	0.69	8,213
9	30	4,100	0.66	8,351
9	40	4,126	0.68	8,353
9	50	4,093	0.73	8,280
9	60	4,114	0.64	8,146
9	70	3,899	0.71	8,133
9	80	4,015	0.68	7,899
9	90	4,711	0.48	10,178
9	100	4,219	0.69	8,541



ภาพที่ 4 - 55 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 56 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 2)



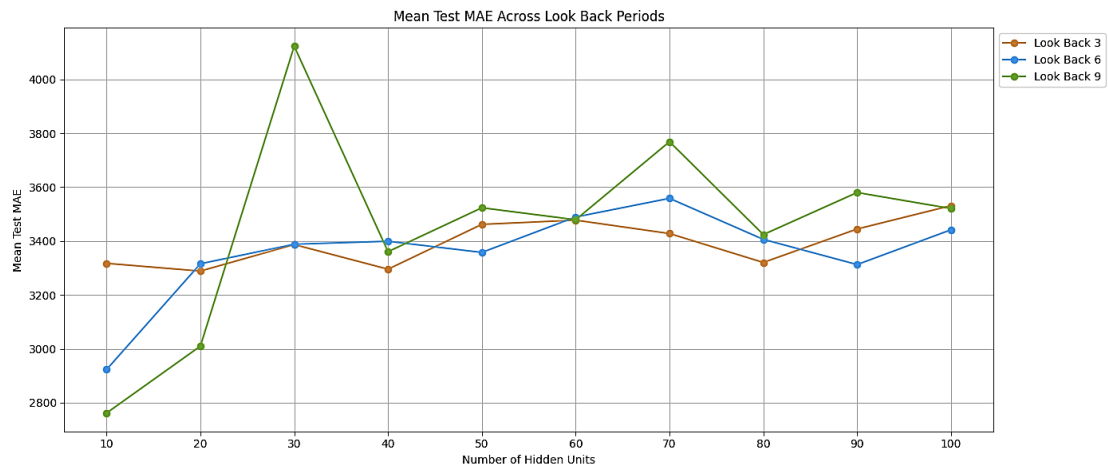
ภาพที่ 4 - 57 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ RNN (Phase 2)

4.2.8 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามภูมิภาค ระยะที่ 2 โดยใช้แบบจำลอง LSTM

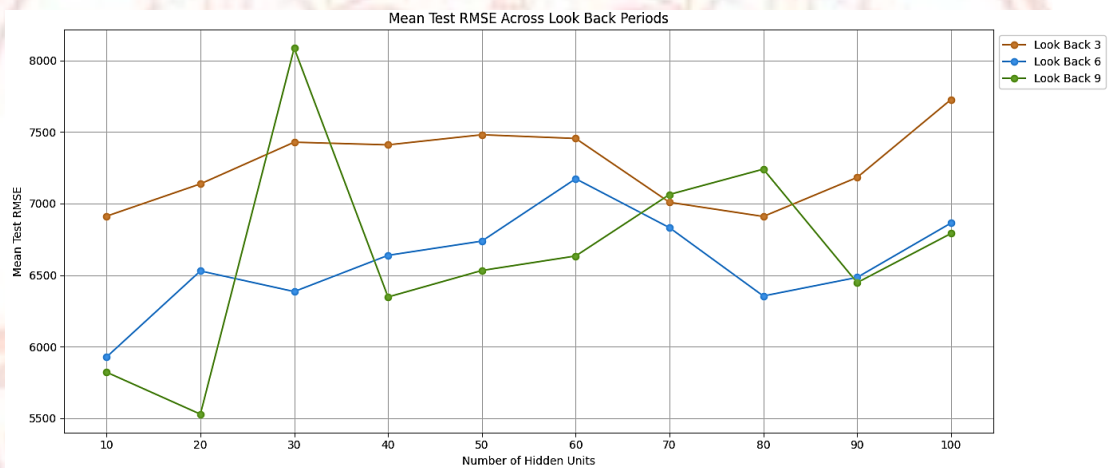
แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 10 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,762, Mean_Test_RMSE = 5,822 และ Mean_Test_R² = 0.85 ซึ่งแม้แบบจำลองอื่นที่ใช้ Look-back = 3 และ 6 เดือน จะให้ค่า R² ใกล้เคียงกัน แต่ยังมีค่า MAE และ RMSE สูงกว่าชัดเจน โดยมีรายละเอียดของผลการทดลอง ดังนี้

ตารางที่ 4 - 20 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 2)

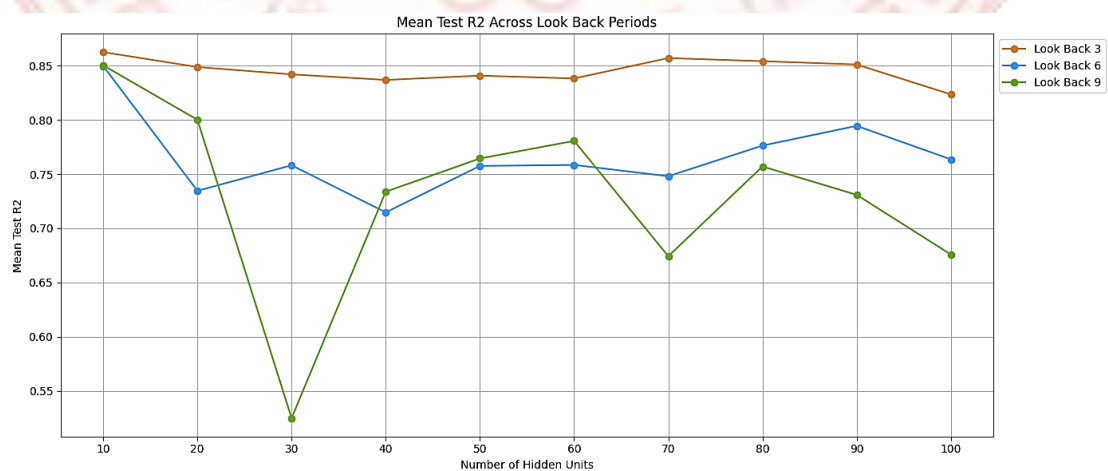
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	3,317	0.86	6,912
3	20	3,289	0.85	7,137
3	30	3,387	0.84	7,428
3	40	3,296	0.84	7,410
3	50	3,462	0.84	7,480
3	60	3,478	0.84	7,454
3	70	3,428	0.86	7,009
3	80	3,321	0.85	6,910
3	90	3,445	0.85	7,182
3	100	3,531	0.82	7,726
6	10	2,923	0.85	5,928
6	20	3,316	0.73	6,529
6	30	3,388	0.76	6,385
6	40	3,399	0.71	6,638
6	50	3,358	0.76	6,738
6	60	3,489	0.76	7,173
6	70	3,559	0.75	6,831
6	80	3,406	0.78	6,354
6	90	3,313	0.79	6,484
6	100	3,442	0.76	6,864
9	10	2,762	0.85	5,822
9	20	3,010	0.80	5,529
9	30	4,123	0.52	8,085
9	40	3,361	0.73	6,347
9	50	3,524	0.76	6,532
9	60	3,479	0.78	6,634
9	70	3,769	0.67	7,063
9	80	3,424	0.76	7,240
9	90	3,580	0.73	6,448
9	100	3,521	0.68	6,791



ภาพที่ 4 - 58 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 59 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 60 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ LSTM (Phase 2)

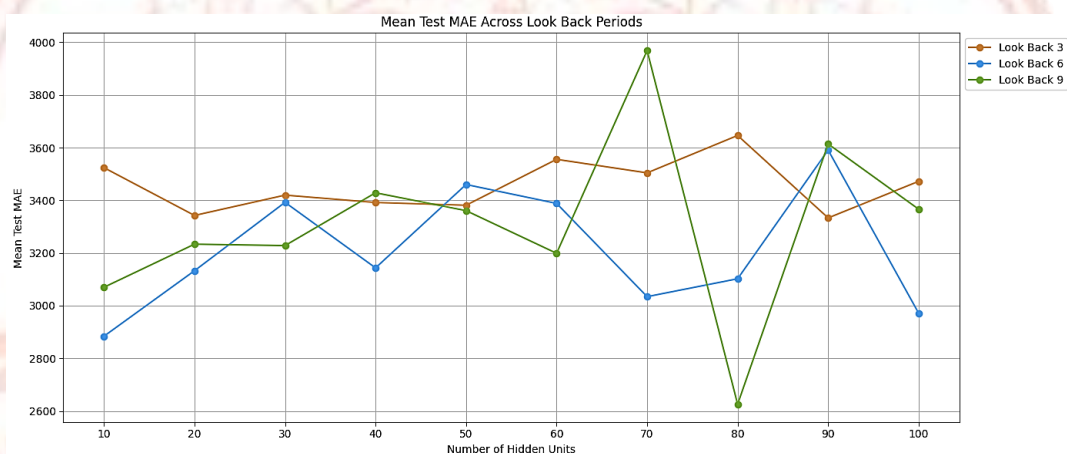
4.2.9 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามภูมิภาค ระยะที่ 2 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 80 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,626, Mean_Test_RMSE = 4,729 และ Mean_Test_R² = 0.87 ซึ่งแม้ MAE จะไม่ต่ำที่สุดในชุดนี้ แต่เมื่อพิจารณาพร้อมกับค่า RMSE และ R² แล้ว แบบจำลองนี้ยังแสดงผลลัพธ์โดยรวมที่ดีกว่ารูปแบบอื่น ขณะที่รูปแบบอื่น เช่น Hidden Units = 70 หรือ 90 หน่วย แม้จะอยู่ในช่วง Look-back เดียวกัน แต่กลับมี R² ต่ำกว่าอย่างชัดเจน โดยมีรายละเอียดของผลการทดลอง ดังนี้

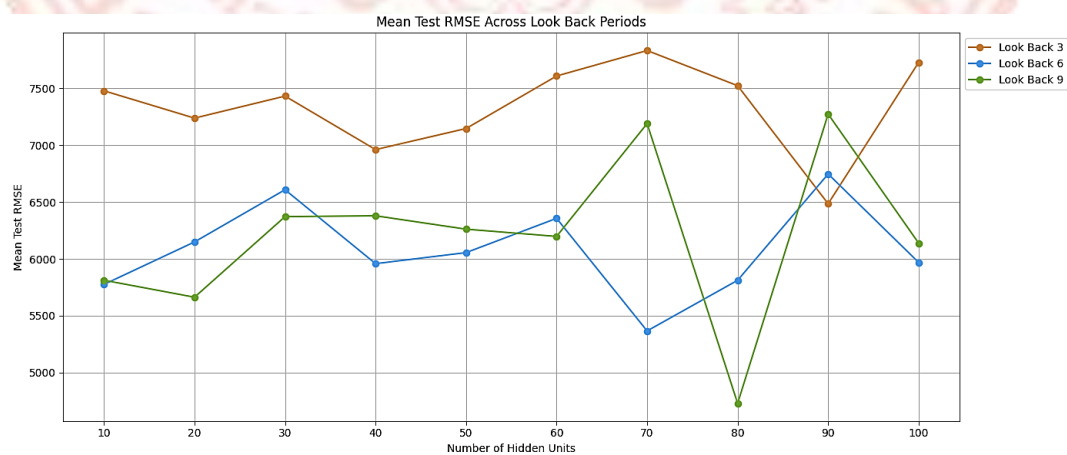
ตารางที่ 4 - 21 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	3,524	0.84	7,479
3	20	3,342	0.84	7,239
3	30	3,420	0.84	7,432
3	40	3,392	0.85	6,962
3	50	3,381	0.86	7,147
3	60	3,555	0.85	7,609
3	70	3,504	0.83	7,833
3	80	3,646	0.84	7,524
3	90	3,333	0.89	6,486
3	100	3,472	0.85	7,729
6	10	2,884	0.85	5,776
6	20	3,133	0.77	6,149
6	30	3,392	0.70	6,609
6	40	3,143	0.79	5,958
6	50	3,460	0.79	6,055
6	60	3,388	0.77	6,357
6	70	3,034	0.85	5,367
6	80	3,102	0.80	5,812
6	90	3,591	0.74	6,747
6	100	2,971	0.82	5,969
9	10	3,070	0.81	5,812

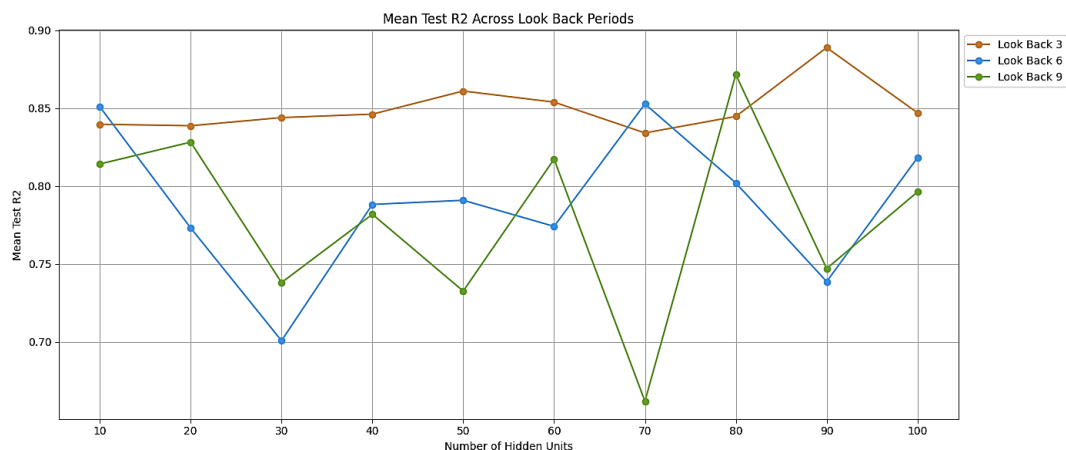
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	20	3,233	0.83	5,663
9	30	3,228	0.74	6,372
9	40	3,428	0.78	6,380
9	50	3,361	0.73	6,262
9	60	3,199	0.82	6,198
9	70	3,968	0.66	7,191
9	80	2,626	0.87	4,729
9	90	3,614	0.75	7,274
9	100	3,366	0.80	6,139



ภาพที่ 4 - 61 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 62 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 63 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามภูมิภาค โดยใช้ GRU (Phase 2)

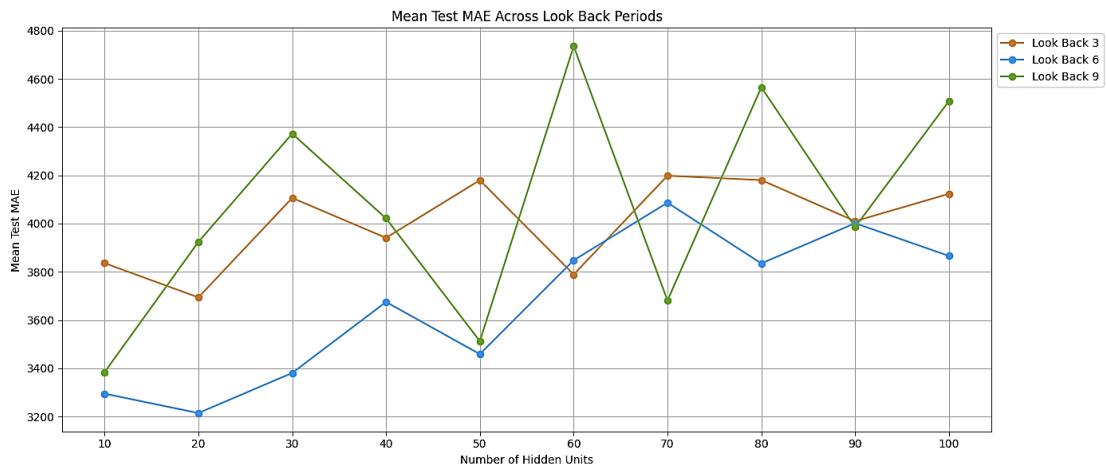
4.2.10 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน ระยะที่ 2 โดยใช้แบบจำลอง RNN

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 10 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 3,384, Mean_Test_RMSE = 7,467 และ Mean_Test_R² = 0.78 ซึ่งแม้ MAE จะไม่ต่ำที่สุด แต่เมื่อพิจารณาพร้อมกับ RMSE และ R² แล้ว แบบจำลองนี้ยังให้ผลลัพธ์ที่สมดุกว่ารูปแบบอื่น ๆ ในการทดลองชุดนี้ ขณะที่พารามิเตอร์อื่น เช่น Hidden Units = 20 ใน Look-back เดียวกัน แม้จะให้ค่า R² ใกล้เคียงกัน แต่ยังมีค่า RMSE สูงกว่า และเมื่อเพิ่ม Hidden Units เกิน 80 หน่วย ค่า R² เริ่มลดลงอย่างชัดเจน โดยมีรายละเอียดของผลการทดลอง ดังนี้

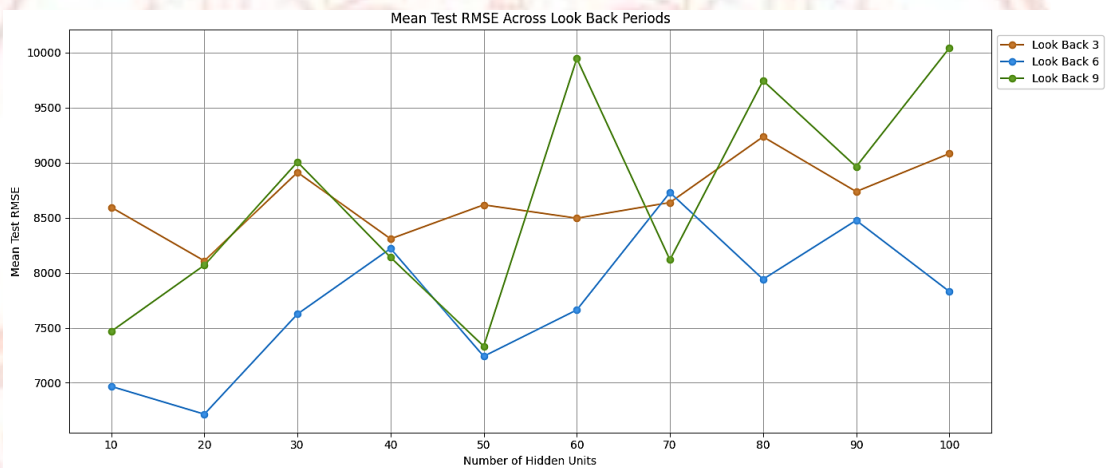
ตารางที่ 4 - 22 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	3,836	0.68	8,594
3	20	3,694	0.73	8,107
3	30	4,106	0.66	8,911
3	40	3,941	0.72	8,308
3	50	4,180	0.66	8,616
3	60	3,789	0.66	8,495
3	70	4,199	0.65	8,638
3	80	4,180	0.56	9,235
3	90	4,010	0.65	8,737

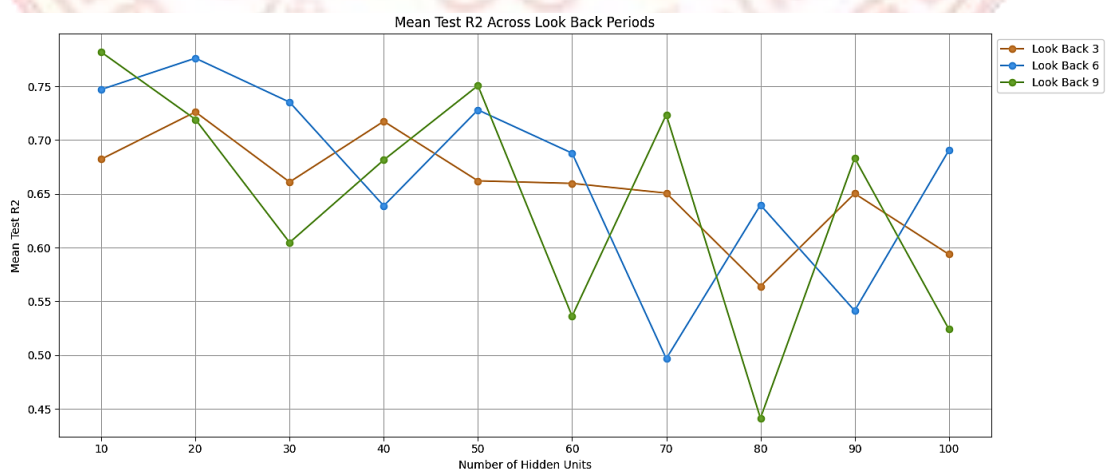
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	100	4,124	0.59	9,082
6	10	3,295	0.75	6,967
6	20	3,215	0.78	6,715
6	30	3,381	0.74	7,624
6	40	3,675	0.64	8,220
6	50	3,459	0.73	7,240
6	60	3,848	0.69	7,662
6	70	4,087	0.50	8,728
6	80	3,835	0.64	7,940
6	90	4,001	0.54	8,476
6	100	3,866	0.69	7,829
9	10	3,384	0.78	7,467
9	20	3,924	0.72	8,068
9	30	4,373	0.60	9,005
9	40	4,022	0.68	8,139
9	50	3,512	0.75	7,332
9	60	4,735	0.54	9,945
9	70	3,681	0.72	8,117
9	80	4,565	0.44	9,745
9	90	3,985	0.68	8,963
9	100	4,508	0.52	10,040



ภาพที่ 4 - 64 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 65 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 2)



ภาพที่ 4 - 66 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ RNN (Phase 2)

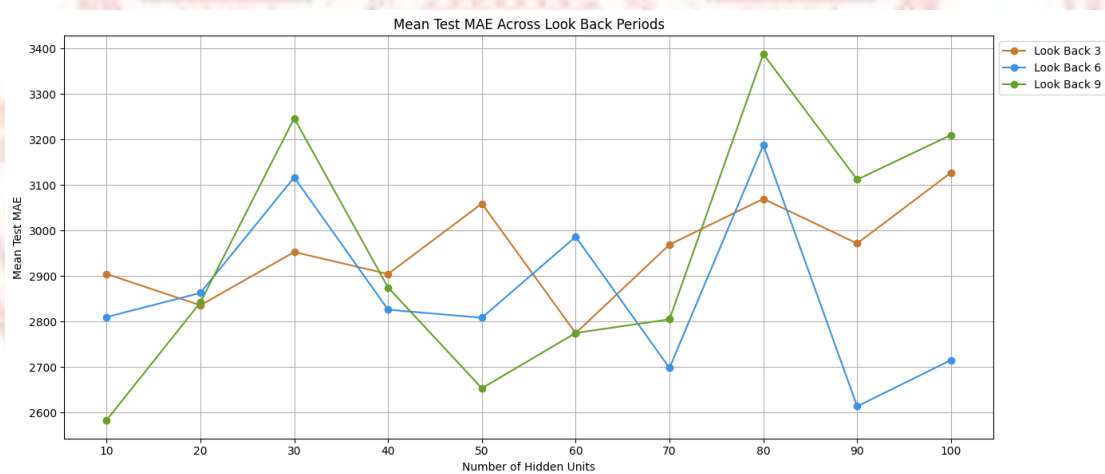
4.2.11 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน ระยะที่ 2 โดยใช้แบบจำลอง LSTM

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 50 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,654, Mean_Test_RMSE = 4,478 และ Mean_Test_R² = 0.89 ซึ่งถือว่ามีความสมมูลระหว่างค่าความคลาดเคลื่อนเชิงตัวเลขและความสอดคล้องกับแนวโน้มข้อมูลมากกว่ารูปแบบอื่น ๆ ในการทดลองชุดนี้ ขณะที่พารามิเตอร์อื่น เช่น Look-back = 6 เดือน, Hidden Units = 90 แม้ให้ MAE ต่ำกว่า แต่ R² ยังไม่สูงเท่าชุดที่เลือก และรูปแบบที่ใช้ Hidden Units มากเกินไป เช่น 100 หน่วย มักมีค่า MAE สูงขึ้นชัดเจน แสดงว่าการใช้ Look-back ที่ยาวควบคู่กับจำนวน Hidden Units ระดับกลางช่วยให้แบบจำลองเรียนรู้ได้ดี โดยมีรายละเอียดของผลการทดลอง ดังนี้

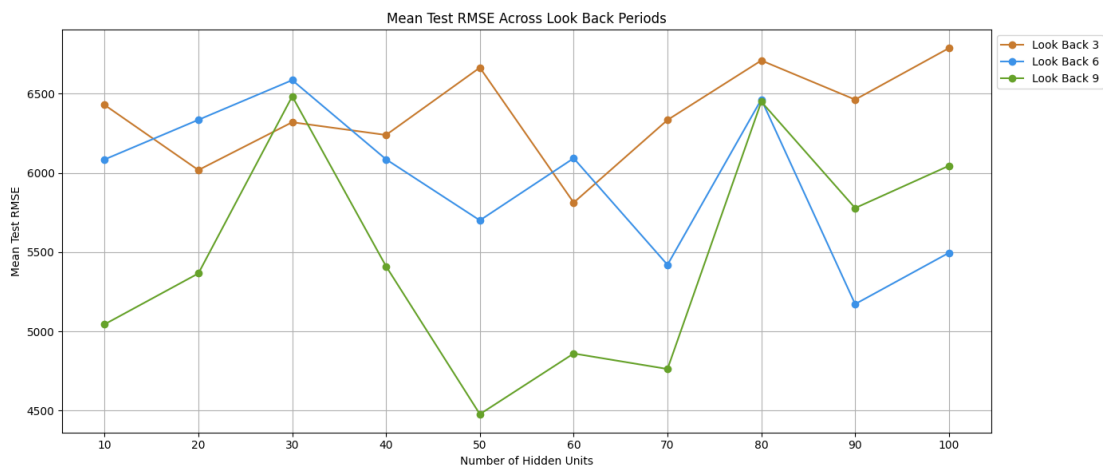
ตารางที่ 4 - 23 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	2,905	0.83	6,429
3	20	2,836	0.85	6,018
3	30	2,953	0.82	6,320
3	40	2,905	0.84	6,240
3	50	3,059	0.82	6,665
3	60	2,775	0.85	5,813
3	70	2,969	0.83	6,335
3	80	3,070	0.82	6,710
3	90	2,972	0.83	6,463
3	100	3,127	0.80	6,788
6	10	2,810	0.80	6,085
6	20	2,863	0.80	6,336
6	30	3,117	0.78	6,586
6	40	2,826	0.81	6,086
6	50	2,809	0.85	5,701
6	60	2,987	0.81	6,093
6	70	2,698	0.84	5,420
6	80	3,188	0.77	6,464

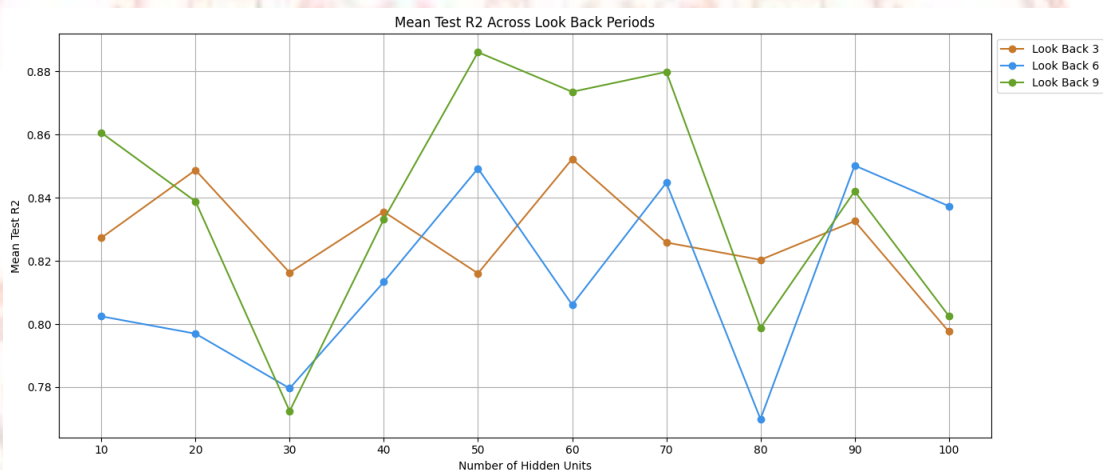
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
6	90	2,614	0.85	5,173
6	100	2,716	0.84	5,496
9	10	2,584	0.86	5,045
9	20	2,842	0.84	5,366
9	30	3,247	0.77	6,484
9	40	2,875	0.83	5,411
9	50	2,654	0.89	4,478
9	60	2,775	0.87	4,861
9	70	2,805	0.88	4,763
9	80	3,388	0.80	6,450
9	90	3,112	0.84	5,779
9	100	3,210	0.80	6,045



ภาพที่ 4 - 67 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 68 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 2)



ภาพที่ 4 - 69 กราฟแสดงค่าเฉลี่ยของ R² จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ LSTM (Phase 2)

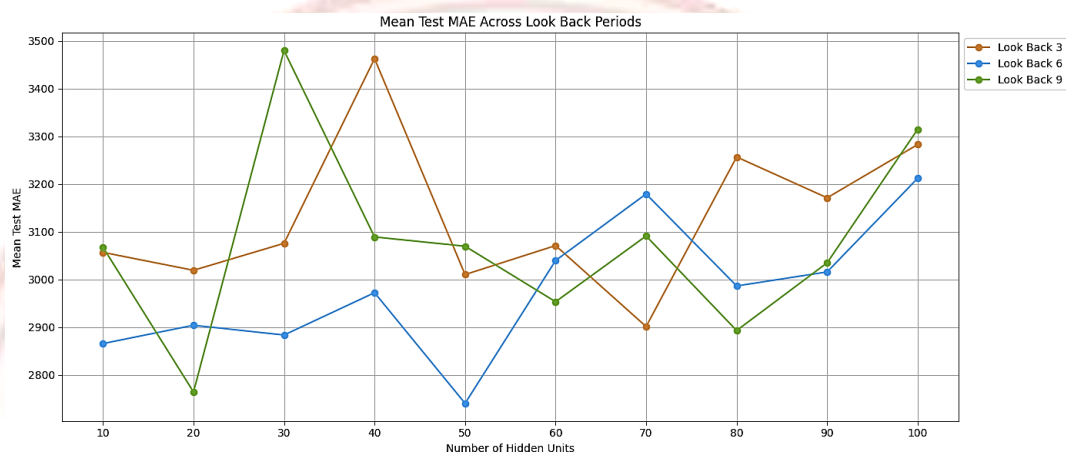
4.2.12 ผลการทดลองแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน ระยะที่ 2 โดยใช้แบบจำลอง GRU

แบบจำลองที่มี Look-back = 9 เดือน และ Hidden Units = 80 หน่วย ให้ผลลัพธ์ที่ดีที่สุด โดยมี Mean_Test_MAE = 2,893, Mean_Test_RMSE = 5,437 และ Mean_Test_R² = 0.84 ซึ่งถือว่าสมมูลทั้งในด้านค่าคลาดเคลื่อนเชิงปริมาณและความสอดคล้องกับแนวโน้มข้อมูลในการทดลองชุดนี้ ขณะที่พารามิเตอร์อื่น เช่น Look-back = 6 เดือน, Hidden Units = 50 แม้ให้ MAE ต่ำกว่าที่ 2,740 แต่ R² ยังไม่สูงเท่ารูปแบบที่เลือก โดยมีรายละเอียดของผลการทดลอง ดังนี้

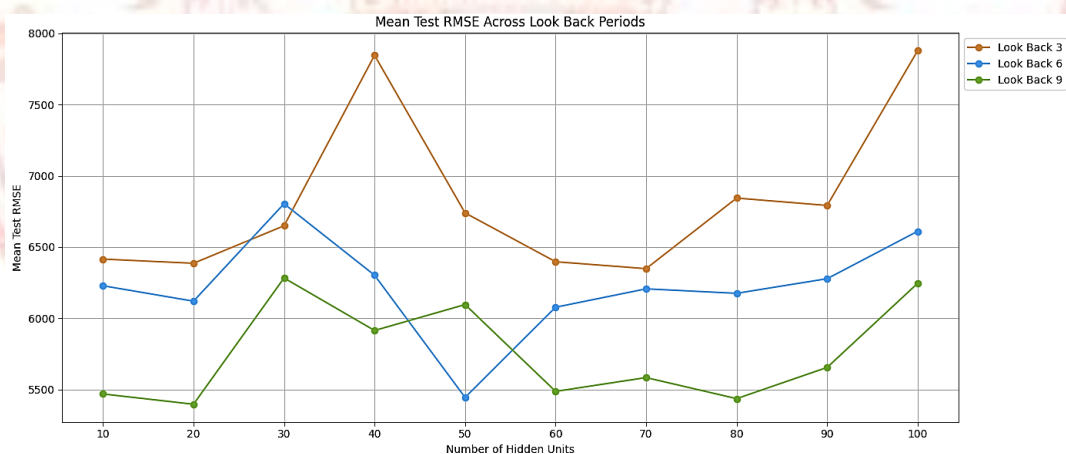
ตารางที่ 4 - 24 ผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 2)

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
3	10	3,057	0.81	6,415
3	20	3,019	0.75	6,386
3	30	3,076	0.76	6,649
3	40	3,463	0.45	7,845
3	50	3,010	0.77	6,739
3	60	3,071	0.80	6,397
3	70	2,901	0.83	6,348
3	80	3,257	0.76	6,843
3	90	3,171	0.78	6,792
3	100	3,283	0.63	7,878
6	10	2,866	0.80	6,228
6	20	2,904	0.80	6,120
6	30	2,884	0.77	6,804
6	40	2,972	0.75	6,302
6	50	2,740	0.83	5,445
6	60	3,040	0.72	6,078
6	70	3,179	0.76	6,207
6	80	2,986	0.72	6,175
6	90	3,016	0.76	6,278
6	100	3,212	0.72	6,611
9	10	3,067	0.82	5,469
9	20	2,764	0.83	5,397
9	30	3,480	0.71	6,283
9	40	3,089	0.80	5,915
9	50	3,070	0.80	6,096
9	60	2,953	0.83	5,488
9	70	3,091	0.83	5,584

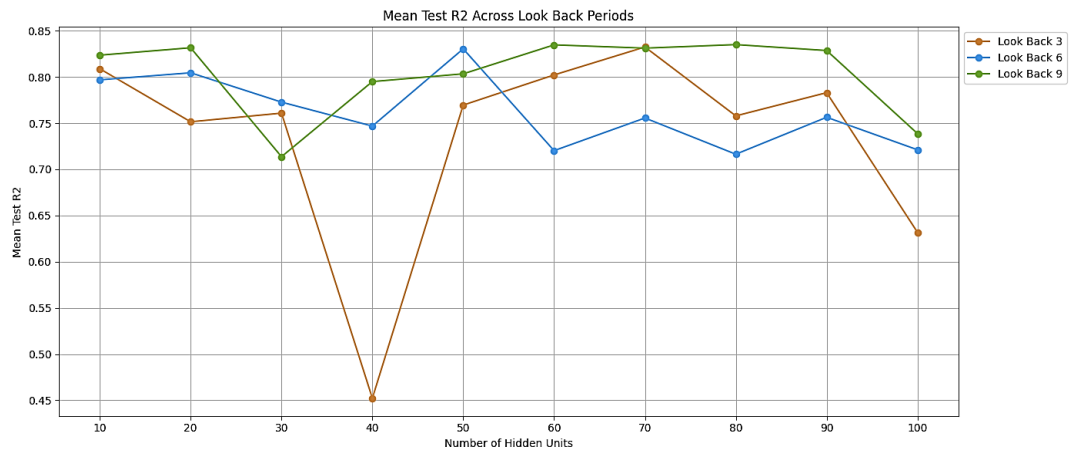
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE
9	80	2,893	0.84	5,437
9	90	3,035	0.83	5,656
9	100	3,315	0.74	6,246



ภาพที่ 4 - 70 กราฟแสดงค่าเฉลี่ยของ MAE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 71 กราฟแสดงค่าเฉลี่ยของ RMSE จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้ GRU (Phase 2)



ภาพที่ 4 - 72 กราฟแสดงค่าเฉลี่ยของ R^2 จากผลการทดลองของแบบจำลองแบบจัดกลุ่มตามลักษณะเด่นของแต่ละอุทยาน โดยใช้แบบจำลอง GRU (Phase 2)

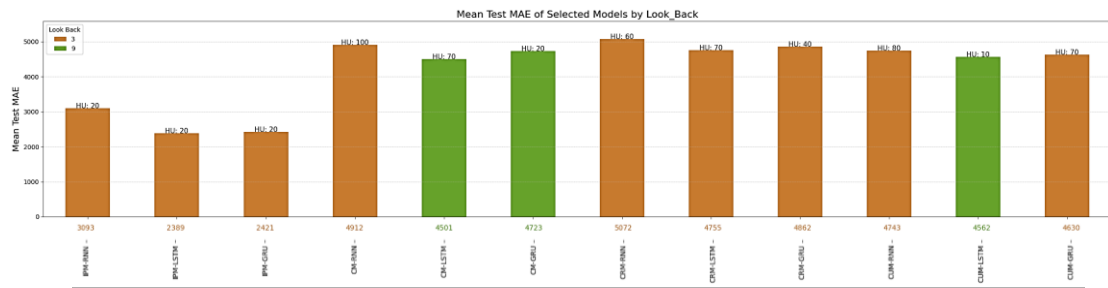
4.3 สรุปผลการทดลองระหว่าง ระยะที่ 1 และ ระยะที่ 2

4.3.1 สรุปผลการทดลองระยะที่ 1

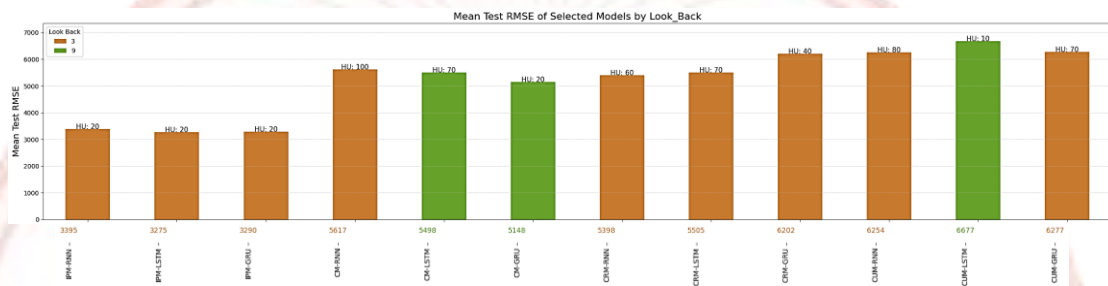
จากทุกแนวทางการทดลองระยะที่ 1 โดยใช้เพียงข้อมูลจำนวนนักท่องเที่ยว (Visitor Count) เป็นฟีเจอร์เดียว ผลสรุปการทดลองแสดงไว้ใน ตารางที่ 4 - 25 และ ภาพที่ 4 - 73 ถึง ภาพที่ 4 - 75 ดังนี้

ตารางที่ 4 - 25 สรุปผลการทดลองของระยะที่ 1

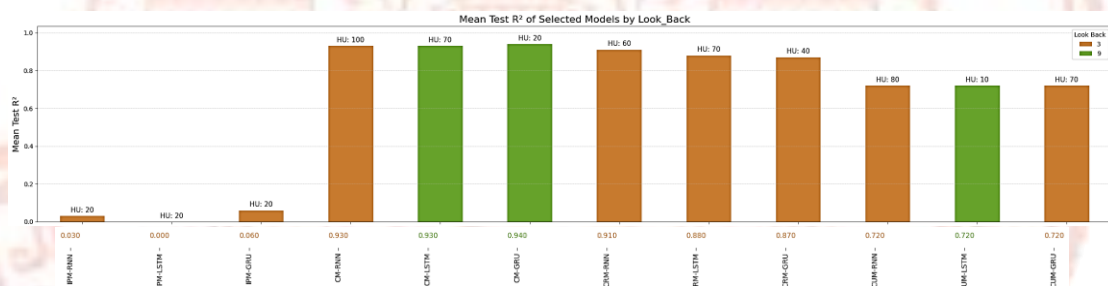
Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE	Name_Model
3	20	3,093	0.03	3,395	IPM-RNN
3	20	2,389	0.00	3,275	IPM-LSTM
3	20	2,421	0.06	3,290	IPM-GRU
3	100	4,912	0.93	5,617	CM-RNN
9	70	4,501	0.93	5,498	CM-LSTM
9	20	4,723	0.94	5,148	CM-GRU
3	60	5,072	0.91	5,398	CRM-RNN
3	70	4,755	0.88	5,505	CRM-LSTM
3	40	4,862	0.87	6,202	CRM-GRU
3	80	4,743	0.72	6,254	CUM-RNN
9	10	4,562	0.72	6,677	CUM-LSTM
3	70	4,630	0.72	6,277	CUM-GRU



ภาพที่ 4 - 73 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ MAE จากการสรุปผลการทดลองของระยะที่ 1



ภาพที่ 4 - 74 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ RMSE จากการสรุปผลการทดลองของระยะที่ 1



ภาพที่ 4 - 75 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ R² จากการสรุปผลการทดลองของระยะที่ 1

จากตารางที่ 4 - 25 ประกอบกับ ภาพที่ 4 - 73, ภาพที่ 4 - 74, ภาพที่ 4 - 75 พบว่าไม่มีค่า Look-back = 6 ปรากฏอยู่ในผลลัพธ์ของแบบจำลองที่ให้ผลดีที่สุด แม้ว่าการทดลองจะครอบคลุมค่า Look-back ที่ 3, 6 และ 9 เดือนก็ตาม ทั้งนี้อาจเป็นเพราะ Look-back = 6 เดือนเป็นช่วงข้อมูลที่อยู่กึ่งกลาง ไม่สามารถแสดงจุดเด่นชัดเจนในการเรียนรู้ข้อมูลได้ โดยอาจยาวเกินไปสำหรับ RNN ที่เหมาะกับลำดับข้อมูลสั้น และสั้นเกินไปสำหรับ LSTM หรือ GRU ที่ทำงานได้ดีในลำดับที่ยาวกว่า ส่งผลให้ค่า Look-back ดังกล่าวไม่สามารถให้ผลลัพธ์ที่ดีที่สุดในแต่ละชุดการทดลองได้ ในการเปรียบเทียบผลลัพธ์ของแต่ละแนวทาง พบว่า

(ก) แบบจำลองรวม (CM) แบบจำลองกลุ่มนี้ให้ค่า R² สูงที่สุด โดยเฉพาะ CM-GRU (Look-back = 9 เดือน, Hidden Units = 20) ให้ค่า R² เท่ากับ 0.94 ซึ่งสูงที่สุดในทุกรายการ และ มีค่า RMSE ต่ำที่สุดในกลุ่มที่ 5,148 แสดงถึงความสามารถของแบบจำลองในการเรียนรู้และทำนายค่าที่มีความสอดคล้องกับข้อมูลจริงได้ดี

(ข) แบบจำลองรายอุทยาน (IPM) แบบจำลอง IPM-LSTM และ IPM-GRU ให้ค่า MAE ต่ำที่สุด (2,389 และ 2,421 ตามลำดับ) แต่มีค่า R^2 ต่ำมาก (0.00 และ 0.06) ซึ่งสะท้อนว่าแม้จะสามารถทำนายค่าตัวเลขได้ใกล้เคียงในเชิงปริมาณ แต่ไม่สามารถอธิบายความสอดคล้องระหว่างค่าทำนายกับค่าจริงได้ดี ผลลัพธ์ในข้อ 4.1 (ผลการทดลอง ระยะที่ 1) ยังพบว่าแบบจำลองรายอุทยานส่วนใหญ่ให้ค่า R^2 ตีลบ โดยเฉพาะเมื่อ Look-back ยาวขึ้น เช่น 6 หรือ 9 เดือน แสดงให้เห็นว่าแบบจำลองไม่สามารถเรียนรู้รูปแบบเชิงเวลาได้อย่างมีประสิทธิภาพ สาเหตุสำคัญอาจมาจากจำนวนข้อมูลต่ออุทยานที่จำกัด ความผันผวนของข้อมูลรายเดือนที่สูง และความแตกต่างของลักษณะการท่องเที่ยวของแต่ละอุทยาน ส่งผลให้แบบจำลองรายอุทยานมีประสิทธิภาพโดยรวมต่ำกว่ากลุ่มอื่นอย่างชัดเจน

(ค) แบบจำลองแบบจัดกลุ่ม (CRM และ CUM) แบบจำลอง CRM-RNN ให้ค่า R^2 เท่ากับ 0.91 ซึ่งอยู่ในระดับค่อนข้างดี แสดงให้เห็นว่าการจัดกลุ่มอุทยานตามภูมิภาคช่วยให้แบบจำลองเรียนรู้แนวโน้มร่วมของอุทยานที่มีลักษณะคล้ายกันได้ อย่างไรก็ตาม ผลลัพธ์ยังไม่สูงเท่ากลุ่ม CM ส่วนแบบจำลอง CUM-GRU และ CUM-LSTM ได้ค่า R^2 ประมาณ 0.72 ซึ่งอยู่ในระดับปานกลาง อาจเนื่องจากความแตกต่างของลักษณะอุทยาน เช่น ภูเขา ทะเล หรือ น้ำตก ที่ทำให้แบบจำลองเรียนรู้แนวโน้มของข้อมูลได้ไม่สม่ำเสมอ

โดยสรุป CM-GRU (Look-back = 9 เดือน, Hidden Units = 20) เป็นแบบจำลองที่มีความสมดุลระหว่างค่าความคลาดเคลื่อนต่ำ (RMSE ต่ำ) และความสามารถในการอธิบายความสอดคล้องของข้อมูลจริง (R^2 สูง = 0.94) จึงถือว่าเป็นแบบจำลองที่ให้ผลลัพธ์ดีที่สุดโดยรวมในระยะที่ 1

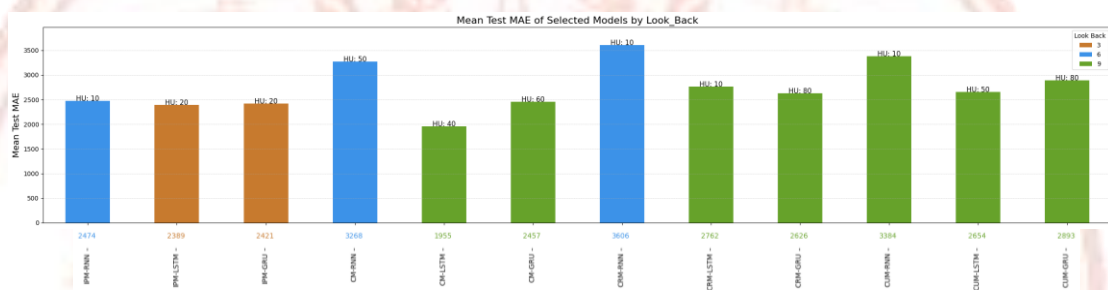
4.3.2 สรุปผลการทดลองระยะที่ 2

จากทุกแนวทางการทดลองเช่นเดียวกับระยะที่ 1 โดยในระยะที่ 2 มีการเพิ่มฟีเจอร์ประกอบอื่น ๆ ได้แก่ ข้อมูลเดือน อุณหภูมิ ปริมาณยานพาหนะ และข้อมูลเชิงพื้นที่ เพื่อประเมินผลของการเพิ่มตัวแปรต่อความสามารถในการพยากรณ์ของแบบจำลอง ผลสรุปการทดลองแสดงไว้ในตารางที่ 4 - 26 และ ภาพที่ 4 - 76 ถึง ภาพที่ 4 - 78 ดังนี้

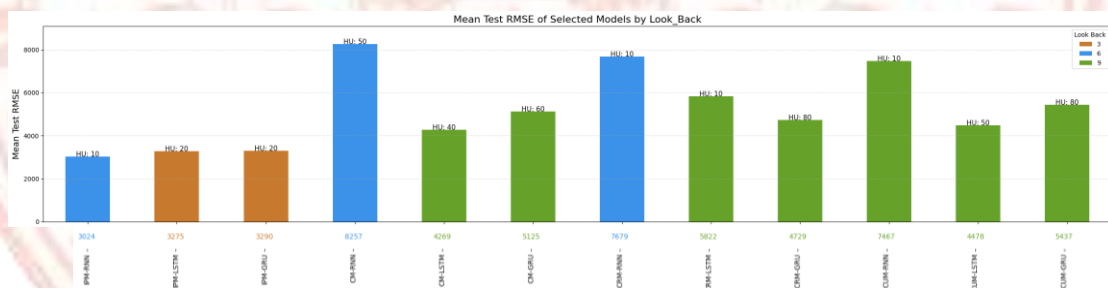
ตารางที่ 4 - 26 สรุปผลการทดลองของระยะที่ 2

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE	Name_Model
6	10	2,474	-0.79	3,024	IPM-RNN
3	20	2,389	0	3,275	IPM-LSTM
3	20	2,421	0.06	3,290	IPM-GRU
6	50	3,268	0.79	8,257	CM-RNN
9	40	1,955	0.95	4,269	CM-LSTM

Look_Back	Hidden_Units	Mean_Test_MAE	Mean_Test_R ²	Mean_Test_RMSE	Name_Model
9	60	2,457	0.93	5,125	CM-GRU
6	10	3,606	0.76	7,679	CRM-RNN
9	10	2,762	0.85	5,822	CRM-LSTM
9	80	2,626	0.87	4,729	CRM-GRU
9	10	3,384	0.78	7,467	CUM-RNN
9	50	2,654	0.89	4,478	CUM-LSTM
9	80	2,893	0.84	5,437	CUM-GRU



ภาพที่ 4 - 76 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ MAE จากการสรุปผลการทดลองของระยะที่ 2



ภาพที่ 4 - 77 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ RMSE จากการสรุปผลการทดลองของระยะที่ 2



ภาพที่ 4 - 78 กราฟแสดงการเปรียบเทียบค่าเฉลี่ยของ R² จากการสรุปผลการทดลองของระยะที่ 2

จากตารางที่ 4 - 26 และ ภาพที่ 4 - 76, ภาพที่ 4 - 77, ภาพที่ 4 - 78 พบว่าแบบจำลองที่มีประสิทธิภาพดีที่สุดของระยะที่ 2 คือ CM-LSTM ซึ่งใช้ Look-back = 9 เดือน และ Hidden Units = 40 หน่วย โดยได้ค่า MAE = 1,955 ต่ำที่สุดในทุกแบบจำลอง, RMSE = 4,269 ต่ำที่สุด และ $R^2 = 0.95$ ซึ่งสูงที่สุดในทุกการทดลอง แสดงให้เห็นว่าแบบจำลองนี้สามารถเรียนรู้ความสอดคล้องของข้อมูลจริงได้ดี และมีความสามารถในการจับลำดับข้อมูลเชิงเวลาในสภาพแวดล้อมที่มีฟีเจอร์หลากหลาย เมื่อวิเคราะห์และเปรียบเทียบในแต่ละแนวทางการทดลอง พบว่า

(ก) แบบจำลองรวม (CM) แบบจำลองกลุ่มนี้ยังคงให้ผลลัพธ์ที่ดีที่สุดในภาพรวม โดยเฉพาะ CM-LSTM และ CM-GRU ซึ่งใช้ข้อมูลรวมจากทุกอุทยาน สามารถเรียนรู้รูปแบบเชิงเวลาที่ซับซ้อนได้ดีกว่าแบบจำลองอื่น โดย CM-LSTM ให้ค่า $R^2 = 0.95$ และ RMSE = 4,269 ซึ่งดีที่สุดในทุกการทดลอง ส่วน CM-GRU (Look-back = 9 เดือน, Hidden Units = 60) ให้ค่า $R^2 = 0.93$ และ RMSE = 5,125 ถือว่าอยู่ในระดับที่ดีมาก แสดงถึงความสามารถของแบบจำลองในการอธิบายความสัมพันธ์ระหว่างค่าทำนายกับค่าจริงได้ใกล้เคียงกัน

(ข) แบบจำลองรายอุทยาน (IPM) ในระยะที่ 2 แบบจำลองรายอุทยานยังคงให้ประสิทธิภาพต่ำกว่ากลุ่มอื่น แม้จะได้ค่า MAE ในระดับต่ำ เช่น IPM-LSTM = 2,389 และ IPM-GRU = 2,421 แต่ค่า R^2 กลับอยู่ในระดับต่ำมาก (0.00 และ 0.06) และในบางกรณี เช่น IPM-RNN (Look-back = 6 เดือน, Hidden Units = 10) พบว่า $R^2 = -0.79$ ซึ่งเป็นค่าติดลบ สะท้อนว่าแบบจำลองไม่สามารถอธิบายความสอดคล้องของข้อมูลจริงได้ดี หรือทำนายได้แย่กว่าการใช้ค่าเฉลี่ยของข้อมูลจริง สาเหตุสำคัญมาจากจำนวนข้อมูลต่ออุทยานที่จำกัด ความผันผวนสูงของจำนวนนักท่องเที่ยวรายเดือน และความแตกต่างด้านลักษณะของแต่ละอุทยาน ซึ่งทำให้แบบจำลองเรียนรู้รูปแบบเชิงเวลาได้ยาก แม้การเพิ่มฟีเจอร์ในระยะนี้จะช่วยให้ค่า R^2 ของ IPM ดีขึ้นเล็กน้อยจากระยะที่ 1 แต่โดยรวมยังอยู่ในระดับต่ำ และมีข้อจำกัดชัดเจนในการพยากรณ์ข้อมูลรายอุทยานแบบแยกพื้นที่

(ค) แบบจำลองแบบจัดกลุ่ม (CRM และ CUM) ผลการทดลองของแบบจำลองจัดกลุ่มดีขึ้นกว่าระยะที่ 1 อย่างเห็นได้ชัด โดยเฉพาะ CRM-LSTM, CRM-GRU และ CUM-LSTM ที่ใช้ Look-back = 9 เดือน ซึ่งได้ค่า R^2 อยู่ในช่วง 0.84 - 0.89 และ RMSE ต่ำกว่า 5,500 ในหลายกรณี เช่น CRM-GRU ($R^2 = 0.87$) และ CUM-LSTM ($R^2 = 0.89$) ผลดังกล่าวชี้ให้เห็นว่าการจัดกลุ่มอุทยานตามลักษณะหรือภูมิภาคช่วยให้แบบจำลองสามารถเรียนรู้รูปแบบร่วมของอุทยานที่มีคุณสมบัติใกล้เคียงกันได้ดีขึ้น โดยเฉพาะแบบจำลองที่มีโครงสร้างสามารถจำลำดับข้อมูลยาว เช่น LSTM และ GRU

นอกจากนี้ ในระยะที่ 2 พบว่า Look-back = 6 และ 9 เดือน ปรากฏในชุดแบบจำลองที่มีผลดีที่สุดถึง 9 ใน 12 แบบ ซึ่งต่างจากระยะที่ 1 ที่ไม่พบ Look-back = 6 เลย สะท้อนว่าการเพิ่มฟีเจอร์ (เช่น เดือน อุณหภูมิ ปริมาณยานพาหนะ และข้อมูลเชิงพื้นที่) ช่วยให้แบบจำลองสามารถ

เรียนรู้บริบทของข้อมูลในอดีตได้ดียิ่งขึ้น และใช้ศักยภาพในการจำลำดับข้อมูลระยะยาวได้อย่างเต็มที่ โดยสรุป CM-LSTM (Look-back = 9 เดือน, Hidden Units = 40) เป็นแบบจำลองที่ให้ผลดีที่สุด ในภาพรวมของระยะที่ 2 ทั้งในด้านค่าความคลาดเคลื่อนที่ต่ำ (MAE และ RMSE ต่ำที่สุด) และ ความสอดคล้องของค่าทำนายกับค่าจริง (R^2 สูงสุด = 0.95) แสดงให้เห็นถึงศักยภาพในการพยากรณ์ และอธิบายความสัมพันธ์ของข้อมูลได้อย่างมีประสิทธิภาพสูงสุด

4.3.3 สรุปผลการเปรียบเทียบระหว่าง ระยะที่ 1 และ ระยะที่ 2

ภายหลังจากการสรุปผลในแต่ละระยะแล้ว ผู้วิจัยได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองที่ดีที่สุดจากระยะที่ 1 และระยะที่ 2 เพื่อวิเคราะห์ความเปลี่ยนแปลงของผลลัพธ์ที่เกิดจากการเพิ่มฟีเจอร์ ผลสรุปการทดลองแสดงไว้ใน ตารางที่ 4 - 27 ดังนี้

ตารางที่ 4 - 27 ผลการเปรียบเทียบแบบจำลองที่คัดเลือกจากระยะที่ 1 และระยะที่ 2

Model	Look-back	Hidden Units	Mean Test MAE	Mean Test RMSE	Mean Test R^2	หมายเหตุ
CM-GRU (Phase 1)	9	20	4,723	5,148	0.94	แบบจำลองที่คัดเลือกจากระยะที่ 1
CM-LSTM (Phase 2)	9	40	1,955	4,269	0.95	แบบจำลองที่คัดเลือกจากระยะที่ 2

จากตารางที่ 4 - 27 พบว่าแบบจำลอง CM-LSTM (Phase 2) ซึ่งใช้ Look-back = 9 เดือน และ Hidden Units = 40 หน่วย ให้ผลลัพธ์โดยรวมดีกว่าแบบจำลอง CM-GRU (Phase 1) ที่ใช้ Look-back และช่วงเวลาข้อมูลเท่ากัน แต่มี Hidden Units = 20 หน่วย โดยค่า MAE ลดลงจาก 4,723 เหลือ 1,955 (ลดลงประมาณ 58.6%) และค่า RMSE ลดลงจาก 5,148 เหลือ 4,269 (ลดลงประมาณ 17.1%) ขณะที่ค่า R^2 เพิ่มขึ้นจาก 0.94 เป็น 0.95 แม้การเพิ่มขึ้นจะไม่มากนัก แต่สะท้อนให้เห็นว่าแบบจำลองในระยะที่ 2 สามารถอธิบายความสอดคล้องระหว่างค่าทำนายกับค่าจริงได้ดีขึ้น

ความแตกต่างของประสิทธิภาพระหว่างสองระยะ เกิดจากการปรับปรุงชุดข้อมูลและโครงสร้างของแบบจำลองในระยะที่ 2 โดยในระยะที่ 1 ใช้เพียงข้อมูลจำนวนนักท่องเที่ยวย้อนหลังเพียงฟีเจอร์เดียว ขณะที่ในระยะที่ 2 ได้เพิ่มฟีเจอร์สนับสนุนอื่น ๆ ได้แก่ ข้อมูลเดือน ข้อมูลอุณหภูมิ ข้อมูลเชิงพื้นที่ และข้อมูลปริมาณรถยนต์ ซึ่งช่วยให้แบบจำลองเข้าใจบริบทของข้อมูลได้ครอบคลุม และสัมพันธ์กับพฤติกรรมของนักท่องเที่ยวมากขึ้น นอกจากนี้ แบบจำลองที่ดีที่สุดในระยะที่ 2 คือ LSTM ซึ่งมีความสามารถในการจดจำลำดับข้อมูลที่ยาวและซับซ้อนได้ดีกว่า GRU ที่ใช้ในระยะที่ 1

แม้จะใช้ค่า Look-back เท่ากันที่ 9 เดือน แต่การเพิ่มจำนวนหน่วยซ่อน (Hidden Units) ที่เหมาะสม และการมีพีเจอร์ประกอบมากขึ้น ทำให้แบบจำลองในระยะที่ 2 สามารถเรียนรู้ความสัมพันธ์เชิงเวลาได้ดีกว่าแบบจำลองในระยะที่ 1 อย่างชัดเจน

โดยสรุป แบบจำลอง CM-LSTM (Phase 2) แสดงประสิทธิภาพที่ดีกว่าในทุกตัวชี้วัด โดยเฉพาะในด้านค่าความคลาดเคลื่อน (MAE และ RMSE) ที่ลดลงอย่างมีนัยสำคัญ และค่า R^2 ที่เพิ่มขึ้นเล็กน้อย แต่มีความหมายในเชิงการอธิบายข้อมูลจริงได้ดียิ่งขึ้น ซึ่งยืนยันว่า การเพิ่มพีเจอร์ และการเลือกโครงสร้างแบบจำลองที่เหมาะสมส่งผลให้ผลลัพธ์ในระยะที่ 2 มีแนวโน้มดีกว่าระยะที่ 1 อย่างชัดเจน

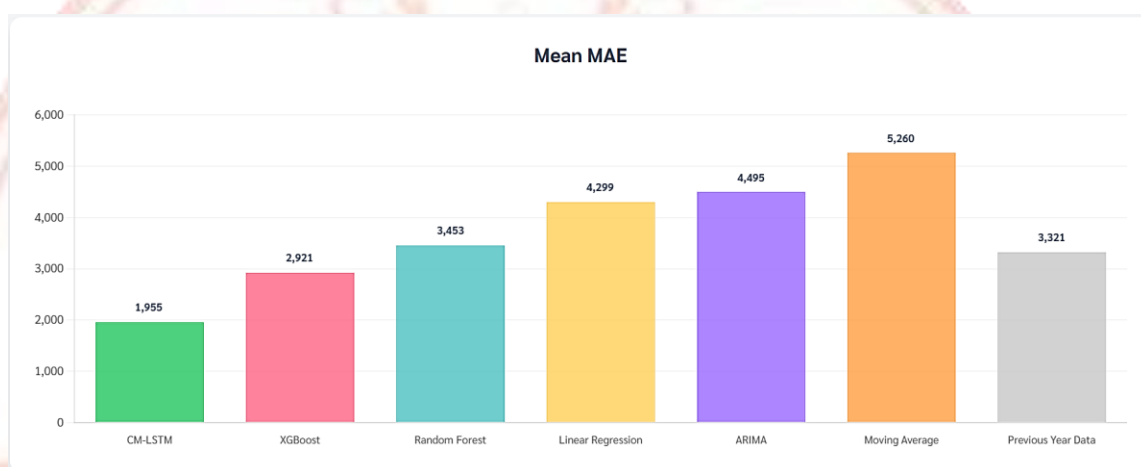
4.4 การเปรียบเทียบผลการทดลองกับเทคนิคอื่น ๆ

ผู้วิจัยได้นำแบบจำลองที่มีประสิทธิภาพดีที่สุด (CM-LSTM) มาเปรียบเทียบเพิ่มเติมกับเทคนิคอื่นที่ส่วนใหญ่มักใช้ในการพยากรณ์ข้อมูลในลักษณะคล้ายกัน เพื่อประเมินว่าผลลัพธ์ของแบบจำลองที่พัฒนาขึ้นมีแนวโน้มแตกต่างหรือใกล้เคียงกับเทคนิคส่วนใหญ่หรือไม่ เทคนิคที่นำมาเปรียบเทียบได้แก่ ARIMA, Moving Average, Linear Regression, Random Forest และ XGBoost โดยเลือกใช้เฉพาะค่าที่ดีที่สุดของแต่ละเทคนิค และจากการทดสอบในแนวทางเดียวกับงานวิจัยนี้ได้แก่ รายอุทยาน (Individual), การรวมทุกอุทยาน (Combined) และการจัดกลุ่ม (Clustered) เพื่อให้มีการเปรียบเทียบภายใต้เงื่อนไขข้อมูลเดียวกัน ทั้งนี้ ผู้วิจัยยังได้เปรียบเทียบร่วมกับ “วิธีการประมาณการจากข้อมูลปีก่อน” ซึ่งใช้ค่าจำนวนนักท่องเที่ยวในเดือนเดียวกันของปีก่อนมาเป็นค่าประมาณในปีถัดไป ซึ่งเป็นแนวทางที่หน่วยงานมักนำไปใช้ในการประมาณการจำนวนนักท่องเที่ยวในปีถัดไป โดยมีรายละเอียดการเปรียบเทียบอยู่ในตารางที่ 4 - 28 และ ภาพที่ 4 - 79 ถึง ภาพที่ 4 - 81 ดังนี้

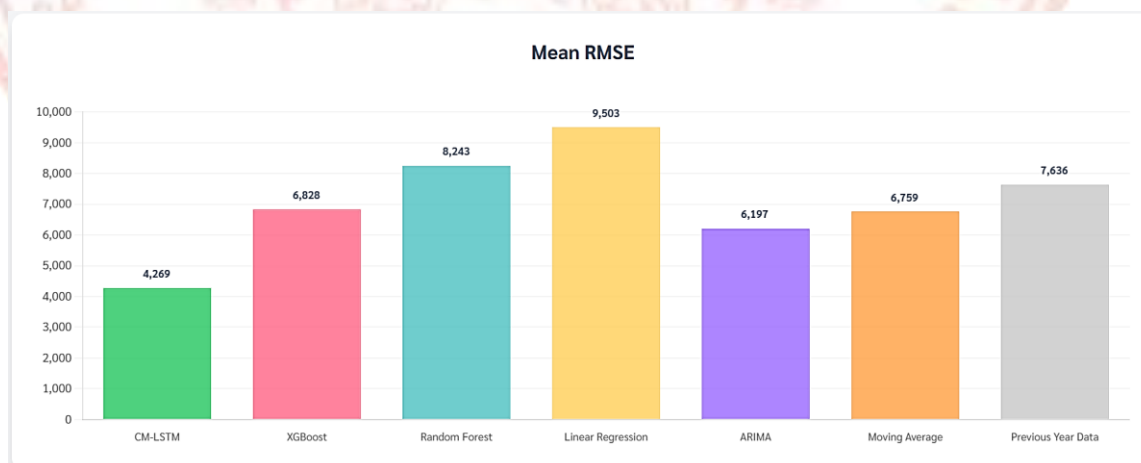
ตารางที่ 4 - 28 การเปรียบเทียบประสิทธิภาพของแบบจำลองที่พัฒนากับเทคนิคทั่วไป

แบบจำลอง	Mean MAE	Mean RMSE	Mean R ²	หมายเหตุ
CM-LSTM	1,955	4,269	0.95	Single model + all feature (ใช้ข้อมูลทุกอุทยาน สร้างแบบจำลองเดี่ยว และใช้พีเจอร์ทั้งหมด)
XGBoost	2,921	6,828	0.91	Cluster model + all feature (แบ่งกลุ่มอุทยาน เพื่อสร้างแบบจำลอง และใช้พีเจอร์ทั้งหมด)
Random Forest	3,453	8,243	0.86	Cluster model + all feature (แบ่งกลุ่มอุทยาน เพื่อสร้างแบบจำลอง และใช้พีเจอร์ทั้งหมด)
Linear Regression	4,299	9,503	0.83	Cluster model + all feature (แบ่งกลุ่มอุทยาน เพื่อสร้างแบบจำลอง และใช้พีเจอร์ทั้งหมด)

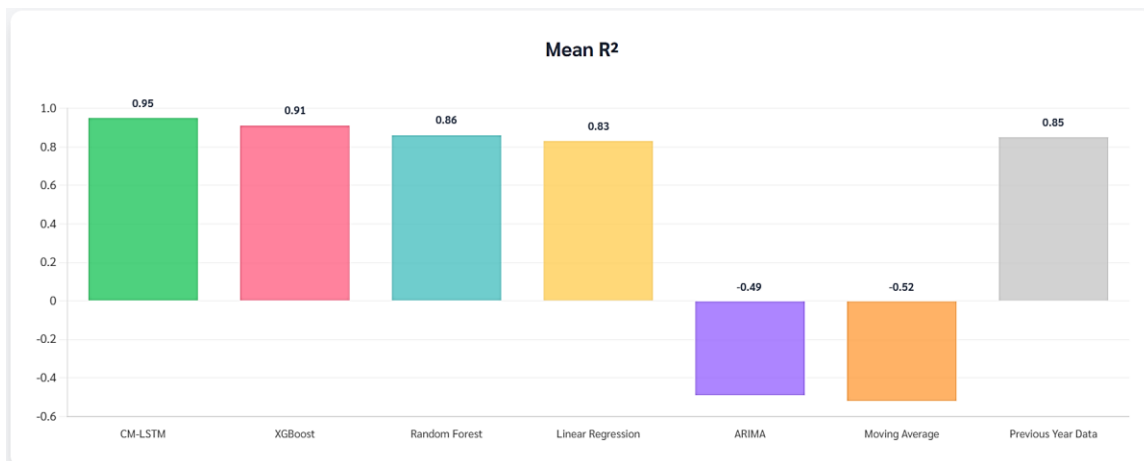
แบบจำลอง	Mean MAE	Mean RMSE	Mean R ²	หมายเหตุ
ARIMA	4,495	6,197	-0.49	Local model + tourists (สร้างแบบจำลองแยก รายอุทยาน และใช้เฉพาะข้อมูลนักท่องเที่ยว)
Moving Average	5,260	6,759	-0.52	Local model + tourists (สร้างแบบจำลองแยก รายอุทยาน และใช้เฉพาะข้อมูลนักท่องเที่ยว)
Baseline	3,321	7,636	0.85	Local model + tourists (ใช้วิธีประมาณการ จากข้อมูลปีก่อนแบบรายอุทยาน Previous Year Data)



ภาพที่ 4 - 79 กราฟเปรียบเทียบค่า MAE ระหว่างแบบจำลองที่พัฒนากับเทคนิคทั่วไป



ภาพที่ 4 - 80 กราฟเปรียบเทียบค่า RMSE ระหว่างแบบจำลองที่พัฒนากับเทคนิคทั่วไป



ภาพที่ 4 - 81 กราฟเปรียบเทียบค่า R^2 ระหว่างแบบจำลองที่พัฒนากับเทคนิคทั่วไป

จากการเปรียบเทียบแบบจำลองในตารางที่ 4 - 28 ประกอบกับ ภาพที่ 4 - 79, ภาพที่ 4 - 80, ภาพที่ 4 - 81 พบว่าแบบจำลอง CM-LSTM ซึ่งพัฒนาขึ้นในงานวิจัยนี้ ที่ใช้ข้อมูลรวมจากทุกอุทยาน และใช้ฟีเจอร์ทั้งหมด ให้ผลการพยากรณ์ที่มีแนวโน้มสอดคล้องกับข้อมูลจริงมากที่สุดเมื่อเทียบกับเทคนิคอื่น ๆ สำหรับเทคนิคกลุ่ม Machine Learning ได้แก่ XGBoost และ Random Forest ที่ใช้แนวทางการจัดกลุ่มตามภูมิภาคและใช้ฟีเจอร์ทั้งหมด พบว่าให้ผลลัพธ์รองลงมา โดย XGBoost มีค่าความคลาดเคลื่อน Mean MAE = 2,921 และ Mean RMSE = 6,828 และค่า Mean R^2 = 0.91 ที่ใกล้เคียงกับ CM-LSTM มากที่สุด ส่วน Random Forest มีแนวโน้มผลลัพธ์ต่ำกว่า XGBoost เล็กน้อย ในขณะที่เทคนิคเชิงสถิติแบบดั้งเดิม เช่น ARIMA และ Moving Average ซึ่งใช้ข้อมูลเฉพาะจำนวนนักท่องเที่ยวแบบรายอุทยานมีข้อจำกัดในด้านการรองรับข้อมูลหลายฟีเจอร์ การรวมข้อมูลและการจัดกลุ่ม ทำให้ผลลัพธ์ที่ได้มีค่าความคลาดเคลื่อนสูงกว่ามาก และ ค่า Mean R^2 ตดลบ ซึ่งแสดงถึงข้อจำกัดในการนำไปใช้งานกับข้อมูลที่มีความผันผวนสูงและมีหลายปัจจัยเกี่ยวข้อง โดยเฉพาะเมื่อมีการหาค่าเฉลี่ยรวมจากทุกอุทยาน สำหรับ Linear Regression แม้จะสามารถรองรับฟีเจอร์หลายตัว และใช้การรวมข้อมูลและการจัดกลุ่มได้ แต่เนื่องจากโครงสร้างของแบบจำลองไม่สามารถเรียนรู้ลำดับเวลาของข้อมูลได้โดยตรง ซึ่งอาจส่งผลให้มีค่าความคลาดเคลื่อนที่ค่อนข้างสูงและค่า Mean R^2 ในระดับต่ำ นอกจากนี้ เมื่อเปรียบเทียบกับ "วิธีการประมาณการจากข้อมูลปีก่อน" ซึ่งเป็นแนวทางพื้นฐานที่หน่วยงานมักนำมาใช้ในการประมาณการจำนวนนักท่องเที่ยว โดยการใช้ค่าของเดือนเดียวกันในปีที่ผ่านมา แม้ว่าวิธีนี้จะง่ายต่อการใช้งาน แต่ผลที่ได้ยังมีความคลาดเคลื่อนที่สูงกว่าแบบจำลองที่พัฒนาขึ้นในงานวิจัยนี้ ซึ่งสรุปโดยรวม ผลลัพธ์จากการเปรียบเทียบนี้แสดงให้เห็นว่าแบบจำลอง CM-LSTM ที่พัฒนาขึ้นมีศักยภาพที่ดีในการนำไปใช้เพื่อการพยากรณ์จำนวนนักท่องเที่ยว โดยสามารถจับแนวโน้มและความสัมพันธ์ของข้อมูลได้ดีกว่าเทคนิคที่มีการใช้กันโดยทั่วไปในหลายแง่มุม ทั้งในด้านความสอดคล้องกับข้อมูลจริงและการลดความคลาดเคลื่อนของการพยากรณ์

4.5 การวิเคราะห์ผลกระทบของฟีเจอร์

ผู้วิจัยได้นำแบบจำลองที่มีประสิทธิภาพสูงสุดซึ่งพัฒนาขึ้นในงานวิจัยนี้ (โดยใช้ข้อมูลย้อนหลังจำนวน 9 เดือน (Look-back = 9 เดือน) และจำนวนหน่วยซ่อน (Hidden Units) เท่ากับ 40 หน่วย) มาทำการวิเคราะห์ผลกระทบของฟีเจอร์ต่าง ๆ เพื่อทำความเข้าใจถึงปัจจัยที่มีอิทธิพลต่อประสิทธิภาพของแบบจำลอง โดยได้ดำเนินการทดสอบความสำคัญของแต่ละฟีเจอร์ที่เกี่ยวข้องทั้งหมด ได้แก่ จำนวนนักท่องเที่ยว (TouristCount), เดือน (Month), อุณหภูมิ (Temperature), คุณลักษณะเชิงพื้นที่ของอุทยาน (AreaFeatures) และปริมาณรถยนต์ (VehicleCount) พร้อมทั้งประเมินผลกระทบที่เกิดขึ้นต่อค่าตัวชี้วัดหลักทั้ง 3 ค่า ได้แก่ MAE, RMSE และ R^2 โดยผลการทดลองมีรายละเอียด ดังนี้

ตารางที่ 4 - 29 รายละเอียดและความหมายของฟีเจอร์

ชื่อฟีเจอร์	ความหมายโดยย่อ	ตัวอย่างค่า
TouristCount	จำนวนผู้เข้าชมอุทยานแห่งชาติในแต่ละเดือนย้อนหลัง	1,072
Month	ลำดับเดือน (1 = ม.ค., ..., 12 = ธ.ค.)	10
Temperature	อุณหภูมิสูงสุด/ต่ำสุดในแต่ละเดือนย้อนหลัง, อุณหภูมิเฉลี่ยและค่าความแปรปรวน	33 (Max) 24 (Min) 27.75 (Avg) 6.97 (Std)
AreaFeatures	ข้อมูลภูมิภาคที่ตั้งและลักษณะเด่นของอุทยาน (ระบุรหัส-กลุ่มเฉพาะ)	A_004 (Region) C_001 (Unique)
VehicleCount	จำนวนยานพาหนะที่เข้าอุทยานในแต่ละเดือนย้อนหลัง	139

ตารางที่ 4 - 30 ผลการวิเคราะห์ผลกระทบของฟีเจอร์ต่อประสิทธิภาพของแบบจำลอง

อันดับ	ชุดฟีเจอร์ (Feature Set)	MAE	RMSE	R^2
1	TouristCount + All Feature	1,955	4,269	0.95
2	TouristCount + Temperature	2,268	5,249	0.93
3	TouristCount + Month + Vehicle	2,266	5,026	0.93
4	TouristCount + Month + Temperature + Vehicle	2,450	5,302	0.93
5	TouristCount + Month + Temperature	2,398	5,323	0.93
6	TouristCount + Month + Temperature + Area	2,290	5,491	0.92
7	TouristCount + Month + Area	2,249	5,547	0.92
8	TouristCount + Month	2,225	5,483	0.92

9	TouristCount + Temperature + Area	2,361	5,793	0.91
10	TouristCount + Month + Vehicle + Area	2,429	5,720	0.91
11	TouristCount + Temperature + Vehicle + Area	2,385	6,070	0.90
12	TouristCount + Temperature + Vehicle	2,593	6,249	0.90
13	TouristCount + Area	2,627	6,539	0.89
14	TouristCount + Area + Vehicle	2,873	6,389	0.89
15	TouristCount + Vehicle	2,868	7,322	0.86

จากตารางที่ 4 - 30 เมื่อพิจารณาผลลัพธ์โดยรวม จะเห็นได้ว่าไม่มีชุดฟีเจอร์ใดให้ค่า R^2 สูงเท่ากับแบบจำลองที่ใช้ฟีเจอร์ทั้งหมดร่วมกัน (TouristCount + All) ซึ่งเป็นลำดับที่ 1 โดยมีค่า MAE ต่ำที่สุด (1,955) และ RMSE ต่ำที่สุด (4,269) รวมถึง R^2 สูงที่สุด (0.95) แสดงให้เห็นว่าฟีเจอร์ทั้งหมดมีผลต่อประสิทธิภาพของแบบจำลองอย่างชัดเจน โดยเฉพาะเมื่อใช้ร่วมกัน ทำให้แบบจำลองสามารถเรียนรู้และจับแนวโน้มของข้อมูลได้ดีที่สุด

เมื่อพิจารณารายละเอียดของฟีเจอร์อื่น ๆ ที่มีลำดับรองลงมา เช่น ลำดับที่ 2, 3, 4 และลำดับที่ 5 พบว่าทุกชุดให้ค่า MAE ต่ำกว่า 2,500 และ R^2 อยู่ในช่วง 0.93 ซึ่งแสดงให้เห็นว่า Month และ Temperature มีอิทธิพลต่อประสิทธิภาพการพยากรณ์อย่างมาก โดยเฉพาะเมื่อใช้ร่วมกับ TouristCount ขณะที่กลุ่มลำดับที่ 6 ถึง 8 แม้จะยังคงให้ค่า R^2 ในระดับ 0.92 แต่ค่าความคลาดเคลื่อนเพิ่มสูงขึ้นเล็กน้อย สะท้อนว่าเมื่อใช้ Month เพียงอย่างเดียว หรือร่วมกับ Area โดยไม่มี Temperature จะทำให้ความสามารถของแบบจำลองลดลง

ลำดับถัดมา ได้แก่ ลำดับที่ 9, 10, 11 และ ลำดับที่ 12 ให้ค่า R^2 ลดลงสู่ช่วง 0.90-0.91 โดยเฉพาะเมื่อ Month ถูกตัดออกไปจากชุดฟีเจอร์ ส่งผลให้แนวโน้มฤดูกาลหายไป ส่งผลต่อความแม่นยำของแบบจำลองอย่างชัดเจน และสุดท้ายคือชุดฟีเจอร์ที่ให้ผลกระทบต่อแบบจำลองน้อยที่สุด ได้แก่ ลำดับที่ 13, 14 และ ลำดับที่ 15 โดยมีค่า MAE และ RMSE สูงที่สุด และ R^2 ต่ำที่สุด อยู่ในช่วง 0.86-0.89 แสดงว่าการขาดฟีเจอร์เชิงเวลา เช่น Month และ Temperature มีผลกระทบโดยตรงต่อประสิทธิภาพของแบบจำลอง

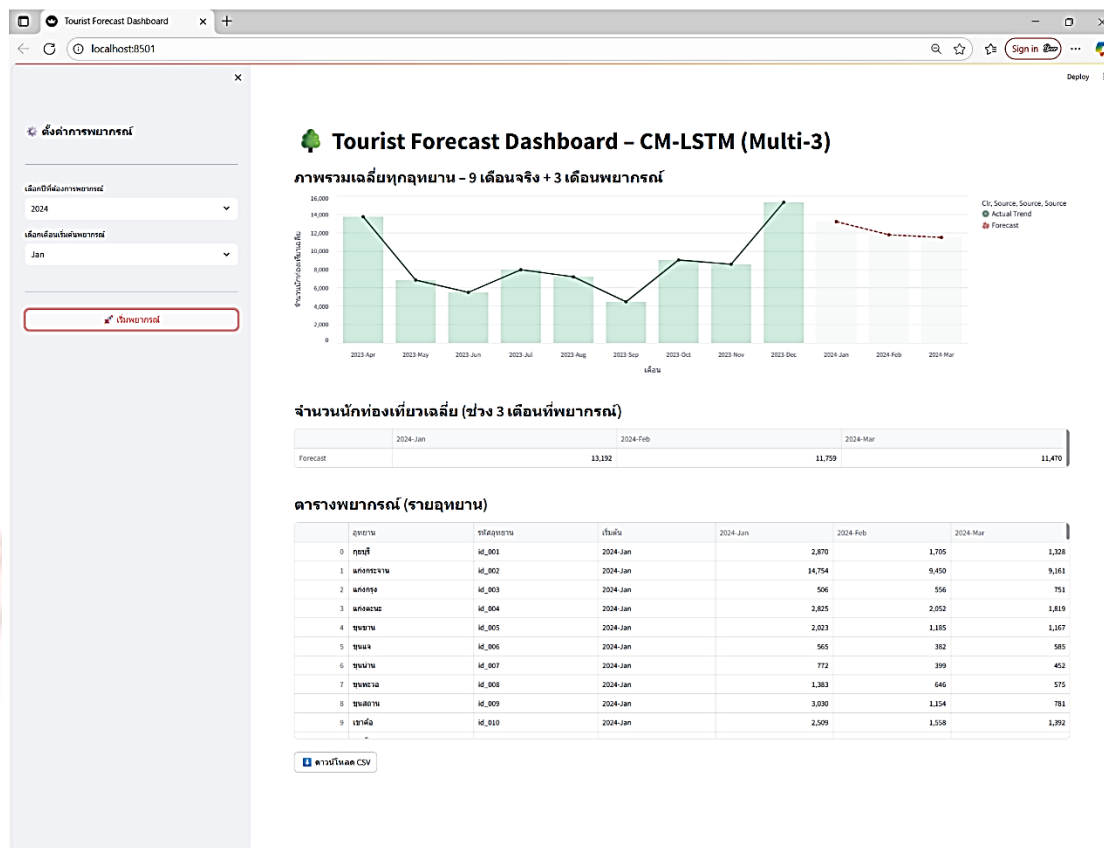
สรุปจากการทดลองชุดนี้ ผู้วิจัยพบว่าฟีเจอร์ Month และ Temperature มีบทบาทสำคัญที่สุดต่อการเพิ่มความแม่นยำและการจับแนวโน้มของข้อมูล ส่วนของ AreaFeatures และ VehicleCount แม้จะมีส่วนในการปรับปรุงประสิทธิภาพอยู่บ้าง แต่ความสำคัญในเชิงการทำนายมีน้อยกว่า Month และ Temperature อย่างเห็นได้ชัด แสดงให้เห็นว่าการรวมฟีเจอร์ที่เกี่ยวข้องกับแนวโน้มเชิงเวลาและฟีเจอร์สภาพแวดล้อมสามารถช่วยเพิ่มประสิทธิภาพของแบบจำลอง LSTM ในการพยากรณ์จำนวนนักท่องเที่ยวได้ดียิ่งขึ้น

4.6 การพัฒนาเว็บแอปพลิเคชันเพื่อการใช้งาน

หลังจากได้แบบจำลองที่มีประสิทธิภาพสูงสุดจากงานวิจัยนี้ ผู้วิจัยได้นำแบบจำลองดังกล่าวมาพัฒนาเป็นเว็บแอปพลิเคชันต้นแบบ เพื่อแสดงตัวอย่างการนำไปใช้งาน โดยมีวัตถุประสงค์เพื่อแสดงให้เห็นถึงการนำไปประยุกต์ใช้งานสำหรับการพยากรณ์จำนวนนักท่องเที่ยวล่วงหน้า โดยให้ผู้ใช้สามารถเลือกปีและเดือนที่ต้องการให้ระบบเริ่มต้นการพยากรณ์ และระบบจะดำเนินการทำนายจำนวนนักท่องเที่ยวล่วงหน้า 3 เดือนถัดไป โดยใช้ข้อมูลย้อนหลัง 9 เดือนตามโครงสร้างของแบบจำลอง CM-LSTM

การพัฒนาเว็บแอปพลิเคชันดำเนินการผ่านเครื่องมือ Streamlit ซึ่งเป็นไลบรารีสำหรับสร้างเว็บแอปพลิเคชันในภาษา Python ที่สามารถเชื่อมโยงกับแบบจำลองและระบบประมวลผลข้อมูลแบบอัตโนมัติ เมื่อผู้ใช้เลือกเดือนและปี ระบบจะตรวจสอบว่าในระบบมีข้อมูลย้อนหลังครบ 9 เดือนหรือไม่ หากมีข้อมูลครบ ระบบจะนำข้อมูลย้อนหลังทั้งหมดของอุทยานแห่งชาติทั้ง 146 แห่งไปประมวลผลผ่านโมเดล CM-LSTM และแสดงผลการพยากรณ์จำนวนนักท่องเที่ยวสำหรับ 3 เดือนถัดไป โดยจะแสดงผลทั้งในรูปแบบตารางผลลัพธ์รายอุทยาน และแสดงกราฟเส้นแนวโน้มรวมซึ่งเป็นค่าเฉลี่ยของแต่ละเดือนจากอุทยานทั้งหมด

กราฟที่แสดงผลจะประกอบด้วยเส้นแสดงข้อมูลจริงย้อนหลัง 9 เดือน และเส้นแสดงค่าที่พยากรณ์ได้สำหรับ 3 เดือนถัดไป โดยใช้เส้นประแทนค่าพยากรณ์เพื่อให้เห็นความต่อเนื่องของแนวโน้มได้ชัดเจน ส่วนในตารางจะแสดงค่าพยากรณ์แยกตามอุทยาน และมีคอลัมน์รวมแสดงค่าเฉลี่ยของแต่ละเดือนเพื่อให้ผู้ใช้สามารถประเมินแนวโน้มภาพรวมได้ โดยมีรูปแบบการแสดงผลการพยากรณ์ผ่านเว็บแอปพลิเคชันดังภาพที่ 4 - 82



ภาพที่ 4 - 82 ตัวอย่างการแสดงผลกราฟแนวโน้มและตารางผลลัพธ์ของการพยากรณ์จำนวนนักท่องเที่ยวผ่านเว็บแอปพลิเคชัน

การพยากรณ์ผ่านเว็บแอปพลิเคชัน พบว่าแบบจำลอง CM-LSTM สามารถพยากรณ์แนวโน้มจำนวนนักท่องเที่ยวเฉลี่ยของทั้ง 146 อุทยาน 3 เดือนล่วงหน้า (มกราคม - มีนาคม 2024) พร้อมกับการแสดงผลแนวโน้ม 9 เดือนย้อนหลัง (เมษายน - ธันวาคม 2023) ซึ่งจากผลการพยากรณ์พบว่าจำนวนนักท่องเที่ยวเฉลี่ยจากทุกอุทยานมีแนวโน้มลดลงอย่างต่อเนื่องในช่วงต้นปี โดยมีค่าเฉลี่ยที่พยากรณ์ได้ในเดือนมกราคม 2024 เท่ากับ 13,192 คน เดือนกุมภาพันธ์ลดลงเฉลี่ย 11,759 คน และเดือนมีนาคมลดลงเฉลี่ยเหลือ 11,470 คน ตามลำดับ นอกจากนี้ เว็บแอปพลิเคชันยังสามารถแสดงค่าพยากรณ์ของจำนวนนักท่องเที่ยวของทุกอุทยาน จำนวน 146 แห่ง ในรูปแบบตาราง ซึ่งจะช่วยให้เห็นถึงแนวโน้มของแต่ละพื้นที่ได้ ทั้งนี้ผู้ใช้งานยังสามารถดาวน์โหลดผลลัพธ์ทั้งหมดเป็นรูปแบบของไฟล์ CSV เพื่อนำไปวิเคราะห์และใช้ประกอบการตัดสินใจในการบริหารจัดการต่อไปได้ (รายละเอียดการใช้งานเว็บแอปพลิเคชันต้นแบบสำหรับการพยากรณ์ (คู่มือ) ระบุไว้ในภาคผนวก ก)

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

จากการดำเนินการพัฒนา ประเมินผล และเปรียบเทียบแบบจำลองโครงข่ายประสาทเทียมแบบเรียกซ้ำ ได้แก่ RNN, LSTM และ GRU รวมถึงการนำผลลัพธ์จากการพัฒนามาเปรียบเทียบกับแบบจำลองที่นิยมใช้กันทั่วไป และการวิเคราะห์ความสำคัญของฟีเจอร์ต่าง ๆ เพื่อหาปัจจัยที่มีอิทธิพลต่อประสิทธิภาพของแบบจำลอง สามารถสรุปผลการวิจัยและข้อเสนอแนะได้ดังนี้

5.1 สรุปผลการวิจัย

การดำเนินการวิจัยแบ่งเป็นสองระยะ โดยระยะที่ 1 ใช้เฉพาะข้อมูลจำนวนนักท่องเที่ยวเพียงอย่างเดียว ส่วนระยะที่ 2 เพิ่มฟีเจอร์สนับสนุน ได้แก่ เดือน (Month) อุณหภูมิ (Temperature) ปริมาณรถยนต์ (VehicleCount) และข้อมูลเชิงพื้นที่ (AreaFeatures) ได้แก่ ภูมิภาค และลักษณะเด่นของแต่ละอุทยาน ซึ่งจากการทดลองทั้งหมด พบว่าแบบจำลองที่มีประสิทธิภาพสูงสุดคือ CM-LSTM (Combined Model with LSTM) ที่ใช้ข้อมูลทุกอุทยานสร้างแบบจำลองเดี่ยว (Single model) จากการทดลองระยะที่ 2 โดยใช้ข้อมูลย้อนหลัง 9 เดือน (Look-back = 9) และใช้หน่วยซ่อน 40 หน่วย (Hidden Units = 40) ให้ผลลัพธ์ที่ดีที่สุด ทั้งนี้ผู้วิจัยได้นำผลลัพธ์ที่ดีที่สุดไปประเมินเปรียบเทียบกับเทคนิคที่นิยมใช้ทั่วไป ได้แก่ ARIMA, Moving Average, Linear Regression, Random Forest, XGBoost และวิธีการประมาณการจากข้อมูลปีก่อน พบว่าแบบจำลอง CM-LSTM ให้ผลลัพธ์ดีกว่าทุกเทคนิคทั้งในด้านความคลาดเคลื่อนเฉลี่ย (MAE, RMSE) และความสามารถในการอธิบายความสัมพันธ์ของข้อมูล (R^2) นอกจากนี้ยังได้มีการวิเคราะห์ผลกระทบของฟีเจอร์แต่ละชุดต่อประสิทธิภาพของแบบจำลอง โดยทำการทดสอบฟีเจอร์ในหลากหลายรูปแบบ ตั้งแต่การใช้ร่วมกันทั้งหมด จนถึงการตัดออกเป็นชุดย่อย ผลลัพธ์ที่ได้แสดงให้เห็นว่า ฟีเจอร์ "Month" และ "Temperature" มีบทบาทสำคัญที่สุดต่อประสิทธิภาพของแบบจำลอง ในขณะที่ฟีเจอร์อื่น ๆ เช่น AreaFeatures และ VehicleCount แม้จะมีส่วนช่วย แต่ไม่เด่นชัดเท่ากับฟีเจอร์เชิงเวลาและสภาพอากาศ และจากการวิเคราะห์เปรียบเทียบแบบจำลองที่ดีที่สุดกับเทคนิคต่าง ๆ และการทดสอบผลของฟีเจอร์ที่ใช้ร่วมกัน พบว่าแบบจำลอง CM-LSTM ที่พัฒนาขึ้นในงานวิจัยนี้มีศักยภาพสูงสำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติของประเทศไทย

5.2 ข้อเสนอแนะสำหรับงานวิจัย

จากการดำเนินงานวิจัยที่ผ่านมา มีข้อเสนอแนะที่สามารถนำไปใช้ในการพัฒนาและต่อยอดงานวิจัยในอนาคต ในเรื่องของการเพิ่มข้อมูลด้านเศรษฐกิจอาจศึกษาผลกระทบจากข้อมูลด้านเศรษฐกิจที่เกี่ยวข้องกับการท่องเที่ยว เช่น รายได้เฉลี่ยต่อหัว, ดัชนีราคาหุ้นในอุตสาหกรรม

การท่องเที่ยว หรืออัตราการจ้างงานในภาคบริการ เพื่อวิเคราะห์เพิ่มเติมว่าตัวแปรทางเศรษฐกิจเหล่านี้มีความสัมพันธ์กับจำนวนนักท่องเที่ยวมากน้อยเพียงใด และจะสามารถเพิ่มประสิทธิภาพของแบบจำลองได้หรือไม่ นอกจากนี้ อาจพิจารณาเพิ่มตัวแปรด้านสิ่งแวดล้อม เช่น ปริมาณน้ำฝน ซึ่งมีความสัมพันธ์โดยตรงกับฤดูกาลและสภาพภูมิอากาศที่ส่งผลต่อการเดินทางและการเข้าถึงพื้นที่อุทยาน การนำข้อมูลดังกล่าวมารวมกับข้อมูลด้านเศรษฐกิจและการท่องเที่ยวที่มีอยู่ จะช่วยให้แบบจำลองสามารถเรียนรู้ปัจจัยเชิงเวลาและความผันผวนในแต่ละช่วงฤดูกาลได้ดียิ่งขึ้น

ในส่วนของการปรับแบบจำลอง อาจขยายการทดลองช่วงข้อมูลย้อนหลัง (Look-back) ให้ยาวขึ้นเป็น 12 เดือน เพื่อให้ครอบคลุมรอบฤดูกาลท่องเที่ยวครบหนึ่งปี และเพื่อทดสอบว่าการใช้ช่วงข้อมูลย้อนหลังที่ยาวขึ้นมีผลต่อความสามารถในการพยากรณ์ของแบบจำลองอย่างไร เมื่อเทียบกับช่วง 3, 6 และ 9 เดือนที่ได้ใช้ในการทดลองก่อนหน้านี้ การนำเทคนิคและอัลกอริทึมใหม่ ๆ มาทดลองเพิ่มเติมเพื่อเปรียบเทียบประสิทธิภาพกับแบบจำลองเดิม การเพิ่มความละเอียดของข้อมูล โดยการศึกษาการใช้ข้อมูลที่มีความถี่และละเอียดมากยิ่งขึ้น เช่น รายสัปดาห์หรือรายวัน หรือข้อมูลจากแหล่งอื่น ๆ เช่น ข้อมูลการจองห้องพัก ข้อมูลจากโซเชียลมีเดีย หรือข้อมูลการจราจรที่เกี่ยวข้องกับการเดินทางของนักท่องเที่ยว เพื่อให้แบบจำลองสามารถเรียนรู้ข้อมูลที่มีพลวัตสูงและสามารถพยากรณ์แนวโน้มได้ดียิ่งขึ้น

บรรณานุกรม

- Boonlha, K. (2019). Forecasting Numbers of Tourists in the Phu Hin Rong Kla National Park. *Nakhon Sawan Rajabhat University Journal of Science and Technology*, 11(14), 1-12.
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis Forecasting and control*.
- Brilliandy, E., Lucky, H., Hartanto, A., Suhartono, D., & Nurzaki, M. (2022). Using Regression to Predict Number of Tourism in Indonesia based of Global COVID-19 Cases. *2022 3rd International Conference on Artificial Intelligence and Data Sciences Championing Innovations in Artificial Intelligence and Data Sciences for Sustainable Future, AiDAS 2022 - Proceedings*, 310-315. <https://doi.org/10.1109/AiDAS56890.2022.9918731>
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3). <https://doi.org/10.5194/gmd-7-1247-2014>
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*. <https://doi.org/10.3115/v1/d14-1179>
- Choden, & Unhapipat, S. (2018). ARIMA model to forecast international tourist visit in Bumthang, Bhutan. *Journal of Physics Conference Series*, 1039(1). <https://doi.org/10.1088/1742-6596/1039/1/012023>
- Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Gated Recurrent Neural Networks on Sequence Modeling. *ArXiv*.

Goodfellow Ian, Bengio Yoshua, C. A. (2016). Deep Learning - Ian Goodfellow, Yoshua Bengio, Aaron Courville - Google Books. In *MIT Press*.

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10).
<https://doi.org/10.1109/TNNLS.2016.2582924>

Hair, F. J., Black C., W., Babin, J. B., & Anderson, E. R. (2019). Multivariate Data Analysis. *E-Jurnal Manajemen Unud*, 5(2).

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8). <https://doi.org/10.1162/neco.1997.9.8.1735>

Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting principles and practice, 3rd edition. *International Journal of Forecasting*, 22(4).

Ibrahim, M. S., Abbas, W., Waseem, M., Lu, C., Lee, H. H., Fan, J., & Loo, K. H. (2023). Long-Term Lifetime Prediction of Power MOSFET Devices Based on LSTM and GRU Algorithms. *Mathematics*, 11(15).
<https://doi.org/10.3390/math11153283>

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). Linear Regression BT - An Introduction to Statistical Learning with Applications in R. In *An Introduction to Statistical Learning*.

Khatri, B. B., Poudel, O., Chandra Kafl, S., & Simkhada, D. (2024). Trend Analysis of Tourist-arrivals in Nepal Using Auto-Regressive Integrated Moving Average (ARIMA) Model Pradeep Acharya. *Journal of Tourism & Adventure*, 7(1), 120-143.

Kulendran, N., & Witt, S. F. (2001). Cointegration versus least squares regression. *Annals of Tourism Research*, 28(2). [https://doi.org/10.1016/S0160-7383\(00\)00031-1](https://doi.org/10.1016/S0160-7383(00)00031-1)

- Kummong, R., & Homdok, S. (2019). Forecasting Foreign Tourists in Northern Thailand Using Artificial Neural Network. *Academic Journal of Science and Applied Science*, 2019(2), 15-31.
- Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting A deep learning approach. *Annals of Tourism Research*, 75, 410-423. <https://doi.org/10.1016/j.annals.2019.01.014>
- Lemmel, J., Babaiee, Z., Kleinlehner, M., Majic, I., Neubauer, P., Scholz, J., Grosu, R., & Neubauer, S. A. (2022). *Deep-Learning vs Regression Prediction of Tourism Flow with Limited Data*. <http://arxiv.org/abs/2206.13274>
- Limpornchitwilai, W., Suksompong, P., Charoenlarnnoppa, C., Kovavisaruch, L. O., Sanpechuda, T., & Nakatani, M. (2022). Improvement of the RF-Based and XGBoost-Based Visitor Data Prediction for Colocated Museums. *19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2022*. <https://doi.org/10.1109/ECTI-CON54298.2022.9795605>
- Liu, X., Chen, Y. N., Qiu, Z., & Chen, M. R. (2019). Forecast of the tourist volume of sanya city by XGBoost model and GM model. *Proceedings - 2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, CyberC 2019*, 166-173. <https://doi.org/10.1109/CyberC.2019.00038>
- Miles, J. (2014). R Squared, Adjusted R Squared. In *Wiley StatsRef Statistics Reference Online*. <https://doi.org/10.1002/9781118445112.stat06627>
- Wichern, D. W., Makridakis, S., & Wheelwright, S. C. (1979). Forecasting Methods and Applications. *Journal of the American Statistical Association*, 74(367), 733. <https://doi.org/10.2307/2287014>

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1).
<https://doi.org/10.3354/cr030079>



ภาคผนวก ก

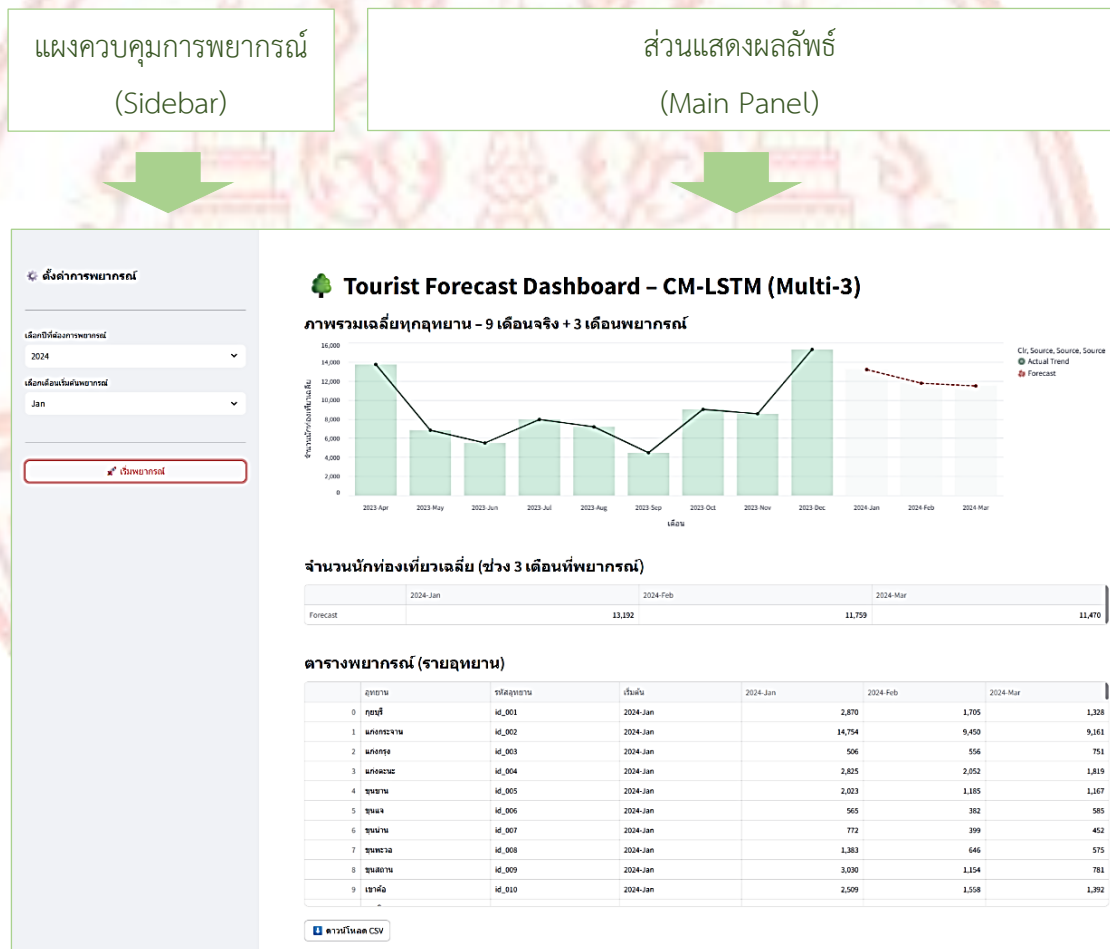
รายละเอียดการใช้งานเว็บแอปพลิเคชันต้นแบบสำหรับการพยากรณ์ (คู่มือ)

1. ส่วนประกอบหลักของหน้าเว็บแอปพลิเคชัน

ประกอบด้วย 2 ส่วนหลัก ได้แก่

1.1 แผงควบคุมการพยากรณ์ (Sidebar) ตั้งอยู่ทางด้านซ้ายของหน้าจอ ใช้สำหรับกำหนดพารามิเตอร์สำหรับการพยากรณ์

1.2 ส่วนแสดงผลลัพธ์ (Main Panel) พื้นที่หลักของหน้าจอสำหรับแสดงผลการพยากรณ์ในรูปแบบกราฟและตาราง

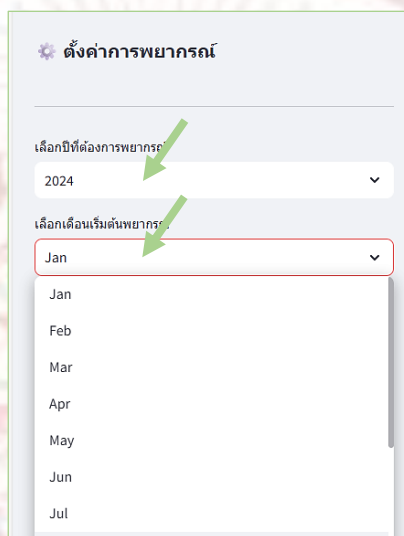


ภาพ ก - 1 ส่วนประกอบหลักของหน้าเว็บแอปพลิเคชัน

2. ขั้นตอนการใช้งาน

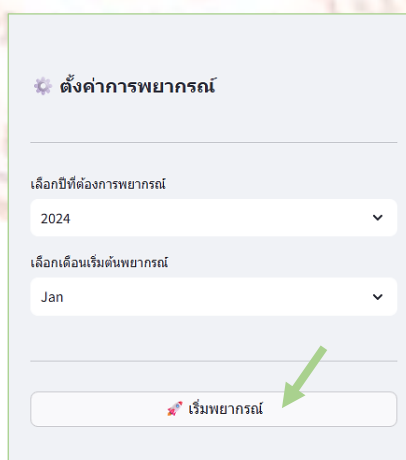
การใช้งานโปรแกรมประกอบด้วยขั้นตอนต่อไปนี้

2.1 ขั้นตอนที่ 1 กำหนดพารามิเตอร์การพยากรณ์ที่แมงควบคุมด้านซ้าย ผู้ใช้ต้องกำหนดค่า 2 ส่วน คือ เลือกปีที่ต้องการพยากรณ์ โดยการระบุปี พ.ศ. ที่ต้องการให้การพยากรณ์เริ่มต้น เลือกเดือนเริ่มต้นพยากรณ์ โดยการระบุเดือนที่จะเริ่มพยากรณ์ข้อมูล ซึ่งมีเงื่อนไขเดือนเริ่มต้นที่เลือกต้องเป็นเดือนถัดจากเดือนล่าสุดที่มีข้อมูลจริงในฐานข้อมูล ดังภาพ ก - 2



ภาพ ก - 2 กำหนดพารามิเตอร์การพยากรณ์

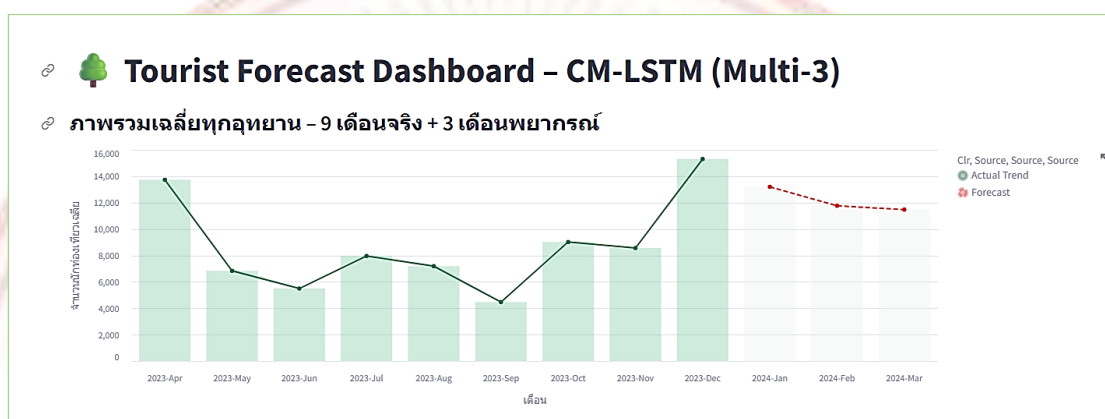
2.2 ขั้นตอนที่ 2 เริ่มกระบวนการพยากรณ์ หลังจากกำหนดค่าปีและเดือนเริ่มต้นแล้ว ผู้ใช้คลิกที่ปุ่ม เริ่มพยากรณ์ เพื่อเริ่มต้นกระบวนการ โปรแกรมจะดึงข้อมูลอนุกรมเวลาย้อนหลัง 9 เดือน สิ้นสุดก่อนเดือนที่ผู้ใช้เลือก เพื่อใช้เป็นข้อมูลนำเข้าสำหรับโมเดล CM-LSTM และสร้างผลการพยากรณ์ล่วงหน้า 3 เดือน ดังภาพ ก - 3



ภาพ ก - 3 เริ่มกระบวนการพยากรณ์

2.3 ขั้นตอนที่ 3 การแสดงผลลัพธ์ เมื่อกระบวนการพยากรณ์เสร็จสิ้น โปรแกรมจะแสดงผลลัพธ์ในส่วนแสดงผลด้านขวา ประกอบด้วย

- กราฟภาพรวมเฉลี่ยทุกอุทยาน กราฟผสมที่แสดงข้อมูล 12 เดือน (9 เดือนข้อมูลจริง และ 3 เดือนข้อมูลพยากรณ์) โดยแท่งกราฟสีเขียวอ่อน คือ ค่าเฉลี่ยจำนวนนักท่องเที่ยวจากข้อมูลจริงในอดีต, เส้นกราฟสีเขียวเข้ม (เส้นทึบ) คือการแสดงผลแนวโน้มของข้อมูลจริงในอดีต, เส้นกราฟสีแดง (เส้นประ) คือการแสดงผลแนวโน้มของข้อมูลพยากรณ์ล่วงหน้า ดังภาพ ก - 4



ภาพ ก - 4 กราฟภาพรวมเฉลี่ยทุกอุทยาน

- ตารางข้อมูลค่าเฉลี่ย เป็นตารางสรุปค่าพยากรณ์เฉลี่ยของทุกอุทยานสำหรับช่วงเวลา 3 เดือนที่พยากรณ์ ดังภาพ ก - 5

จำนวนนักท่องเที่ยวเฉลี่ย (ช่วง 3 เดือนที่พยากรณ์)

	2024-Jan	2024-Feb	2024-Mar
Forecast	13,192	11,759	11,470

ภาพ ก - 5 ตารางข้อมูลค่าเฉลี่ย

- ตารางพยากรณ์ (รายอุทยาน) ตารางแสดงผลการพยากรณ์จำนวนนักท่องเที่ยวจำแนกตามรายอุทยานและรายเดือน ดังภาพ ก - 6

ตารางพยากรณ์ (รายอุทยาน)

	อุทยาน	รหัสอุทยาน	เริ่มต้น	2024-Jan	2024-Feb	2024-Mar
0	กุยบุรี	id_001	2024-Jan	2,870	1,705	1,328
1	แก่งกระจาน	id_002	2024-Jan	14,754	9,450	9,161
2	แก่งกรุง	id_003	2024-Jan	506	556	751
3	แก่งตะนะ	id_004	2024-Jan	2,825	2,052	1,819
4	ขุนเขา	id_005	2024-Jan	2,023	1,185	1,167
5	ขุนแจ	id_006	2024-Jan	565	382	585
6	ขุนน้ำ	id_007	2024-Jan	772	399	452
7	ขุนพะวอ	id_008	2024-Jan	1,383	646	575
8	ขุนสถาน	id_009	2024-Jan	3,030	1,154	781
9	เขาค้อ	id_010	2024-Jan	2,509	1,558	1,392

ภาพ ก - 6 ตารางพยากรณ์ (รายอุทยาน)

3. การดาวน์โหลดข้อมูล

ผู้ใช้งานสามารถบันทึกข้อมูลผลลัพธ์จาก ตารางพยากรณ์ (รายอุทยาน) ในรูปแบบไฟล์ CSV โดยคลิกที่ปุ่มดาวน์โหลด CSV ดังภาพ ก - 7

4	ขุนขาน	id_005
5	ขุนแจ	id_006
6	ขุนน่าน	id_007
7	ขุนพะวอ	id_008
8	ขุนสถาน	id_009
9	เขาค้อ	id_010

ดาวน์โหลด CSV

ภาพ ก - 7 ดาวน์โหลดข้อมูล

4. ข้อความแจ้งเตือนและการแก้ไข

ระหว่างการใช้งานโปรแกรม อาจปรากฏข้อความแจ้งเตือนซึ่งมีสาเหตุและแนวทางแก้ไข ดังนี้

4.1 การแสดงข้อความ “ไม่สามารถพยากรณ์ได้ ข้อมูลล่าสุดที่มีคือ... แต่การพยากรณ์นี้ต้องการข้อมูลถึง...”

- สาเหตุ เดือนเริ่มต้นที่เลือกอยู่ห่างจากข้อมูลล่าสุดที่มีอยู่ เกินกว่าที่โปรแกรมจะหาข้อมูลย้อนหลัง 9 เดือนได้

- วิธีแก้ไข เลือกเดือนเริ่มต้นเป็นเดือนที่ถัดจากเดือนล่าสุดที่มีข้อมูลจริง

4.2 การแสดงข้อความ “ข้อมูลย้อนหลังไม่พอ”

- สาเหตุ ข้อมูลในระบบ ณ จุดเริ่มต้นของอนุกรมเวลา มีไม่เพียงพอที่จะนับย้อนหลังไปครบ 9 เดือน

- วิธีแก้ไข เลือกช่วงเวลาเริ่มต้นพยากรณ์ที่ใหม่ขึ้น

4.3 การแสดงข้อความ “ไม่สามารถสร้างการพยากรณ์ - ไม่มีอุทยานใดที่มีข้อมูลย้อนหลังครบ 9 เดือนตามเงื่อนไข”

- สาเหตุ ในช่วงเวลาที่เลือก ไม่มีข้อมูลของอุทยานใดที่มีความต่อเนื่องครบ 9 เดือน เพื่อใช้เป็นฐานในการพยากรณ์

- วิธีแก้ไข เปลี่ยนช่วงเวลาเริ่มต้นการพยากรณ์

ภาคผนวก ข

ลำดับรายชื่ออุทยานและตัวอย่างชุดข้อมูล

ภาคผนวกนี้แสดงรหัสอุทยานแห่งชาติที่ใช้ในงานวิจัย พร้อมชื่ออุทยานที่เกี่ยวข้อง รวมถึงตัวอย่างข้อมูลทั้ง 4 ประเภท ได้แก่ ข้อมูลจำนวนนักท่องเที่ยว ข้อมูลยานพาหนะ ข้อมูลสภาพอากาศ และข้อมูลลักษณะของอุทยาน เพื่อแสดงโครงสร้างและรูปแบบของข้อมูลที่ใช้ในการพัฒนาแบบจำลอง

ตาราง ข - 1 ข้อมูลรายชื่ออุทยานแห่งชาติ

id_National_Park	National_Park	Region_Park	Unique_Park
id_001	กุยบุรี	A_004	C_001
id_002	แก่งกระจาน	A_004	C_001
id_003	แก่งกรุง	A_005	C_001
id_004	แก่งตะนะ	A_002_001	C_002
id_005	ขุนขาน	A_001	C_001
id_006	ขุนแจ	A_001	C_001
id_007	ขุนน้ำ	A_001	C_001
id_008	ขุนพะวอ	A_001	C_001
id_009	ขุนสถาน	A_001	C_001
id_010	เขาค้อ	A_001	C_001
id_011	เขาคิชฌกูฏ	A_002	C_003
id_012	เขาชะเมา-เขาวง	A_002	C_001
id_013	เขานัน	A_005	C_001
id_014	เขาน้ำค้าง	A_005	C_001
id_015	เขาปู่-เขาย่า	A_005	C_001
id_016	เขาพนมเบญจา	A_005	C_001
id_017	เขาพระวิหาร	A_002_001	C_003
id_018	เขาลำปี-หาดท้ายเหมือง	A_005	C_004
id_019	เขาสก	A_005	C_002
id_020	เขาสาร้อยยอด	A_004	C_001
id_021	เขาสิบห้าชั้น	A_002	C_002
id_022	เขาหลวง	A_005	C_001

id_National_Park	National_Park	Region_Park	Unique_Park
id_023	เขาหลัก-ลำรู่	A_005	C_004
id_024	เขาแหลม	A_002	C_002
id_025	เขาแหลมหญ้า-หมู่เกาะเสม็ด	A_002	C_004
id_026	เขาใหญ่	A_002_001	C_001
id_027	เขื่อนศรีนครินทร์	A_004	C_002
id_028	คลองพนม	A_005	C_002
id_029	คลองลาน	A_001	C_002
id_030	คลองวังเจ้า	A_001	C_002
id_031	แจ้ซ้อน	A_001	C_002
id_032	เฉลิมพระเกียรติไทยประจัน	A_004	C_003
id_033	เฉลิมรัตนโกสินทร์	A_004	C_002
id_034	ดอยขุนตาล	A_001	C_001
id_035	ดอยจาง	A_001	C_001
id_036	ดอยผาแดง	A_001	C_001
id_037	ดอยผ้าห่มปก	A_001	C_001
id_038	ดอยภูคา	A_001	C_001
id_039	ดอยภูนาง	A_001	C_001
id_040	ดอยเวียงผา	A_001	C_001
id_041	ดอยสุเทพ-ปุย	A_001	C_003
id_042	ดอยหลวง	A_001	C_001
id_043	ดอยอินทนนท์	A_001	C_001
id_044	ต้นสักใหญ่	A_001	C_003
id_045	ตะรุเตา	A_005	C_004
id_046	ตากสินมหาราช	A_001	C_001
id_047	ตาดโตน	A_002_001	C_002
id_048	ตาดหมอก	A_001	C_002
id_049	ตาพระยา	A_002	C_001
id_050	ไผ่ล้อมเย็น	A_005	C_001
id_051	ถ้ำปลา-น้ำตกผาเสื่อ	A_001	C_002

id_National_Park	National_Park	Region_Park	Unique_Park
id_052	ถ้ำผาไท	A_001	C_001
id_053	ถ้ำสะเกิน	A_001	C_001
id_054	ทองผาภูมิ	A_004	C_001
id_055	ทะเลบัน	A_005	C_004
id_056	ทับลาน	A_002	C_001
id_057	ทุ่งแสลงหลวง	A_001	C_001
id_058	ไทรทอง	A_002_001	C_002
id_059	ไทรโยค	A_004	C_002
id_060	ธารโบกขรณี	A_005	C_002
id_061	ธารเสด็จ-เกาะพะงัน	A_005	C_004
id_062	นันทบุรี	A_001	C_001
id_063	นาเยื้อง-น้ำโสม	A_002_001	C_001
id_064	น้ำตกคลองแก้ว	A_002	C_002
id_065	น้ำตกเจ็ดสาวน้อย	A_003	C_002
id_066	น้ำตกชาติตระการ	A_001	C_002
id_067	น้ำตกซีโป	A_005	C_002
id_068	น้ำตกทรายขาว	A_005	C_002
id_069	น้ำตกพลี	A_002	C_002
id_070	น้ำตกพาเจริญ	A_001	C_002
id_071	น้ำตกแม่สุรินทร์	A_001	C_002
id_072	น้ำตกโยง	A_005	C_002
id_073	น้ำตกสามหลั่น	A_003	C_002
id_074	น้ำตกสี่ขีด	A_005	C_002
id_075	น้ำตกหงาว	A_005	C_002
id_076	น้ำตกห้วยยาง	A_004	C_002
id_077	น้ำพอง	A_002_001	C_002
id_078	น้ำหนาว	A_001	C_001
id_079	บางยาง	A_005	C_001
id_080	บูโด-สุโงปาดิ	A_005	C_001

id_National_Park	National_Park	Region_Park	Unique_Park
id_081	ปางสีดา	A_002	C_002
id_082	ป่าหินงาม	A_002_001	C_001
id_083	ผาแดง	A_001	C_001
id_084	ผาแต้ม	A_002_001	C_003
id_085	พุเตย	A_003	C_001
id_086	ภูกระดึง	A_002_001	C_001
id_087	ภูเก้า-ภูพานคำ	A_002_001	C_001
id_088	ภูจองนายอย	A_002_001	C_001
id_089	ภูซาง	A_001	C_002
id_090	ภูผาทิพย์	A_002_001	C_001
id_091	ภูผาม่าน	A_002_001	C_001
id_092	ภูผายล	A_002_001	C_001
id_093	ภูผาเหล็ก	A_002_001	C_001
id_094	ภูพาน	A_002_001	C_001
id_095	ภูเรือ	A_002_001	C_001
id_096	ภูลังกา	A_002_001	C_001
id_097	ภูแลนคา	A_002_001	C_001
id_098	ภูเวียง	A_002_001	C_001
id_099	ภูสระดอกบัว	A_002_001	C_002
id_100	ภูสวนทราย	A_002_001	C_001
id_101	ภูสอยดาว	A_001	C_002
id_102	ภูหินร่องกล้า	A_001	C_001
id_103	แม่เงา	A_001	C_002
id_104	แม่จรม	A_001	C_001
id_105	แม่ตะไคร้	A_001	C_001
id_106	แม่โถ	A_001	C_001
id_107	แม่ปิง	A_001	C_001
id_108	แม่ปืม	A_001	C_002
id_109	แม่เมย	A_001	C_001

id_National_Park	National_Park	Region_Park	Unique_Park
id_110	แม่ยม	A_001	C_002
id_111	แม่วังก์	A_001	C_001
id_112	แม่วะ	A_001	C_002
id_113	แม่วาง	A_001	C_002
id_114	รามคำแหง	A_001	C_001
id_115	ลานสาง	A_001	C_001
id_116	ลำคลองงู	A_004	C_001
id_117	ลำน้ำกก	A_001	C_001
id_118	ลำน้ำกระบุรี	A_005	C_001
id_119	ลำน้ำน่าน	A_001	C_002
id_120	เวียงโกศัย	A_001	C_001
id_121	ศรีน่าน	A_001	C_001
id_122	ศรีพังงา	A_005	C_004
id_123	ศรีลานนา	A_001	C_001
id_124	ศรีสัชนาลัย	A_001	C_003
id_125	สันกาลาศีรี	A_005	C_001
id_126	สาละวิน	A_001	C_001
id_127	สิรินาถ	A_005	C_004
id_128	หมู่เกาะช้าง	A_002	C_004
id_129	หมู่เกาะชุมพร	A_005	C_004
id_130	หมู่เกาะเกตรา	A_005	C_004
id_131	หมู่เกาะระนอง	A_005	C_004
id_132	หมู่เกาะลันตา	A_005	C_004
id_133	หมู่เกาะสิมิลัน	A_005	C_004
id_134	หมู่เกาะสุรินทร์	A_005	C_004
id_135	หมู่เกาะอ่างทอง	A_005	C_004
id_136	ห้วยน้ำดัง	A_001	C_002
id_137	หาดขนอม-หมู่เกาะทะเลใต้	A_005	C_004
id_138	หาดเจ้าไหม	A_005	C_004

id_National_Park	National_Park	Region_Park	Unique_Park
id_139	หาดนพรัตน์ธารา-หมู่เกาะพีพี	A_005	C_004
id_140	หาดวนกร	A_004	C_004
id_141	แหลมสน	A_005	C_004
id_142	ออบขาน	A_001	C_001
id_143	ออบหลวง	A_001	C_001
id_144	อ่าวพังงา	A_005	C_004
id_145	อ่าวมะนาว-เขาดันหยง	A_005	C_004
id_146	เอร์วีน	A_004	C_002

ตาราง ข - 2 คำอธิบายชุดข้อมูลจำนวนนักท่องเที่ยว

ชื่อคอลัมน์	คำอธิบาย	ตัวอย่างข้อมูล
No	ลำดับข้อมูล	2
id_National_Park	รหัสอุทยานแห่งชาติ	id_002
Year	ปี พ.ศ. ของข้อมูล	2014
Jan_Tourist	จำนวนนักท่องเที่ยวเดือน ม.ค.	19689
Feb_Tourist	จำนวนนักท่องเที่ยวเดือน ก.พ.	12124
Mar_Tourist	จำนวนนักท่องเที่ยวเดือน มี.ค.	11390
Apr_Tourist	จำนวนนักท่องเที่ยวเดือน เม.ย.	20905
May_Tourist	จำนวนนักท่องเที่ยวเดือน พ.ค.	17692
Jun_Tourist	จำนวนนักท่องเที่ยวเดือน มิ.ย.	7910
Jul_Tourist	จำนวนนักท่องเที่ยวเดือน ก.ค.	11429
Aug_Tourist	จำนวนนักท่องเที่ยวเดือน ส.ค.	10090
Sep_Tourist	จำนวนนักท่องเที่ยวเดือน ก.ย.	5190
Oct_Tourist	จำนวนนักท่องเที่ยวเดือน ต.ค.	4384
Nov_Tourist	จำนวนนักท่องเที่ยวเดือน พ.ย.	9235
Dec_Tourist	จำนวนนักท่องเที่ยวเดือน ธ.ค.	30324
Total_VisitorCount	จำนวนนักท่องเที่ยวรวมทั้งปี	160362

ตาราง ข - 3 คำอธิบายชุดข้อมูลยานพาหนะ

ชื่อคอลัมน์	คำอธิบาย	ตัวอย่างข้อมูล
No	-	2
id_National_Park	รหัสอุทยาน	id_002
Year	ปี พ.ศ. ของข้อมูล	2014
Jan_Vehicle	จำนวนยานพาหนะเดือน ม.ค.	4850
Feb_Vehicle	จำนวนยานพาหนะเดือน ก.พ.	3095
Mar_Vehicle	จำนวนยานพาหนะเดือน มี.ค.	2757
Apr_Vehicle	จำนวนยานพาหนะเดือน เม.ย.	4993
May_Vehicle	จำนวนยานพาหนะเดือน พ.ค.	4713
Jun_Vehicle	จำนวนยานพาหนะเดือน มิ.ย.	2105
Jul_Vehicle	จำนวนยานพาหนะเดือน ก.ค.	3021
Aug_Vehicle	จำนวนยานพาหนะเดือน ส.ค.	2259
Sep_Vehicle	จำนวนยานพาหนะเดือน ก.ย.	1460
Oct_Vehicle	จำนวนยานพาหนะเดือน ต.ค.	1055
Nov_Vehicle	จำนวนยานพาหนะเดือน พ.ย.	2063
Dec_Vehicle	จำนวนยานพาหนะเดือน ธ.ค.	7090
Total_VehicleCount	จำนวนยานพาหนะทั้งปี	39461

ตาราง ข - 4 คำอธิบายชุดข้อมูลสภาพอากาศ

ชื่อคอลัมน์	คำอธิบาย	ตัวอย่างข้อมูล
No	-	2
id_National_Park	รหัสอุทยาน	id_002
Year	ปี พ.ศ. ของข้อมูล	2014
Jan_TempMax	อุณหภูมิสูงสุดในเดือน ม.ค. (°C)	23
Jan_TempMin	อุณหภูมิต่ำสุดในเดือน ม.ค. (°C)	15

ชื่อคอลัมน์	คำอธิบาย	ตัวอย่างข้อมูล
Feb_TempMax	อุณหภูมิสูงสุดในเดือน ก.พ. (°C)	35
Feb_TempMin	อุณหภูมิต่ำสุดในเดือน ก.พ. (°C)	21
Mar_TempMax	อุณหภูมิสูงสุดในเดือน มี.ค. (°C)	36
Mar_TempMin	อุณหภูมิต่ำสุดในเดือน มี.ค. (°C)	22
Apr_TempMax	อุณหภูมิสูงสุดในเดือน เม.ย. (°C)	37
Apr_TempMin	อุณหภูมิต่ำสุดในเดือน เม.ย. (°C)	23
May_TempMax	อุณหภูมิสูงสุดในเดือน พ.ค. (°C)	37
May_TempMin	อุณหภูมิต่ำสุดในเดือน พ.ค. (°C)	24
Jun_TempMax	อุณหภูมิสูงสุดในเดือน มิ.ย. (°C)	37
Jun_TempMin	อุณหภูมิต่ำสุดในเดือน มิ.ย. (°C)	24
Jul_TempMax	อุณหภูมิสูงสุดในเดือน ก.ค. (°C)	35
Jul_TempMin	อุณหภูมิต่ำสุดในเดือน ก.ค. (°C)	24
Aug_TempMax	อุณหภูมิสูงสุดในเดือน ส.ค. (°C)	36
Aug_TempMin	อุณหภูมิต่ำสุดในเดือน ส.ค. (°C)	24
Sep_TempMax	อุณหภูมิสูงสุดในเดือน ก.ย. (°C)	35
Sep_TempMin	อุณหภูมิต่ำสุดในเดือน ก.ย. (°C)	24
Oct_TempMax	อุณหภูมิสูงสุดในเดือน ต.ค. (°C)	34
Oct_TempMin	อุณหภูมิต่ำสุดในเดือน ต.ค. (°C)	24
Nov_TempMax	อุณหภูมิสูงสุดในเดือน พ.ย. (°C)	34
Nov_TempMin	อุณหภูมิต่ำสุดในเดือน พ.ย. (°C)	22
Dec_TempMax	อุณหภูมิสูงสุดในเดือน ธ.ค. (°C)	34
Dec_TempMin	อุณหภูมิต่ำสุดในเดือน ธ.ค. (°C)	20
Avg_Temp	ค่าเฉลี่ยของอุณหภูมิ (°C)	28.33
Std_Temp	ส่วนเบี่ยงเบนมาตรฐานของอุณหภูมิ (°C)	6.84

ตาราง ข - 5 คำอธิบายชุดข้อมูลเชิงพื้นที่ (ภูมิภาคและลักษณะของอุทยาน)

ชื่อคอลัมน์	คำอธิบาย	ตัวอย่างข้อมูล
No	-	2
id_National_Park	รหัสอุทยาน	id_002
Year	ปี พ.ศ. ของข้อมูล	2014
Region_Park	รหัสภูมิภาคที่อุทยานตั้งอยู่	A_004
Unique_Park	รหัสกลุ่มลักษณะเฉพาะของอุทยาน	C_001



ประวัติผู้เขียน

ชื่อ	นางสาวณัฐชานันท์ ลายเงินมณี
ชื่อการค้นคว้าอิสระ	การประยุกต์ใช้โครงข่ายประสาทเทียมแบบเรียกซ้ำ สำหรับการพยากรณ์จำนวนนักท่องเที่ยวในอุทยานแห่งชาติ ของประเทศไทย The Application of Recurrent Neural Networks for Forecasting the Number of Tourists in Thailand's National Parks
สาขาวิชา	วิทยาการคอมพิวเตอร์
ประวัติ	สำเร็จการศึกษาระดับปริญญาตรี สาขาระบบสารสนเทศ ทางคอมพิวเตอร์ คณะบริหารธุรกิจ มหาวิทยาลัยเทคโนโลยี ราชมงคลธัญบุรี และระดับปริญญาโท สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ประยุกต์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้า พระนครเหนือ