



การพยากรณ์ฝุ่น PM2.5 ล่วงหน้าในจังหวัดเชียงใหม่
โดยใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว

FORECASTING PM2.5 IN CHIANGMAI
USING LONG SHORT-TERM MEMORY MODELS

ปรัชญา สิงหรวงศ์

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2563

การพยากรณ์ฝุ่น PM2.5 ล่วงหน้าในจังหวัดเชียงใหม่
โดยใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว



สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2563
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

FORECASTING PM2.5 IN CHIANGMAI
USING LONG SHORT-TERM MEMORY MODELS



A Master's Project Submitted in Partial Fulfillment of the Requirements
for the Degree of MASTER OF SCIENCE
(Information Technology)

Faculty of Science, Srinakharinwirot University

2020

Copyright of Srinakharinwirot University

สารนิพนธ์

เรื่อง

การพยากรณ์ฝุ่น PM2.5 ล่วงหน้าในจังหวัดเชียงใหม่

โดยใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว

ของ

ปรัชญา สิงหรวงศ์

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

คณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบปากเปล่าสารนิพนธ์

..... ที่ปรึกษาหลัก ประธาน
(ผู้ช่วยศาสตราจารย์ ดร.นุรีย์ วิวัฒน์วัฒนา) (อาจารย์ ดร.สุทธิพงศ์ รัชชพงษ์)

..... กรรมการ
(อาจารย์ ดร.โสภณ มงคลลักษณ์)

ชื่อเรื่อง	การพยากรณ์ฝุ่น PM2.5 ล่วงหน้าในจังหวัดเชียงใหม่ โดยใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว
ผู้วิจัย	ปรัชญา สิงหวรรณ
ปริญญา	วิทยาศาสตร์มหาบัณฑิต
ปีการศึกษา	2563
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. นุรีย์ วิวัฒน์วัฒนา

ในปัจจุบันปัญหาปริมาณฝุ่นละออง PM2.5 เป็นปัญหาทางสิ่งแวดล้อมที่เกิดขึ้นทั้งในประเทศและต่างประเทศโดยเฉพาะพื้นที่จังหวัดเชียงใหม่ของประเทศไทยที่มีค่าฝุ่นละอองเข้าสู่ชั้นวิกฤตเป็นอันตรายต่อสุขภาพ งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อสรรหาและเลือกใช้แบบจำลองการเรียนรู้ของเครื่องในการพยากรณ์ปริมาณฝุ่นละออง PM2.5 ในพื้นที่จังหวัดเชียงใหม่ล่วงหน้า 24 ชั่วโมง ในอันที่จะสามารถนำไปกำหนดมาตรการที่จะรับมือกับปัญหาฝุ่นละอองที่จะเกิดขึ้นได้ งานวิจัยได้ใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว Long Short-Term Memory (LSTM) ที่เป็นส่วนหนึ่งในเทคนิคการเรียนรู้เชิงลึก (Deep Learning) โดยชุดข้อมูลที่ใช้ได้จากการเก็บรวบรวมข้อมูลของกรมควบคุมมลพิษในจังหวัดเชียงใหม่แบบรายชั่วโมง ชุดข้อมูลที่ใช้มีจำนวน 2 ชุด ได้แก่ สถานีศูนย์ราชการจังหวัดเชียงใหม่ ระหว่างวันที่ 1 มกราคม ค.ศ. 2018 ถึง 31 พฤษภาคม ค.ศ. 2019 และสถานีโรงเรียนยุพราชวิทยาลัยจังหวัดเชียงใหม่ ระหว่างวันที่ 1 กรกฎาคม ค.ศ. 2018 ถึง 31 พฤษภาคม ค.ศ. 2019 การประเมินเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลอง คือ RMSE, MAE และ MAPE งานวิจัยได้ปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ได้แก่ Adam, RMSprop และ AdaGrad พบว่าชุดข้อมูลที่มีค่าข้อมูลสูญหายไม่มาก ควรเลือกใช้ Optimizer เป็น Adam และ AdaGrad แต่หากชุดข้อมูลที่มีค่าข้อมูลสูญหายมาก ควรเลือกใช้ Optimizer เป็น RMSprop จะทำให้สามารถพยากรณ์ฝุ่นละออง PM2.5 ล่วงหน้า 24 ชั่วโมงได้ดีที่สุด

คำสำคัญ : แบบจำลองหน่วยความจำระยะสั้นแบบยาว, การเรียนรู้เชิงลึก, ฝุ่นละออง PM2.5

Title	FORECASTING PM2.5 IN CHIANGMAI USING LONG SHORT-TERM MEMORY MODELS
Author	PRADYA SINGHAWORAWONG
Degree	MASTER OF SCIENCE
Academic Year	2020
Thesis Advisor	Assistant Professor Dr. Nuwee Wiwatwattana

PM2.5 has now become a major source of severe environmental air pollution both domestically and internationally, especially in Chiang Mai, where pollutant concentration levels are high, and are harmful to human health. The objective of this research is to select and fine-tune machine learning models for a 24-hour PM2.5 forecast in Chiang Mai that could help determine appropriate measures to cope with the haze problem in the future. The Long Short-Term Memory models (LSTM), one of the Deep Learning models, were trained using two sets of hourly air pollution data from the Pollution Control Department, Thailand. The first dataset was from the monitoring station at Chiang Mai Provincial Government Center dated between 1 January 2018 and 31 May 2019 and the second dataset was from the monitoring station at Yupparaj Wittayalai School, dated between 1 July 2018 and 31 May 2019. Three metrics, namely RMSE, MAE and MAPE, were used to evaluate the machine learning models. It was found that the models built for different stations gave different evaluation results. Furthermore, Adam, RMSprop and AdaGrad were three optimizers considered for fine-tuning. In terms of forecasting 24-hours ahead, the best results were obtained with Adam and AdaGrad. For the datasets with a lot of missing values, the best results were obtained with RMSprop.

Keyword : Long Short-Term Memory Models, Deep Learning, PM2.5

กิตติกรรมประกาศ

สารนิพนธ์นี้สำเร็จลุล่วงได้ด้วยความช่วยเหลือจาก ผศ.ดร.นุรีย์ วิวัฒน์วัฒนา อาจารย์ที่ปรึกษาที่ให้คำปรึกษา คำแนะนำในการทำสารนิพนธ์ ตลอดจนสนับสนุนข้อมูลทางวิชาการและข้อมูลสำหรับทำสารนิพนธ์นี้

ขอขอบคุณคณะกรรมการสอบวิทยานิพนธ์ได้แก่ ดร.สุทธิพงศ์ รัชชพงษ์ และ อาจารย์ ดร.โสภณ มงคลลักษณ์ ที่กรุณาเป็นประธานและกรรมการสอบสารนิพนธ์ และให้คำแนะนำ คำติชม สารนิพนธ์เล่มนี้ ซึ่งข้อติชมเหล่านั้นผู้จัดทำได้นำมาปรับปรุงสารนิพนธ์ฉบับนี้และเก็บไว้เป็นความรู้ในการศึกษาต่อไป

ขอขอบคุณอาจารย์และพนักงานประจำภาควิชาวิทยาการคอมพิวเตอร์ที่ให้ความรู้ คำแนะนำ และขอขอบพระคุณเจ้าหน้าที่ห้องธุรการประจำภาควิชาวิทยาการคอมพิวเตอร์ทุกท่านที่อำนวยความสะดวกด้านเอกสาร ตลอดจนคำปรึกษาในขั้นตอนดำเนินการจนสารนิพนธ์เล่มนี้เสร็จสมบูรณ์

ปรีชญา สิงหรวงศ์

สารบัญ

หน้า

บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ	ช
สารบัญตาราง.....	ญ
สารบัญรูปภาพ	ฎ
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและความเป็นมาของงานวิจัย.....	1
1.2 วัตถุประสงค์.....	3
1.3 ขอบเขตงานวิจัย	4
1.4 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย	5
บทที่ 2 ทบทวนวรรณกรรม.....	6
2.1 ข้อมูลพื้นฐานเกี่ยวกับความเข้มข้นฝุ่นละออง PM2.5	6
2.1.1 ความเข้มข้นฝุ่นละออง PM2.5	6
2.1.2 ระดับสีของค่าความเข้มข้น PM2.5	7
2.2 เทคนิคการเรียนรู้เชิงลึก (Deep Learning)	9
2.3 โครงสร้างเครือข่ายประสาทเทียม (Neural Network)	10
2.3.1 ส่วนประกอบของ Neural Network	11
2.4 แบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory)	12
2.4.1 Forget gate	14
2.4.2 Input gate และ Input Modulation gate	15

2.4.3 Update Cell State	17
2.4.4 Output Gate หรือ Read	18
2.5 ตัวปรับพารามิเตอร์แบบจำลองระยะสั้นแบบยาว LSTM (LSTM Optimizer)	19
2.5.1 Adaptive Gradient Descent (AdaGrad)	19
2.5.2 Root Mean Square Propagation (RMSprop)	20
2.5.3 Adaptive Moment Estimation (Adam)	20
2.6 ฟังก์ชันกระตุ้น (Activation Function).....	21
2.6.1 Sigmoid Function.....	21
2.6.2 Tanh function (Hyperbolic Tangent Activation Function)	22
2.7 ทฤษฎีที่เกี่ยวข้องกับการประเมินวัดความแม่นยำของแบบจำลอง	23
2.7.1 รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root mean square error: RMSE).....	24
2.7.2 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error: MAE).....	24
2.7.3 ค่าเฉลี่ยของร้อยละความคลาดเคลื่อนสัมบูรณ์ (Mean absolute percentage error: MAPE).....	25
2.8 งานวิจัยที่เกี่ยวข้อง	25
บทที่ 3 วิธีการดำเนินการวิจัย.....	29
3.1 การสร้างเครื่องมือที่ใช้การวิจัย	29
3.2 การสำรวจข้อมูล (Exploratory Data Analysis) ก่อนทำความสะอาดข้อมูล.....	30
3.3 การทำความสะอาดข้อมูล (Data Cleaning).....	33
3.4 การสำรวจข้อมูล (Exploratory Data Analysis) หลังทำความสะอาดข้อมูล	33
3.5 การสร้างแบบจำลอง Long Short-Term Memory (LSTM).....	44
3.5.1 การ Shift ข้อมูลย้อนหลัง.....	44
3.5.2 การแบ่งข้อมูล (Split data)	44

3.5.3 การแปลงข้อมูลโดยใช้ Min-Max Scaler	45
3.5.4 การปรับพารามิเตอร์ (Optimizing LSTM)	45
บทที่ 4 ผลการศึกษา	46
4.1 ผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วย การเลือกใช้ Optimizer ของสถานีศูนย์ราชการจังหวัดเชียงใหม่.....	46
4.2 ผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วย การเลือกใช้ Optimizer ของสถานีโรงเรียนยุพราชวิทยาลัย.....	48
บทที่ 5 สรุป อภิปรายผล และข้อเสนอแนะ.....	55
5.1 สรุปผลการวิจัย.....	55
5.2 อภิปรายผลการวิจัย	57
5.3 ข้อเสนอแนะ.....	58
บรรณานุกรม	60

สารบัญตาราง

หน้า

ตาราง 1 แสดงค่าความเข้มข้นของ PM2.5 (กรมควบคุมมลพิษ, 2563)	8
ตาราง 2 แสดงตารางเปรียบเทียบค่าเฉลี่ยของแต่ละตัวแปรของข้อมูลศูนย์ราชการจังหวัด เชียงใหม่และโรงเรียนยุพราชวิทยาลัย.....	35
ตาราง 3 แสดงผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมาก ขึ้นด้วยการเลือกใช้ Optimizer ของสถานีศูนย์ราชการจังหวัดเชียงใหม่.....	47
ตาราง 4 แสดงผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมาก ขึ้นด้วยการเลือกใช้ Optimizer ของสถานีโรงเรียนยุพราชวิทยาลัย.....	49
ตาราง 5 เปรียบเทียบการเลือก Optimizer ตามช่วงเวลาของข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	56
ตาราง 6 เปรียบเทียบการเลือก Optimizer ตามช่วงเวลาของข้อมูลโรงเรียนยุพราชวิทยาลัย.....	57

สารบัญรูปภาพ

หน้า

ภาพประกอบ 1 แสดงพื้นที่สถานีศูนย์ราชการจังหวัดเชียงใหม่และโรงเรียนยุพราชวิทยาลัย	4
ภาพประกอบ 2 แสดงขนาดฝุ่นละออง PM2.5 กับเส้นขน.....	7
ภาพประกอบ 3 แสดงระดับสีของค่าความเข้มข้น PM2.5.....	7
ภาพประกอบ 4 แสดงข้อแตกต่างระหว่าง Machine Learning และ Deep Learning	9
ภาพประกอบ 5 แสดงการเปรียบเทียบช่วงเวลาการเกิดของแต่ละเทคนิค	10
ภาพประกอบ 6 แสดงลักษณะทางชีววิทยาของโครงข่ายประสาท (สงวนสิทธิ์, 2562)	11
ภาพประกอบ 7 แสดงลักษณะโครงสร้าง Neural Network (Tanty & Desmukh, 2015)	12
ภาพประกอบ 8 แสดงแบบจำลอง Long Short-Term Memory (LSTM) (Fister & Jagric, 2019)	13
ภาพประกอบ 9 แสดง Forget gate.....	14
ภาพประกอบ 10 แสดง Input gate และ Input modulation gate	15
ภาพประกอบ 11 แสดง Update Cell State	17
ภาพประกอบ 12 แสดง Output gate	18
ภาพประกอบ 13 แสดง Sigmoid Function (ประกอบผล, 2552)	22
ภาพประกอบ 14 แสดง Tanh Function (ประกอบผล, 2552)	23
ภาพประกอบ 15 แสดงขั้นตอนการสร้างแบบจำลองการเรียนรู้ของเครื่อง	29
ภาพประกอบ 16 แสดงค่า NaN.....	30
ภาพประกอบ 17 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (ก่อนทำการทำความสะอาดข้อมูล)	30
ภาพประกอบ 18 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย (ก่อน ทำการทำความสะอาดข้อมูล)	31

ภาพประกอบ 19 แสดงกราฟ Boxplot (ก่อนทำความสะอาดข้อมูล) ของชุดข้อมูลศูนย์ราชการ จังหวัดเชียงใหม่	32
ภาพประกอบ 20 แสดงกราฟ Boxplot (ก่อนทำความสะอาดข้อมูล) ของชุดข้อมูลโรงเรียนยุพราช วิทยาลัย	32
ภาพประกอบ 21 แสดงตัวอย่างการแทนค่า NaN ด้วยค่าเฉลี่ย	33
ภาพประกอบ 22 แสดงการสำรวจข้อมูลสถิติเบื้องต้นชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	34
ภาพประกอบ 23 แสดงการสำรวจข้อมูลสถิติเบื้องต้นชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (ต่อ)	34
ภาพประกอบ 24 แสดงการสำรวจข้อมูลสถิติเบื้องต้นชุดข้อมูลโรงเรียนยุพราชวิทยาลัย	34
ภาพประกอบ 25 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (หลังทำความสะอาดข้อมูล)	36
ภาพประกอบ 26 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย (หลัง ทำความสะอาดข้อมูล)	36
ภาพประกอบ 27 แสดงกราฟ Boxplot ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (หลังทำความสะอาด ข้อมูล)	37
ภาพประกอบ 28 แสดงกราฟ Boxplot ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย (หลังทำความสะอาด ข้อมูล)	37
ภาพประกอบ 29 แสดงค่า Correlation ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	38
ภาพประกอบ 30 แสดงค่า Correlation ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย	39
ภาพประกอบ 31 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ t-168 ของชุด ข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	40
ภาพประกอบ 32 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ t-1 ของชุด ข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	41
ภาพประกอบ 33 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ t-168 ของชุด ข้อมูลโรงเรียนยุพราชวิทยาลัย	42

ภาพประกอบ 34 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ $t-1$ ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย	43
ภาพประกอบ 35 ตัวอย่างการ Shift ย้อนหลัง lag 7 วัน หรือ 168 ชั่วโมง	44
ภาพประกอบ 36 กราฟแสดงค่า RMSE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	50
ภาพประกอบ 37 กราฟแสดงค่า MAE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	51
ภาพประกอบ 38 กราฟแสดงค่า MAPE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่	52
ภาพประกอบ 39 กราฟแสดงค่า RMSE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย	52
ภาพประกอบ 40 กราฟแสดงค่า MAE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย	53
ภาพประกอบ 41 กราฟแสดงค่า MAPE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย	54

บทที่ 1

บทนำ

1.1 ความสำคัญและความเป็นมาของงานวิจัย

มลพิษทางอากาศเป็นปัญหาทางสิ่งแวดล้อมที่เกิดขึ้นทั้งในประเทศและต่างประเทศทั่วโลก ปัญหาเหล่านี้เกิดจากพฤติกรรม กิจกรรมต่าง ๆ ของมนุษย์ในชีวิตประจำวันที่ทำให้เกิดมลพิษทางอากาศที่เป็นอันตรายต่อสุขภาพ โดยปัจจัยที่สำคัญที่ทำให้เกิดปัญหามลพิษทางอากาศ ได้แก่ การเผาในที่โล่ง การคมนาคมขนส่ง การผลิตไฟฟ้า การทำอุตสาหกรรม และการรวมตัวของก๊าซอื่น ๆ ในบรรยากาศ เป็นต้น ด้วยปัจจัยที่ก่อให้เกิดปัญหาเหล่านี้ ทำให้แต่ละประเทศทั่วโลกได้กำหนดมาตรการในการแก้ไขปัญหาหมอกควันมลพิษทางอากาศที่เกิดขึ้น สาเหตุหลักที่ทำให้เกิดปัญหามลพิษทางอากาศ คือ ความเข้มข้นฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน หรือความเข้มข้นฝุ่นละออง PM2.5 ที่เป็นอันตรายต่อสุขภาพ เนื่องจากมีขนาดที่ค่อนข้างเล็กเปรียบเทียบได้กับขนาด 1 ใน 25 ของเส้นผ่านศูนย์กลางของเส้นผมมนุษย์ ทำให้คนจมูกไม่สามารถกรองได้ โดยในปัจจุบันปริมาณความเข้มข้นฝุ่นละออง PM2.5 ที่เพิ่มมากขึ้นในแต่ละวัน และเกิดขึ้นในทั่วโลก เข้าสู่ชั้นบรรยากาศได้เกินค่ามาตรฐาน โดยเฉพาะในประเทศไทย จังหวัดเชียงใหม่เป็นจังหวัดที่ติดอันดับพื้นที่ที่มีค่าความเข้มข้นฝุ่นละออง PM2.5 เข้าสู่ชั้นบรรยากาศได้เกินค่ามาตรฐานต่อเนื่องนานนับเดือนและเป็นอันตรายต่อคนในพื้นที่จังหวัดเชียงใหม่ เนื่องจากจังหวัดเชียงใหม่มีปัญหาเผาไหม้เป็นปัญหาหลักที่มีมาอย่างต่อเนื่องทำให้เกิดปัญหาความเข้มข้นฝุ่นละอองต่อเนื่องมา เมื่อฝุ่นละอองมีปริมาณมากหากแพร่เข้าสู่ทางเดินหายใจ กระแสเลือด และสามารถซึมผ่านอวัยวะต่าง ๆ ของร่างกายได้ ทำให้คนในพื้นที่เกิดปัญหาส่งผลเสียต่อร่างกายในระยะยาว

นอกจากปัญหาทางสุขภาพที่เกิดขึ้นจากมลพิษทางอากาศ ความเข้มข้นฝุ่นละออง PM2.5 แล้ว ปัญหาที่ตามมา คือ ปัญหาการท่องเที่ยวภายในประเทศไทย จังหวัดเชียงใหม่เป็นจังหวัดที่มีการท่องเที่ยวที่มากของประเทศไทย ฝุ่นละอองที่เกิดขึ้นในอากาศทำให้ท้องฟ้าเป็นสีฝุ่นเทา ทำให้ทัศนียภาพของจังหวัดเชียงใหม่ไม่สวยงามเหมือนแต่ดั้งเดิมในอดีต ส่งผลให้นักท่องเที่ยวภายในประเทศ และต่างประเทศเดินทางมาท่องเที่ยวในจังหวัดเชียงใหม่ลดลง จากข้อมูลนางสาวภคินันท์ วิณิชัย ผู้อำนวยการการท่องเที่ยวแห่งประเทศไทย (ททท.) สำนักงานเชียงใหม่เปิดเผยข้อมูลว่าในช่วงเดือนมกราคมถึงเดือนกรกฎาคม ปี พ.ศ. 2562 จังหวัดเชียงใหม่มีจำนวนนักท่องเที่ยว ประมาณ 5.7 ล้านคนต่อครั้ง ซึ่งมีจำนวนลดลงจากช่วงเดือนเดียวกันของปี พ.ศ. 2561 ประมาณร้อยละ 1.8 โดยจำนวนนักท่องเที่ยวที่ลดลงมีสาเหตุหลักที่เกิดขึ้นจากปัญหา

ความเข้มข้นฝุ่นละออง PM2.5 โดยในเดือนมีนาคม ลดลงไปถึงร้อยละ 12.21 และเดือนเมษายน ลดลงร้อยละ 6.46% (ประชาชาติธุรกิจ, 2562)

จากปัญหาปริมาณความเข้มข้นฝุ่นละออง PM2.5 ที่เข้าสู่ชั้นวิกฤตบริเวณจังหวัดเชียงใหม่ของประเทศไทย ทำให้ผู้วิจัยเห็นความสำคัญของการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 ในพื้นที่จังหวัดเชียงใหม่ล่วงหน้า 24 ชั่วโมง เพื่อให้ทราบข้อมูลล่วงหน้าก่อนที่จะเกิดขึ้นได้ ทำให้สามารถรับมือกับปัญหาฝุ่นละออง PM2.5 ที่เกิดขึ้นได้ สามารถกำหนดมาตรการ หรือวางแผนนโยบายก่อนฝุ่นละออง PM2.5 ล่วงหน้าได้ ทางผู้วิจัยได้ทำการศึกษาเทคนิคการเรียนรู้เชิงลึก (Deep Learning) ที่ใช้พยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 โดยเทคนิคการเรียนรู้เชิงลึกนี้เป็นส่วนหนึ่งของเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เทคนิคการเรียนรู้เชิงลึกเป็นเทคนิคที่ได้รับความนิยมเป็นอย่างมากในช่วง 10 ปีหลังมานี้ เนื่องจากเป็นเทคโนโลยีใหม่ที่มีนักวิจัย นักพัฒนาสนใจเป็นอย่างมาก เทคนิคการเรียนรู้เชิงลึกเป็นเทคโนโลยีที่มีประสิทธิภาพดีกว่าเทคนิคการเรียนรู้ของเครื่อง เนื่องจากเป็นเทคนิคที่สามารถทำงานได้ดีกับข้อมูลจำนวนมาก มีการประมวลผลได้ดี และเทคนิคนี้ยังมีคุณสมบัติเด่น คือ สามารถทำการแยกคุณลักษณะของข้อมูล (Feature Extraction) ได้ด้วยตัวเอง ทำให้สามารถแบ่งแยกคุณลักษณะที่เหมาะสมกับแบบจำลองได้ดี แตกต่างจากจากเทคนิคการเรียนรู้ของเครื่องที่มีความจำเป็นที่ต้องทำการแยกคุณลักษณะก่อนที่จะนำข้อมูลเข้าในแบบจำลอง ทำให้เทคนิคการเรียนรู้เชิงลึกทำงานได้ดีกว่าเทคนิคการเรียนรู้ของเครื่อง จึงเหมาะสมสำหรับการนำไปใช้ในแบบจำลองที่เป็นอนุกรมเวลา (Time series) ดังนั้น จึงเหมาะสำหรับการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 ในพื้นที่จังหวัดเชียงใหม่ล่วงหน้า 24 ชั่วโมง

จังหวัดเชียงใหม่เป็นจังหวัดที่มีพื้นที่ต่ำบด และอากาศค่อนข้างมากทำให้มีการเก็บข้อมูลจากกรมควบคุมมลพิษในจังหวัดเชียงใหม่อยู่หลายสถานี ทางผู้วิจัยจึงทำการเลือกชุดของข้อมูลเพื่อมาใช้สำหรับพยากรณ์ฝุ่นละอองในจังหวัดเชียงใหม่ แต่เนื่องจากการพยากรณ์เพียงหนึ่งสถานี อาจทำให้ไม่สามารถพิสูจน์การพยากรณ์ของแบบจำลองได้ดี ทำให้การเลือกชุดข้อมูลมากกว่าหนึ่งสถานีจะทำให้สามารถเปรียบเทียบข้อมูลแบบจำลองได้ดีกว่า ทางผู้วิจัยจึงได้ทำการศึกษาชุดข้อมูลจาก 2 สถานี เพื่อทำการศึกษาและเปรียบเทียบการพยากรณ์ล่วงหน้าในจังหวัดเชียงใหม่ล่วงหน้า 24 ชั่วโมง โดยสถานีที่เลือกมาพยากรณ์ คือ สถานีโรงเรียนยุพราชวิทยาลัย และสถานีศูนย์ราชการจังหวัดเชียงใหม่

เทคนิคการเรียนรู้เชิงลึก หรือ เครือข่ายประสาทเทียมเชิงลึก (Deep Neural Network) เป็นแบบจำลองที่ถูกสร้างขึ้นโดยมีลักษณะโครงสร้างและการทำงานคล้ายคลึงกับสมองของมนุษย์ ซึ่งโครงสร้างเครือข่ายประสาทเทียมที่สร้างมาจากการลอกเลียนแบบมาจากลักษณะทางชีววิทยาของสมองมนุษย์ โดยเครือข่ายประสาทเทียมมีหลายประเภท ทางผู้วิจัยได้เลือกมาใช้คือ แบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory) เป็นแบบจำลองที่พัฒนามาจากเครือข่ายประสาทเทียมตัวเดิม คือ Recurrent Neural Network (RNN) โดยที่แบบจำลองหน่วยความจำระยะสั้นแบบยาวจะเข้ามาแก้ไขปัญหา Vanishing Gradient และปัญหา Exploding Gradient ของ RNN และแบบจำลองหน่วยความจำระยะสั้นแบบยาวสามารถจดจำ (Memory) สถานะของแต่ละโหนดทำให้สามารถย้อนกลับไปหาค่าที่แท้จริงของเดิมได้ ยังมีประตู (Gate) ที่ทำหน้าที่ในการตัดสินใจว่าจะให้ข้อมูลนั้นผ่านไปได้หรือไม่

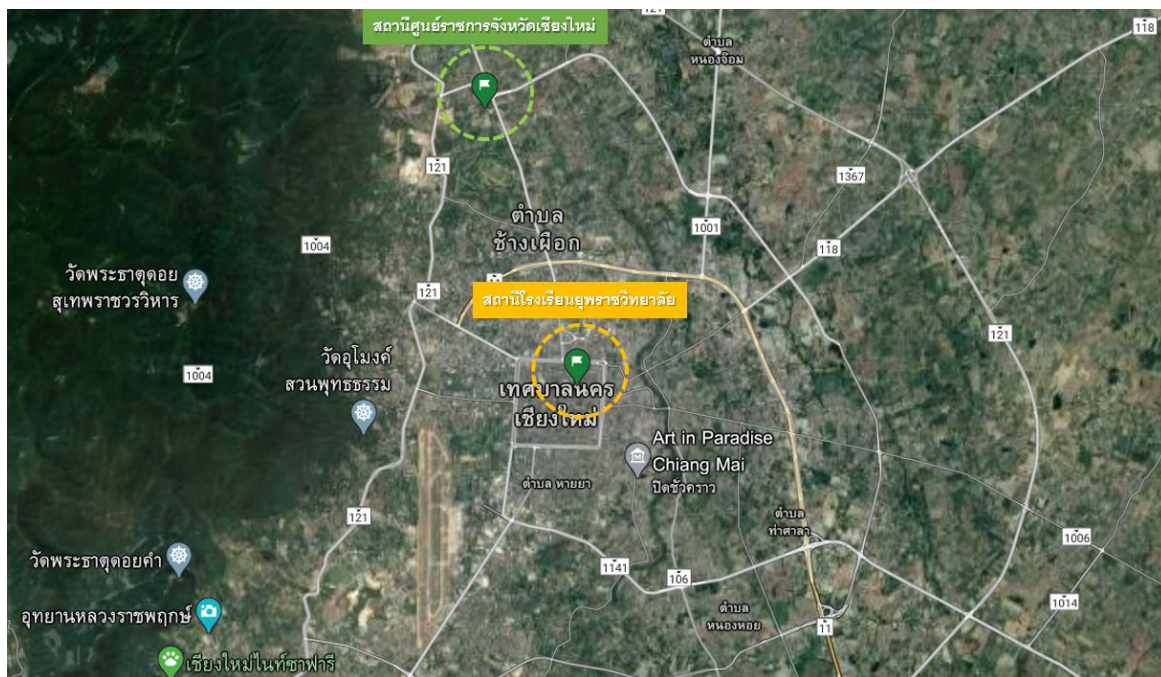
เทคนิคการเรียนรู้เชิงลึกสามารถทำงานได้อย่างมีประสิทธิภาพมากขึ้น หากทำการปรับค่า Optimizer ของแบบจำลอง เพื่อให้แบบจำลองพยากรณ์ฝุ่นละออง PM2.5 ล่วงหน้า 24 ชั่วโมงในจังหวัดเชียงใหม่มีประสิทธิภาพการทำงานที่ดีขึ้นดังกล่าว ทางผู้วิจัยจึงต้องการทดสอบความแตกต่างของ Optimizer สามประเภท โดยได้แก่ Adagrad, RMSprop และ Adam เนื่องจาก Optimizer แต่ละตัวถูกพัฒนามาเพื่อให้เทคนิคการเรียนรู้เชิงลึกสามารถพยากรณ์ได้มีประสิทธิภาพสูงสุด แต่ยังไม่มีการวิจัยที่พิสูจน์ได้ชัดเจนว่า Optimizer ตัวไหนเหมาะสมที่สุด เพื่อให้แบบจำลองสามารถนำไปใช้ในการพยากรณ์ได้อย่างมีประสิทธิภาพแล้ว ดังนั้น การปรับค่า Optimizer ก็อาจจะทำให้มีผลกระทบต่อผลลัพธ์ของการพยากรณ์ ผู้วิจัยจึงต้องการทดสอบเพื่อให้ผลลัพธ์ที่ดีที่สุด และสามารถพยากรณ์ฝุ่นละออง PM2.5 ล่วงหน้าในจังหวัดเชียงใหม่ 24 ชั่วโมงได้อย่างแม่นยำ

1.2 วัตถุประสงค์

1. เพื่อสร้างแบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory) ในการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 จากข้อมูลสถานีในจังหวัดเชียงใหม่
2. เพื่อเปรียบเทียบความแตกต่างของแบบจำลองในการพยากรณ์ระหว่างชุดข้อมูลสถานีศูนย์ราชการจังหวัดเชียงใหม่ และชุดข้อมูลสถานีโรงเรียนยุพราชวิทยาลัย
3. เพื่อเปรียบเทียบประสิทธิภาพแบบจำลองในการพยากรณ์ให้แม่นยำด้วยการปรับ Optimizer ที่เกี่ยวข้องกับแบบจำลองที่แตกต่างกัน

1.3 ขอบเขตงานวิจัย

งานวิจัยได้ใช้ชุดข้อมูล (Dataset) ที่ได้จากการเก็บรวบรวมข้อมูลจากกรมควบคุมมลพิษ ในจังหวัดเชียงใหม่ โดยเก็บข้อมูลในลักษณะเป็นรายชั่วโมง ข้อมูลที่ใช้เป็นข้อมูลจังหวัดเชียงใหม่ จาก 2 สถานีตรวจวัด ได้แก่ ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ และชุดข้อมูลโรงเรียนยุพราชวิทยาลัย



ภาพประกอบ 1 แสดงพื้นที่สถานีศูนย์ราชการจังหวัดเชียงใหม่และโรงเรียนยุพราชวิทยาลัย

ที่มา : <https://www.google.co.th/maps/>

1. ชุดข้อมูลสถานีศูนย์ราชการจังหวัดเชียงใหม่ ข้อมูลที่ใช้ระหว่างวันที่ 1 มกราคม 2018 0:00 ถึง 31 พฤษภาคม 2019 23:00 ได้แก่ NO (ไนตริกออกไซด์), NOX (ไนโตรเจนออกไซด์), NO2 (ไนโตรเจนไดออกไซด์), SO2 (ซัลเฟอร์ไดออกไซด์), O3 (โอโซน), PM10, Wind dir (ทิศทางของลม), Temp (อุณหภูมิ), Rel hum (ความชื้นสัมพัทธ์), Rain (ปริมาณฝน), Wind speed (ความเร็วลม), PM2.5 จำนวน 12 ตัวแปร

2. ชุดข้อมูลสถานีโรงเรียนยุพราชวิทยาลัย ข้อมูลที่ใช้ระหว่างวันที่ 1 กรกฎาคม 2018 0:00 ถึง 31 พฤษภาคม 2019 23:00 ได้แก่ CO (คาร์บอนมอนนอกไซด์), NO (ไนตริกออกไซด์), NOX (ไนโตรเจนออกไซด์), NO2 (ไนโตรเจนไดออกไซด์), SO2 (ซัลเฟอร์ไดออกไซด์), PM10,

Wind dir (ทิศทางของลม), Temp (อุณหภูมิ), Rel hum (ความชื้นสัมพัทธ์), Wind speed (ความเร็วลม), PM2.5 จำนวน 10 ตัวแปร

เพื่อสร้างแบบจำลองในการทำนายฝุ่น PM2.5 โดยใช้แบบจำลอง Long Short-Term Memory (LSTM) และใช้ทฤษฎีในการประเมินความแม่นยำแบบจำลองโดยใช้เทคนิคการหาค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error: RMSE), ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) และค่าเฉลี่ยของร้อยละความคลาดเคลื่อนสัมบูรณ์ (Mean absolute percentage error: MAPE)

1.4. ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย

1. ทำให้สามารถทำนายปริมาณฝุ่นละออง PM2.5 ล่วงหน้า 24 ชั่วโมงได้อย่างแม่นยำ
2. ทำให้สามารถนำแบบจำลองไปปรับใช้กับข้อมูลในจังหวัดอื่น ๆ ได้
3. สามารถนำแบบจำลองไปใช้ในการกำหนดมาตรการในการแก้ไขปัญหาฝุ่นละอองได้

บทที่ 2

ทบทวนวรรณกรรม

ในการวิจัยครั้งนี้ ผู้วิจัยได้ทำการศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง และได้นำเสนอตามหัวข้อต่อไปนี้

1. ข้อมูลพื้นฐานเกี่ยวกับความเข้มข้นฝุ่นละออง PM2.5
2. เทคนิคการเรียนรู้เชิงลึก (Deep Learning)
3. โครงสร้างเครือข่ายประสาทเทียม (Neural Network)
4. แบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory)
5. ตัวปรับพารามิเตอร์แบบจำลองระยะสั้นแบบยาว LSTM (LSTM Optimizer)
6. ฟังก์ชันกระตุ้น (Activation Function)
7. ทฤษฎีที่เกี่ยวข้องกับการประเมินวัดความแม่นยำของแบบจำลอง
8. งานวิจัยที่เกี่ยวข้อง

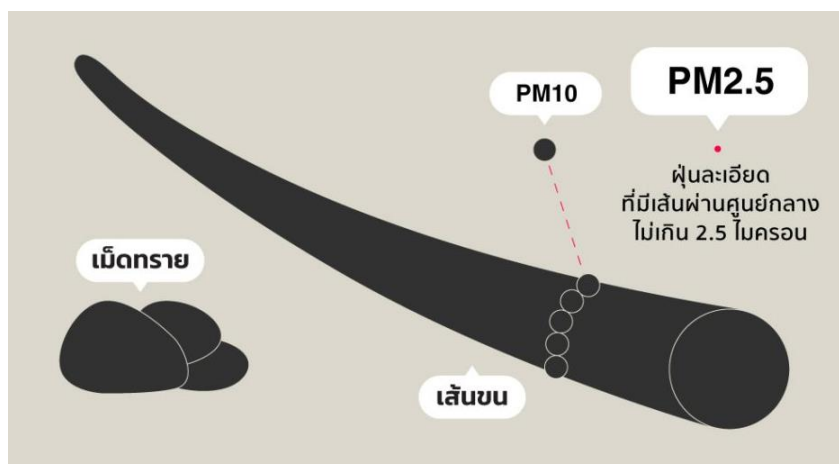
2.1 ข้อมูลพื้นฐานเกี่ยวกับความเข้มข้นฝุ่นละออง PM2.5

2.1.1 ความเข้มข้นฝุ่นละออง PM2.5

คำว่า PM ย่อมาจาก Particulate Matter เป็นคำเรียกค่ามาตรฐานของฝุ่นละอองขนาดเล็กที่เป็นอันตรายต่อสุขภาพซึ่งมีอยู่ด้วยกัน 2 ชนิดได้ PM10 และ PM2.5 ซึ่งตัวเลข 10 และ 2.5 นั้นหมายถึงหน่วยไมโครกรัมต่อลูกบาศก์เมตร

PM2.5 คือ ฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมโครกรัมต่อลูกบาศก์เมตร เปรียบเทียบได้กับขนาดประมาณ 1 ใน 25 ของเส้นผ่านศูนย์กลางของเส้นผมมนุษย์ มีขนาดเล็กที่ขมุกขมูยของมนุษย์นั้นไม่สามารถกรองได้ ทำให้ฝุ่นละอองแพร่กระจายเข้าสู่ทางเดินหายใจกระแสน้ำเลือดและแทรกซึมสู่กระบวนการทำงานในอวัยวะต่าง ๆ ของร่างกาย เพิ่มความเสี่ยงในการเป็นโรคต่าง ๆ อีก

ที่มาของฝุ่นละออง PM2.5 เกิดจากปัจจัย 2 ปัจจัยหลัก เกิดจากแหล่งกำเนิดโดยตรง เช่น การเผาในที่โล่ง ปล่อย PM2.5 มากที่สุดถึง 209,937 ตันต่อปี, การคมนาคมขนส่ง ปล่อย PM2.5 รวม 50,240 ตันต่อปี, การผลิตไฟฟ้า ปล่อย PM2.5 รวม 31,793 ตันต่อปี, อุตสาหกรรมการผลิต ปล่อย PM2.5 รวม 65,140 ตันต่อปี และการรวมตัวของก๊าซอื่น ๆ ในบรรยากาศ เช่น ซัลเฟอร์ไดออกไซด์ (SO₂) และไนโตรเจนออกไซด์ (NO_x) รวมทั้งสารอื่น ๆ ที่เป็นสารพิษที่เป็นอันตรายต่อร่างกายมนุษย์ (Standard, 2561)



ภาพประกอบ 2 แสดงขนาดฝุ่นละออง PM2.5 กับเส้นขน

ที่มา : ประชุมพันธ์,จ. (2018). PM2.5 ฝุ่นละอองขนาดเล็กในอากาศ กับวิกฤตสุขภาพที่คนไทยจะต้องแลก Retrieved from <https://thestandard.co/pm-2-5-environmental-nano-pollutants/>

2.1.2 ระดับสีของค่าความเข้มข้น PM2.5

ระดับสีของค่าความเข้มข้น PM2.5 คือ ตั้งแต่ 0-91 ขึ้นไป ซึ่งแต่ละระดับจะใช้สีเป็นสัญลักษณ์เปรียบเทียบกับระดับของผลกระทบต่อสุขภาพอนามัย ตามมาตรฐานกรมควบคุมมลพิษ (กรมควบคุมมลพิษ, 2563)



ภาพประกอบ 3 แสดงระดับสีของค่าความเข้มข้น PM2.5

ตาราง 1 แสดงค่าความเข้มข้นของ PM2.5 (กรมควบคุมมลพิษ, 2563)

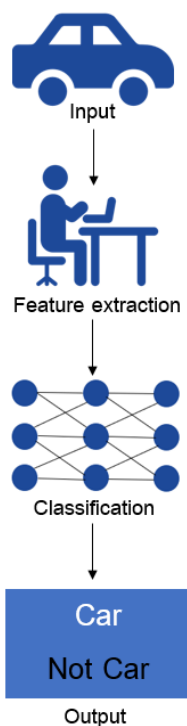
PM2.5 ($\mu\text{g}/\text{m}^3$)	ความหมาย	สีที่ใช้	คำอธิบาย
0-25	คุณภาพอากาศดีมาก	ฟ้า	คุณภาพอากาศดีเหมาะสำหรับกิจกรรมกลางแจ้งและการท่องเที่ยว
26-37	คุณภาพอากาศดี	เขียว	คุณภาพอากาศดี สามารถทำกิจกรรมกลางแจ้งและการท่องเที่ยวได้ตามปกติ
38-50	ปานกลาง	เหลือง	ประชาชนทั่วไป: สามารถทำกิจกรรมกลางแจ้งได้ตามปกติ ผู้ที่ต้องดูแลสุขภาพเป็นพิเศษ: หากมีอาการ เช่น ไอ หายใจลำบาก ระคายเคืองตา ควรลดระยะเวลาทำกิจกรรมกลางแจ้ง
51-90	เริ่มมีผลกระทบต่อสุขภาพ	ส้ม	ประชาชนทั่วไป: ควรพิจารณาถึงสุขภาพ ถ้ามีอาการไอ หายใจลำบาก ระคายเคืองตา ควรลดระยะเวลาทำกิจกรรมกลางแจ้งหรือใช้อุปกรณ์ป้องกันตนเองหากมีความจำเป็น ผู้ที่ต้องดูแลสุขภาพเป็นพิเศษ: ควรลดระยะเวลาการทำกิจกรรมกลางแจ้ง หรือใช้อุปกรณ์ป้องกันหากมีความจำเป็น ถ้ามีอาการทางสุขภาพ เช่น ไอ หายใจลำบาก ตาอักเสบ แสบหน้าอก ปวดศีรษะ หัวใจเต้นไม่ปกติ คลื่นไส้ อ่อนเพลีย ควรปรึกษาแพทย์
91 ขึ้นไป	มีผลกระทบต่อสุขภาพ	แดง	ทุกคนควรหลีกเลี่ยงกิจกรรมกลางแจ้ง และพื้นที่ที่มีมลพิษทางอากาศสูงหรือใช้อุปกรณ์ป้องกันตนเองหากมีความจำเป็น หากมีอาการทางสุขภาพของปรึกษาแพทย์

2.2 เทคนิคการเรียนรู้เชิงลึก (Deep Learning)

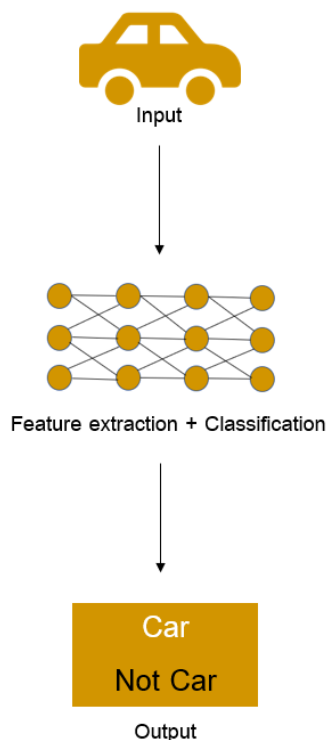
เทคนิคการเรียนรู้เชิงลึก (Deep Learning) เป็นการพัฒนาระบบปัญญาประดิษฐ์ (AI: Artificial Intelligence) ให้แบบจำลองหรือเครื่องจักรสามารถเรียนรู้ สร้างองค์ความรู้ใหม่ (Knowledge) และสามารถพยากรณ์สิ่งที่จะเกิดขึ้นในอนาคตมาได้

เทคนิคการเรียนรู้เชิงลึก เป็นส่วนหนึ่งของเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยข้อแตกต่างระหว่างการเรียนรู้ของเครื่องและเทคนิคการเรียนรู้เชิงลึก คือ การทำ Feature extraction ที่เทคนิคการเรียนรู้เชิงลึกสามารถทำได้ด้วยตนเอง แต่สำหรับเทคนิคการเรียนรู้ของเครื่องจะต้องทำวิธี Feature extraction ก่อนจะนำไปเข้าในแบบจำลอง ซึ่งถือว่าเป็นจุดเด่นของเทคนิคการเรียนรู้เชิงลึก แต่ข้อเสียของเทคนิคการเรียนรู้เชิงลึก คือ จำเป็นต้องใช้ทรัพยากรของเครื่องคอมพิวเตอร์มาก หากมีจำนวน Neural network มีโครงสร้างสถาปัตยกรรมขนาดใหญ่มากก็จำเป็นต้องใช้ข้อมูลทรัพยากรของเครื่องคอมพิวเตอร์มากเช่นกัน

Machine Learning

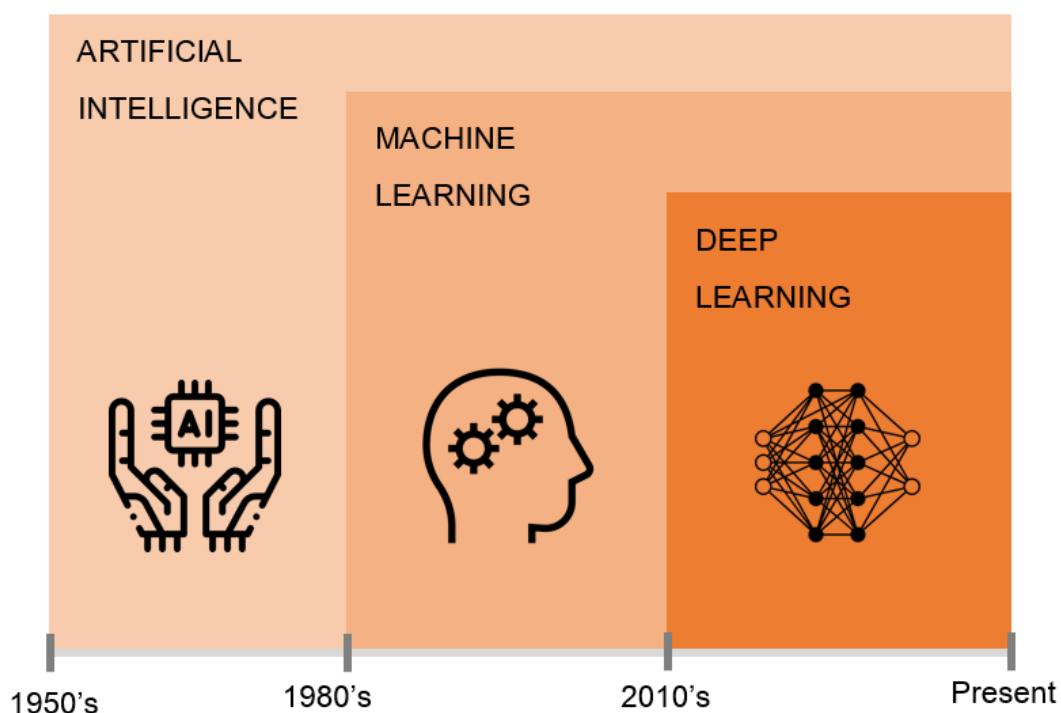


Deep Learning



ภาพประกอบ 4 แสดงข้อแตกต่างระหว่าง Machine Learning และ Deep Learning

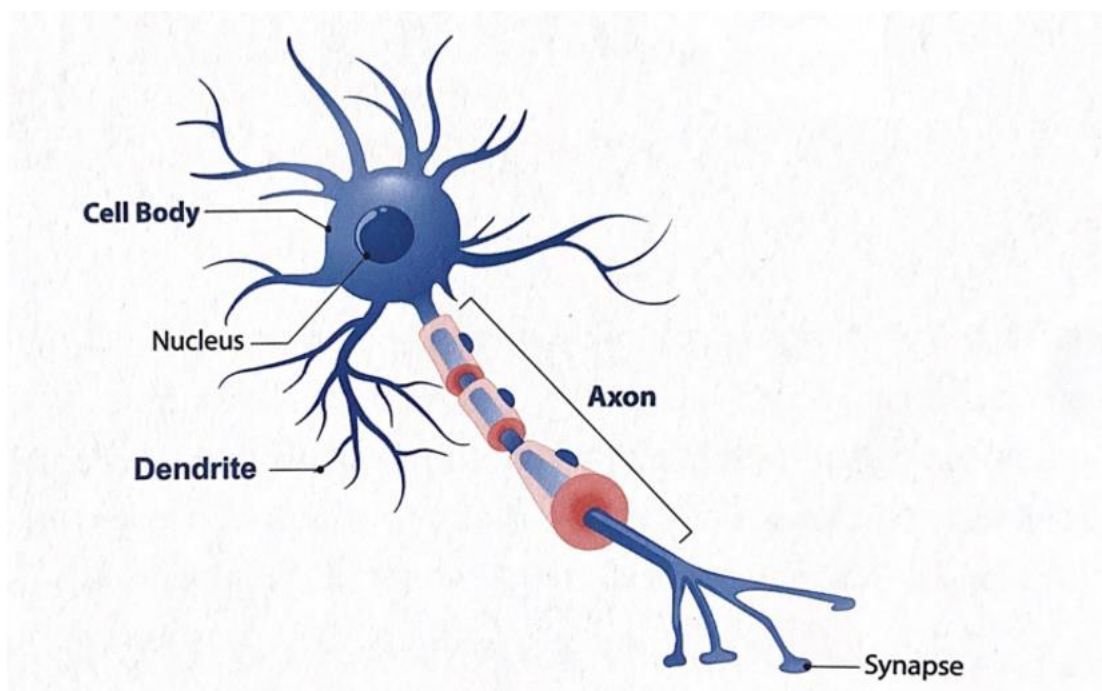
หากเปรียบเทียบช่วงอายุของการเรียนรู้เชิงลึกจะเห็นว่า การเรียนรู้เชิงลึกเกิดขึ้นมาในช่วงปี ค.ศ. 2010 ซึ่งต่างเทคนิคจากการเรียนรู้ของเครื่องที่เกิดขึ้นมานานมากแล้ว ทำให้เทคนิคการเรียนรู้เชิงลึกเป็นที่ให้ความสนใจของนักวิจัยและนักพัฒนามากในช่วง 10 ปีหลังมานี้



ภาพประกอบ 5 แสดงการเปรียบเทียบช่วงเวลาการเกิดของแต่ละเทคนิค

2.3 โครงสร้างเครือข่ายประสาทเทียม (Neural Network)

โครงสร้างเครือข่ายประสาทเทียม (Neural Network) เป็นหนึ่งในเทคนิคการเรียนรู้เชิงลึกเป็นแบบจำลองที่ถูกสร้างขึ้นมามีลักษณะโครงสร้างและการทำงานคล้ายคลึงกับสมองของมนุษย์ การประมวลผลของแบบจำลองจะมีการคำนวณแบบคอนเนคชันนิสต์ (Connectionist) จุดเริ่มต้นของแบบจำลองเกิดขึ้นจากแนวคิดที่ได้มาจากการศึกษาโครงข่ายไฟฟ้าชีวภาพ (Bioelectric Network) ในสมอง ซึ่งประกอบไปด้วย เซลล์ประสาท (Neurons) และ จุดประสานประสาท (Synapse) ตามแบบจำลอง โครงข่ายประสาทเกิดจากการเชื่อมต่อกันระหว่างเซลล์ประสาท ทำให้เกิดเป็นเครือข่ายทำงานร่วมกัน ซึ่งโครงสร้างเครือข่ายประสาทเทียมที่สร้างมาจากการลอกเลียนแบบมาจากลักษณะทางชีววิทยาของสมองมนุษย์



ภาพประกอบ 6 แสดงลักษณะทางชีววิทยาของโครงข่ายประสาท (สงวนสิทธิ์, 2562)

2.3.1 ส่วนประกอบของ Neural Network

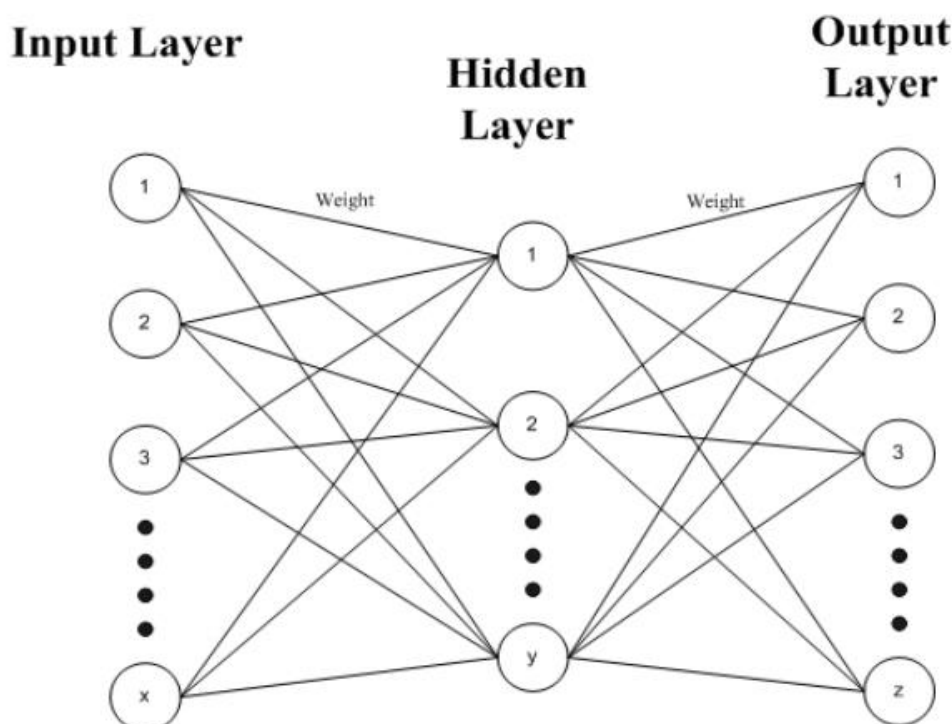
ประกอบไปด้วย Neurons หรือ โหนด (Node) จำนวนมากเชื่อมต่อกันเป็นโครงข่ายประสาท โดยที่แต่ละโหนดจะถูกแบ่งออกแตกต่างกันตามระดับชั้น (layer) ที่มันอยู่ แต่ละชั้นประกอบไปด้วย Input layer, Hidden layer และ Output layer

- **Input layer** ทำหน้าที่รับข้อมูลเข้ามาในโครงข่ายประสาทโดย Input Layer จะมีเพียงชั้นเดียวเท่านั้นและจะมีหน้าที่ส่งข้อมูลต่อไปยังชั้น Hidden Layer

- **Hidden Layer** ทำหน้าที่รับข้อมูลจาก Layer ชั้นก่อนหน้า โดยที่ Hidden Layer จะสามารถมีจำนวนมากกว่า 1 ชั้นได้ และหากต้องการให้แบบจำลองมีความแม่นยำมากขึ้นเราจะเพิ่มจำนวนชั้นของ Hidden Layer และเพิ่มจำนวน Neurons ให้มากขึ้นจะยิ่งทำให้แม่นยำมากขึ้นโดยพื้นฐาน

ภายในทุก Hidden Layer จะมี bias และค่า weight ซ่อนอยู่ภายใน bias จะเชื่อมต่อเพื่อให้ทุกการคำนวณเพื่อส่งต่อมีความเท่าเทียมกันเพื่อให้ bias เข้าไปคำนวณ decision boundary โดยไม่จำเป็นต้องผ่านจุด origin ส่วนค่า weight จะเป็นน้ำหนักที่ส่งผลกับทุก neurons มีค่าไม่เท่ากันทำให้แต่ละคลาสมีน้ำหนักไม่เท่ากันเวลาคำนวณว่าเป็นคลาสไหน ทำให้สามารถแยกได้ว่าข้อมูลนี้เป็นคลาสอะไร

- Output Layer ทำหน้าที่รับข้อมูลจาก Hidden Layer เป็นลำดับสุดท้ายที่ได้ผลลัพธ์ของแบบจำลอง

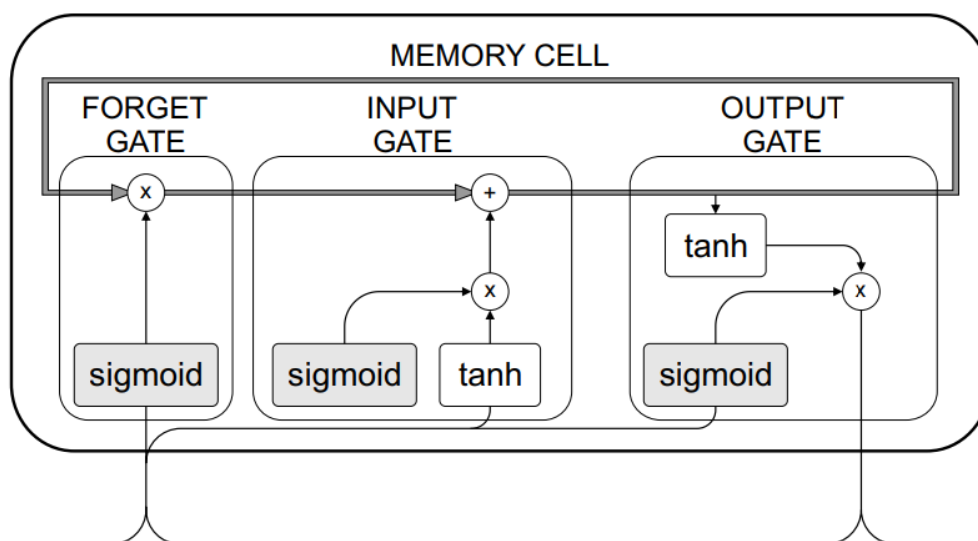


ภาพประกอบ 7 แสดงลักษณะโครงสร้าง Neural Network (Tanty & Desmukh, 2015)

2.4 แบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory)

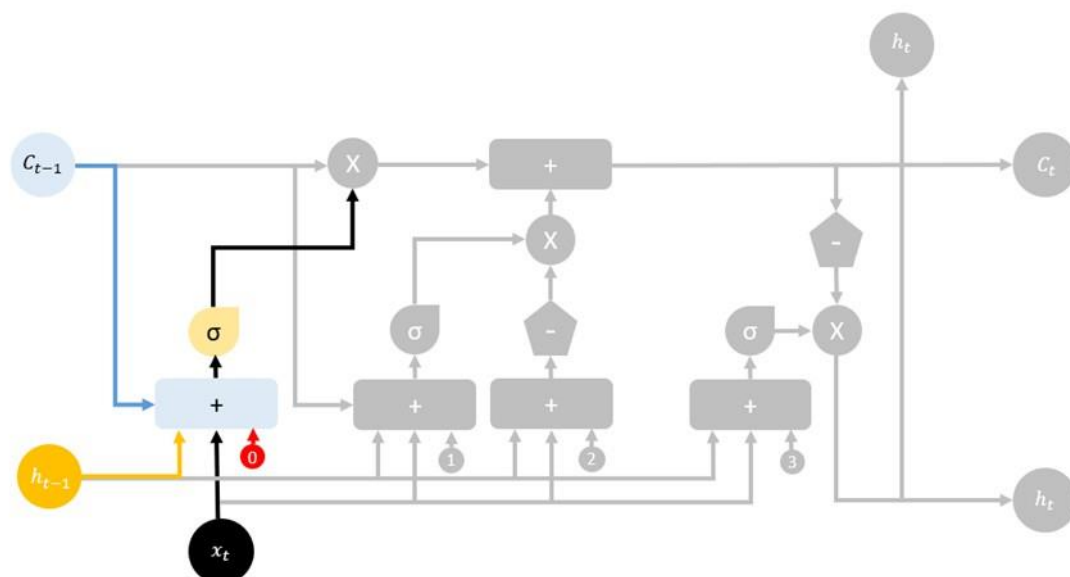
แบบจำลองหน่วยความจำระยะสั้นแบบยาว หรือ Long Short-Term Memory (LSTM) เป็นแบบจำลองที่เหมาะสมสำหรับการวิเคราะห์ข้อมูลแบบลำดับ (Sequence) เช่น การวิเคราะห์คำพูด การวิเคราะห์ข้อมูลวิดีโอ หรือรูปภาพ แบบจำลองนี้ถูกพัฒนาขึ้นจากแบบจำลอง Recurrent Neural Network (RNN) เพื่อแก้ปัญหา Vanishing Gradient Problem เป็นปัญหาที่เกิดขึ้นในแบบจำลอง RNN พบว่าระหว่างการทำ Training การทำงาน back-propagation ทำให้ค่า Gradient จะมีขนาดที่เล็กลงจนมีค่าเท่ากับศูนย์ส่งผลกับค่า Weight ที่ได้ ทำให้แบบจำลองไม่สามารถใช้งานได้ต่อไป และ Exploding Gradient Problem เป็นปัญหาที่เกิดขึ้นในแบบจำลอง RNN โดยมีลักษณะที่ตรงกันข้ามกับ Vanishing Gradient Problem ทำให้ค่า Gradient มีขนาดใหญ่ขึ้น จนมีค่าเป็นอนันต์ส่งผลทำให้เกินกว่าแบบจำลองจะทำงานได้

แบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory) มีประสิทธิภาพในการทำงานได้ดีขึ้นจากแบบจำลอง RNN โดยสามารถจดจำ (Memory) สถานะของแต่ละโหนดทำให้สามารถย้อนกลับไปหาค่าที่แท้จริงของเดิมได้ว่าเป็นอะไรมาก่อน โดยแบบจำลอง LSTM ประกอบไปด้วย Cell State, Input gate, Output gate และ Forget gate โดยที่เซลล์จะจดจำช่วงเวลาต่าง ๆ และประตู (Gate) ทำหน้าที่ควบคุมการไหลของข้อมูลเข้า-ออกจากเซลล์ ดังภาพประกอบ 8



ภาพประกอบ 8 แสดงแบบจำลอง Long Short-Term Memory (LSTM) (Fister & Jagric, 2019)

2.4.1 Forget gate



ภาพประกอบ 9 แสดง Forget gate

Forget เป็นการลบข้อมูลใน Cell State เพื่อเตรียมพร้อมสำหรับการรับข้อมูลใหม่ (Input) โดยที่ Forget Gate จะทำหน้าที่เป็นเหมือนประตูที่ควบคุมให้ลบข้อมูลนั้นทิ้งไปหรือไม่ โดยถ้ามีค่าเป็น 0 คือ ลบข้อมูล แต่ถ้าค่าเป็น 1 คือ ไม่ลบข้อมูล การสร้าง Forget Gate เริ่มจากการที่ Input data เข้ามารวมกับ Hidden State ก่อนหน้า โดยใช้ Sigmoid function เป็นตัวตัดสินว่าจะให้ผ่านหรือไม่ (Medium, 2016b) สูตรในการคำนวณดังนี้ (1)

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (1)$$

จากสมการ

f_t คือ Forget gate

σ คือ ฟังก์ชัน Sigmoid

W_{xf} คือ ค่าน้ำหนักของ Input ใน Forget gate

W_{hf} คือ ค่าน้ำหนักของ Output ใน Forget gate

W_{cf} คือ ค่าน้ำหนักของ Cell state ใน Forget gate

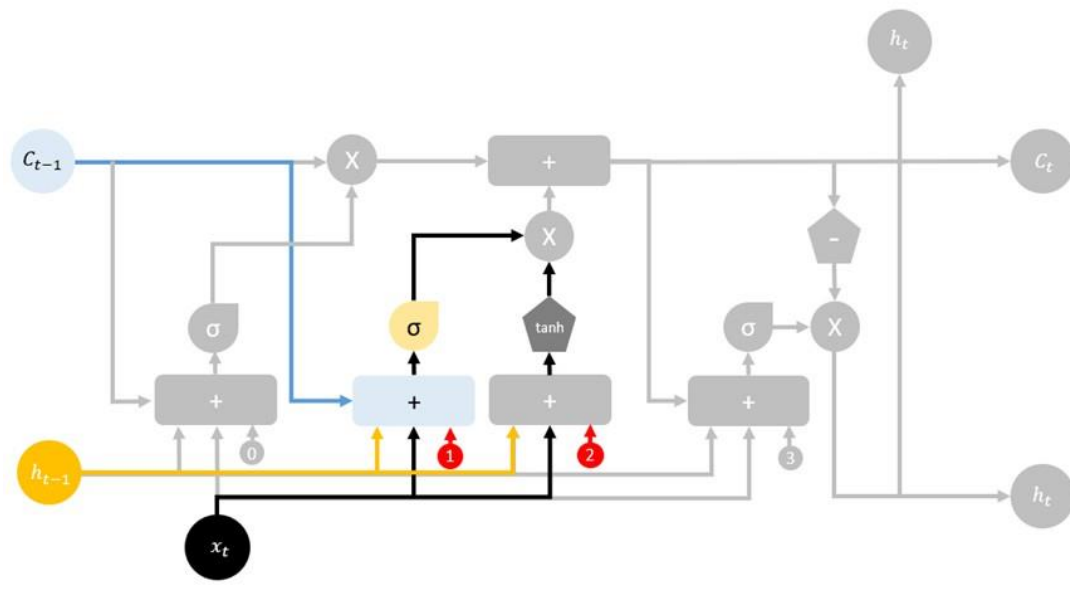
x_t คือ ค่า Input ของ Cell state ณ เวลาที่ t

h_{t-1} คือ ค่า Output ของ cell state ก่อนหน้า ณ เวลาที่ $t-1$

c_{t-1} คือ ค่า Update cell state ณ เวลาที่ t

b_f คือ ค่า Bias

2.4.2 Input gate และ Input Modulation gate



ภาพประกอบ 10 แสดง Input gate และ Input modulation gate

Input gate การคำนวณว่าจะอัปเดต cell state ด้วย Input data ใหม่หรือไม่ ควบคุมโดย Input gate ที่ใช้ Sigmoid function เป็นตัวตัดสินใจว่าอนุญาตให้อัปเดตหรือไม่ เป็นการคำนวณโดยใช้ input data ที่เข้ามากับ hidden state ก่อนหน้านั้นมีสูตรในการคำนวณดังนี้

(2)

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (2)$$

จากสมการ

i_t คือ Input gate

σ คือ ฟังก์ชัน Sigmoid

W_{xi} คือ ค่าน้ำหนักของ Input ใน Input gate

W_{hi} คือ ค่าน้ำหนักของ Output ใน Input gate

W_{ci} คือ ค่าน้ำหนักของ Cell state ใน Input gate

x_t คือ ค่า Input ของ Cell state ณ เวลาที่ t

h_{t-1} คือ ค่า output ของ Cell State ก่อนหน้า ณ เวลาที่ $t-1$

c_{t-1} คือ ค่า Update Cell state ณ เวลาที่ $t-1$

b_i คือ ค่า Bias ของ Input gate

Input Modulation Gate การอัปเดต Cell State จะใช้ Input modulation gate เป็น
ปรับค่า Input data โดยใช้ Tanh Function ซึ่งค่าที่ได้นั้นจะมองเป็น Cell State Candidate (3)

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (3)$$

จากสมการ

$\tilde{c}_t$ คือ ค่า Candidate ของ Cell state ณ เวลาที่ t

$\tanh$ คือ ฟังก์ชัน Tanh

W_{xc} คือ ค่าน้ำหนักของ Input ใน Cell state candidate

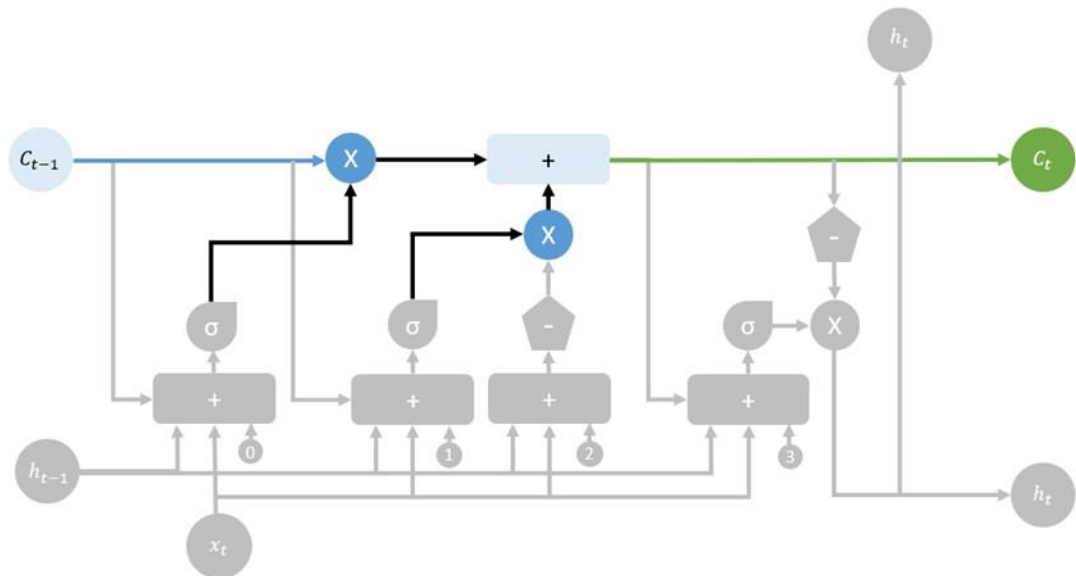
W_{hc} คือ ค่าน้ำหนักของ Output ใน Cell State candidate

x_t คือ ค่า Input ของ Cell state ณ เวลาที่ t

h_{t-1} คือ ค่า Output ของ Cell state ก่อนหน้า ณ เวลาที่ $t-1$

b_c คือ ค่า Bias ของ Cell state candidate

2.4.3 Update Cell State



ภาพประกอบ 11 แสดง Update Cell State

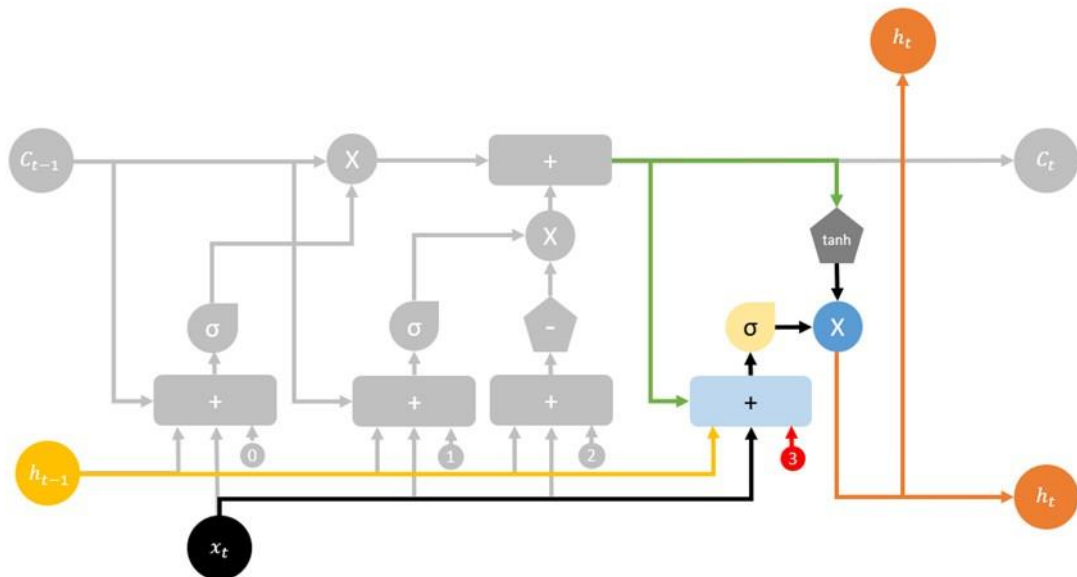
Update Cell State เป็นการทำงานหลังจากที่ได้รับข้อมูลจากประตูทั้งหมดแล้ว ได้แก่ Forget gate, Input gate และ Input modulation gate ทำให้เพียงพอต่อการ Update cell state แล้ว มีสูตรในการคำนวณดังนี้ (4)

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (4)$$

จากสมการ

- c_t คือ Update cell state ณ เวลาที่ t
- f_t คือ Forget gate ณ เวลาที่ t
- c_{t-1} คือ Update cell state ณ เวลาที่ t-1
- i_t คือ Input gate ณ เวลาที่ t
- x_t คือ ค่า Input ของ Cell state ณ เวลาที่ t
- $\tilde{c}_t$ คือ ค่า Candidate ของ Cell State ณ เวลาที่ t

2.4.4 Output Gate หรือ Read



ภาพประกอบ 12 แสดง Output gate

Read เป็นการอนุญาตให้มาอ่าน Output ได้หรือไม่ได้โดยใช้ output gate มาช่วยในการตัดสินใจ โดยใช้ sigmoid function, ค่า hidden state ตัวก่อนหน้า และ input data ที่เข้ามา ณ ขณะนั้นมีสูตรในการคำนวณดังนี้ (5)

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (5)$$

จากสมการ

o_t คือ Output gate ณ เวลาที่ t

σ คือ ฟังก์ชัน Sigmoid

W_{xo} คือ ค่าน้ำหนักของ Input ใน Output gate

W_{ho} คือ ค่าน้ำหนักของ Output ใน Output gate

W_{co} คือ ค่าน้ำหนักของ Update cell stage ใน Output gate

x_t คือ ค่า Input ของ Cell state ณ เวลาที่ t

h_{t-1} คือ ค่า Output ของ Cell state ก่อนหน้า ณ เวลาที่ t-1

c_t คือ ค่า Update cell state ณ เวลาที่ t

b_o คือ ค่า Bias ของ Output gate

ค่า Output ที่จะเป็นค่า h_t สำหรับ sequence ถัดไป มีสูตรในการคำนวณดังนี้ (6)

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

จากสมการ

h_t คือ ค่า Output ของ cell state ก่อนหน้า ณ เวลาที่ t

o_t คือ Output gate ณ เวลาที่ t

$\tanh$ คือ ฟังก์ชัน Tanh

c_t คือ ค่า Update cell state ณ เวลาที่ t

2.5 ตัวปรับพารามิเตอร์แบบจำลองระยะสั้นแบบยาว LSTM (LSTM Optimizer)

2.5.1 Adaptive Gradient Descent (AdaGrad)

เป็นอัลกอริทึมจะทำการปรับอัตราการเรียนรู้จากค่าเริ่มต้นที่กำหนดด้วยตัวเองให้เหมาะสมกับพารามิเตอร์ โดยทำการอัปเดตค่าอัตราการเรียนรู้ต่ำ สำหรับพารามิเตอร์ที่เกี่ยวข้องกับแบบจำลองจะมีการอัปเดตอัตราการเรียนรู้ที่สูงขึ้น วิธีนี้เหมาะสำหรับข้อมูล Sparse Data สำหรับค่าอัตราการเรียนรู้จะถูกเปลี่ยนแปลงทุกครั้งโดยอ้างอิงจากทิศทางที่ผ่านการคำนวณจาก Gradient descent มาแล้ว มีสูตรการคำนวณดังนี้ (7), (8) และ (9) ตามลำดับ

$$v_t = \beta_1 \cdot v_{t-1} - (1 - \beta_1) \cdot g_t \quad (7)$$

$$\hat{v}_t = \max(\hat{v}_{t-1}, v_t) \quad (8)$$

$$\omega_{t+1} = \omega_t - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \cdot v_t \quad (9)$$

จากสมการ

η คือ ค่าอัตราการเรียนรู้

g_t คือ ค่าเกรเดียน ณ เวลาที่ t

v_t คือ ค่าเฉลี่ยเคลื่อนที่ของ g_t ณ เวลาที่ t

β_1 คือ ไฮเปอร์พารามิเตอร์ของ g_t

ϵ คือ ค่า Epsilon

ω_{t+1} คือ ค่าน้ำหนักที่ปรับค่า

2.5.2 Root Mean Square Propagation (RMSprop)

เป็นวิธีการทำให้อัตราการเรียนรู้ถูกปรับให้เหมาะสมกับพารามิเตอร์แต่ละตัว โดยวิธีนี้จะเก็บค่าเกรเดียนของครั้งก่อนหน้ามาใช้ในการเรียนรู้ในครั้งต่อไป โดยใช้ Mean Square Error ในการเก็บค่าเกรเดียนมีสูตรในการคำนวณดังนี้ (10), (11) และ (12) ตามลำดับ

$$v_t = \rho v_{t-1} + (1 - \rho) \cdot g_t^2 \quad (10)$$

$$\Delta\omega_t = -\frac{\eta}{\sqrt{v_t + \epsilon}} \cdot g_t \quad (11)$$

$$\omega_{t+1} = \omega_t + \Delta\omega_t \quad (12)$$

จากสมการ

η คือ ค่าอัตราการเรียนรู้

g_t คือ ค่าเกรเดียน ณ เวลาที่ t

v_t คือ ค่าเฉลี่ยเคลื่อนที่ของ g_t^2 ณ เวลาที่ t

ρ คือ อัตราการใช้เกรเดียนในการเรียนรู้ ณ เวลาที่ $t-1$

ω_{t+1} คือ ค่าน้ำหนักที่ปรับค่า

2.5.3 Adaptive Moment Estimation (Adam)

เป็นวิธีการที่ได้รับความนิยมมากที่สุด เนื่องจากเป็นวิธีที่รวมสิ่งที่ดีที่สุด และลดสิ่งที่ไม่ดีที่สุดออกไปของ AdaGrad และ RMSprop อีกทั้งยังมีความเร็ว และลดปัญหาการแกว่งของพารามิเตอร์ มีสูตรในการคำนวณดังนี้ (13), (14), (15) และ (16) ตามลำดับ

$$v_t = \beta_1 \cdot v_{t-1} + (1 - \beta_1) \cdot g_t \quad (13)$$

$$s_t = \beta_2 \cdot s_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (14)$$

$$\Delta\omega_t = -\eta \frac{v_t}{\sqrt{s_t + \epsilon}} \cdot g_t \quad (15)$$

$$\omega_{t+1} = \omega_t + \Delta\omega_t \quad (16)$$

จากสมการ

η คือ ค่าอัตราการเรียนรู้

g_t คือ ค่าเกรเดียน ณ เวลาที่ t

v_t	คือ ค่าเฉลี่ยเคลื่อนที่ของ g_t ณ เวลาที่ t
s_t	คือ ค่าเฉลี่ยเคลื่อนที่ของ g_t^2 ณ เวลาที่ t
β_1	คือ ไฮเปอร์พารามิเตอร์ของ g_t
β_2	คือ ไฮเปอร์พารามิเตอร์ของ g_t^2
ϵ	คือ ค่า Epsilon
ω_{t+1}	คือ ค่าน้ำหนักที่ปรับค่า

ความแตกต่างของ Optimizer ทั้งสามแบบได้แก่ AdaGrad, RMSprop และ Adam คือ การคำนวณค่าเกรเดียน โดยที่ RMSprop จะใช้การคำนวณโดยใช้การยกกำลังสองแตกต่างจาก AdaGrad ที่ไม่ได้ทำการยกกำลังสองก่อน แต่จะทำปรับอัตราการเรียนรู้จากค่าเริ่มต้นที่กำหนดด้วยตัวเองให้เหมาะสมกับพารามิเตอร์ และ Adam เป็น Optimizer ที่มีคำนวณโดยใช้เกรเดียน และเกรเดียนยกกำลังสอง ทำให้มีความแม่นยำ และเป็นที่ยอมรับมากที่สุด แต่ยังไม่มีความที่บ่งบอกชัดเจนว่า Adam ทำงานได้ดีและมีประสิทธิภาพมากกว่า Optimizer ตัวอื่นยังอยู่ในการศึกษาและวิจัยต่อไป

2.6 ฟังก์ชันกระตุ้น (Activation Function)

Activation Function คือ ฟังก์ชันภายใน Neural Network ที่รับผลรวมการประมวลผลทั้งหมด จากทุก Input ภายใน 1 โหนด แล้วจะทำการพิจารณาโดยฟังก์ชันคำนวณแล้วจะการส่งต่อเป็น Output โดยที่ Activation Function ที่ใช้ในแบบจำลอง LSTM ของงานวิจัยจะมี 2 ประเภท ได้แก่ Sigmoid Function และ Tanh Function

2.6.1 Sigmoid Function

Sigmoid Function เป็นฟังก์ชันที่มีลักษณะเป็นรูปร่างเป็น Curve รูปตัว S ทำให้สามารถเข้าใจง่าย และเนื่องจาก Output ของ Sigmoid Function มีค่าอยู่ระหว่างค่า 0-1 ทำให้เป็นที่นิยมในการถูกใช้งานที่ต้องการ Output เช่น 1=Yes, 0=No มีสูตรการคำนวณดังนี้ (17)

$$S(z) = \phi(z) = \frac{1}{1+e^{-z}} \quad (17)$$

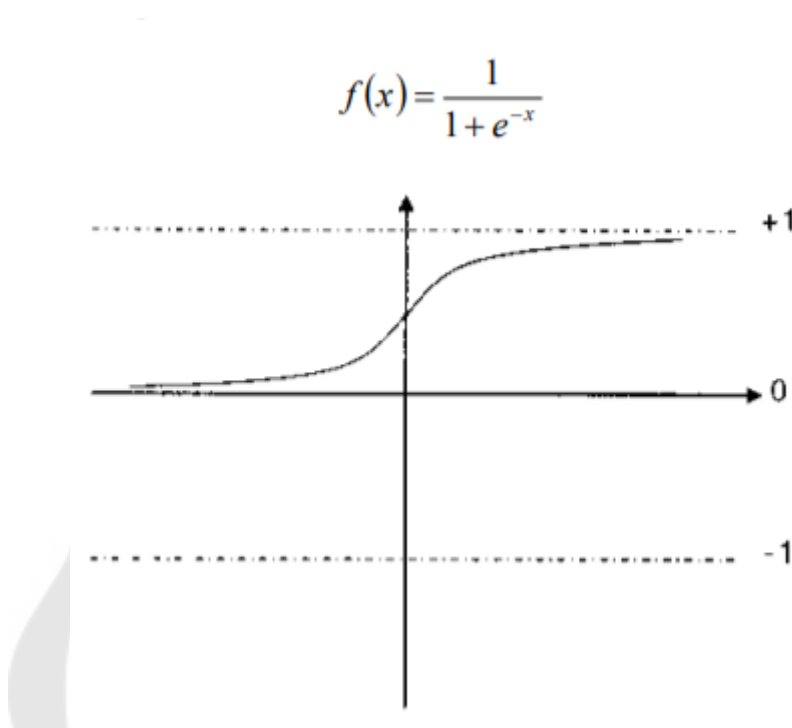
จากสมการ

$\phi(z)$ คือ ฟังก์ชันซิกมอยด์

e คือ ค่าลอการิทึมธรรมชาติ ประมาณ 2.71828

z คือ ข้อมูล Input

เพื่อให้แบบจำลองโครงข่ายประสาทเทียมทำงานได้ดีขึ้นในการฝึกฝนข้อมูล (training dataset) จึงมี Tanh function ที่เป็นที่นิยมในการใช้งาน



ภาพประกอบ 13 แสดง Sigmoid Function (ประกอบผล, 2552)

2.6.2 Tanh function (Hyperbolic Tangent Activation Function)

Tanh Function เป็นฟังก์ชันที่แก้ไขข้อเสียในหลายเรื่องของ Sigmoid Function แต่มีลักษณะรูปร่างที่เหมือนกันโดยเป็นตัว S แตกต่างกันว่า Sigmoid จะมีค่า Output ระหว่าง 0-1 แต่ Tanh จะมีค่า Output ระหว่าง -1 ถึง 1 มีสูตรคำนวณดังนี้ (18)

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (18)$$

จากสมการ

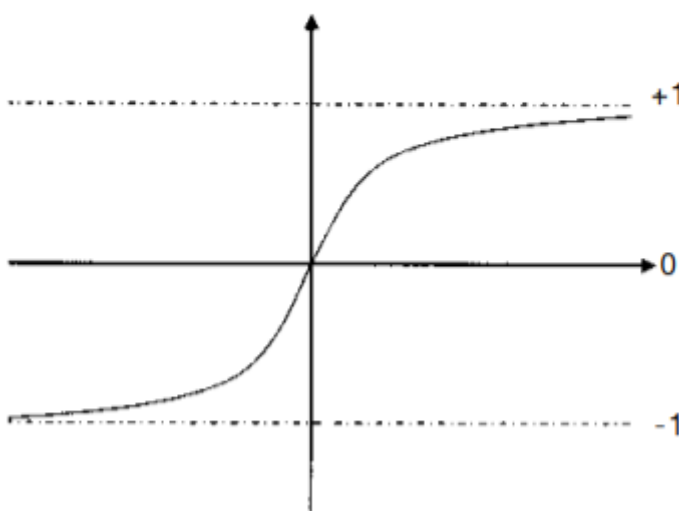
$\tanh(x)$ คือ ฟังก์ชัน Tanh

e คือ ค่าลอการิทึมธรรมชาติ ประมาณ 2.71828

x คือ ข้อมูล Input

ข้อดีของ Tanh Function ทำให้ Output มีความสมดุล (Balance) เนื่องจากมีค่าเฉลี่ย (Mean) เท่ากับ 0 ทำให้ Optimize ได้ง่ายขึ้น และ Output มีค่าระหว่าง -1 ถึง 1 ทำให้นิยมใช้งานเกี่ยวกับ Classification ระหว่าง 2 อย่างและทำให้ Normalize ให้เป็น 1 Unit

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



ภาพประกอบ 14 แสดง Tanh Function (ประกอบผล, 2552)

2.7 ทฤษฎีที่เกี่ยวข้องกับการประเมินวัดความแม่นยำของแบบจำลอง

สำหรับปัญหา Regression เพื่อใช้ในการทำนายค่า y ข้อมูลจะมีลักษณะเป็นตัวเลข (continuous) สามารถคำนวณหาค่า Loss function เพื่อเป้าหมายในการ train แบบจำลองเพื่อประเมินวัดความแม่นยำของแบบจำลอง ให้มีค่า loss ที่มีค่าต่ำที่สุด โดยในงานวิจัยนี้ได้ใช้เทคนิครากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error: RMSE) ,ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Error: MAE) และค่าเฉลี่ยของร้อยละความคลาดเคลื่อนสัมบูรณ์ (Mean absolute percentage error: MAPE)

ค่าการประเมินหากมีค่าน้อยหมายถึงการพยากรณ์มีผลลัพธ์และมีประสิทธิภาพ

y_t คือ ค่าความเข้มข้นฝุ่นละอองจริง ณ เวลาที่ t

$\hat{y}_t$ คือ ค่าความเข้มข้นฝุ่นละอองจากการพยากรณ์ ณ เวลาที่ t

n คือ จำนวนข้อมูลกลุ่มตัวอย่าง

2.7.1 รากของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root mean square error: RMSE)

เป็นวิธีวัดความแม่นยำค่าพยากรณ์โดยการคำนวณรากที่สองกำลังเฉลี่ย โดย RMSE จะทำการยกกำลังสองค่าคลาดเคลื่อนก่อนที่จะเฉลี่ยค่าเฉลี่ยทำให้ RMSE ให้ค่าน้ำหนักค่อนข้างสูงถึงค่าความคลาดเคลื่อนค่อนข้างสูงถึงข้อผิดพลาดขนาดใหญ่ ดังนั้น RMSE จะมีประโยชน์สำหรับความคลาดเคลื่อนที่ไม่พึงประสงค์เป็นพิเศษ ค่า RMSE มีค่าระหว่างศูนย์ถึงอนันต์ หรือ infinity หากค่า RMSE เป็นศูนย์หมายถึงประสิทธิภาพที่ดีสมบูรณ์แบบ (Medium, 2016a) มีสูตรในการคำนวณดังนี้ (18)

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (18)$$

2.7.2 ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error: MAE)

เป็นตัวชี้วัดที่ใช้งานง่ายที่สุดเนื่องจากเพียงแค่วัดความแตกต่างที่แน่นอนระหว่างข้อมูลและการคาดการณ์ของแบบจำลอง เนื่องจากเราใช้ค่าสัมบูรณ์ (Absolute) ของค่าคลาดเคลื่อน MAE ไม่ได้บ่งชี้ประสิทธิภาพที่ต่ำกว่าหรือมากเกินไปของแบบจำลอง (Medium, 2016a) มีสูตรคำนวณดังนี้ (19)

$$MAE = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n} \quad (19)$$

ค่า MAE ที่มีขนาดเล็กสามารถบอกได้ว่าแบบจำลองนั้นมีประสิทธิภาพที่ดีมาก ส่วนค่า MAE ที่มีขนาดใหญ่สามารถบอกได้ว่าแบบจำลองนั้นอาจมีปัญหาในบางพื้นที่ ค่า MAE มีค่าระหว่างศูนย์ถึงอนันต์ หรือ infinity ถ้า MAE เป็นศูนย์ หมายถึง แบบจำลองมีประสิทธิภาพที่ดีสมบูรณ์แบบ ค่า MAE สามารถตีความได้ง่ายโดยการใช้ค่าสัมบูรณ์ของค่าคลาดเคลื่อนแต่ก็ไม่เป็นนิยามเท่ากับค่า RMSE ที่มีการยกกำลังสองก่อน

2.7.3 ค่าเฉลี่ยของร้อยละความคลาดเคลื่อนสัมบูรณ์ (Mean absolute percentage error: MAPE)

ลักษณะคล้ายกับค่า MAE แต่มีการปรับเปลี่ยนทุกอย่างให้เป็นร้อยละ โดย MAPE เป็นระยะการคาดการณ์ของแบบจำลองที่ห่างจากผลลัพธ์ที่สอดคล้องกันโดยเฉลี่ย ดีความที่ชัดเจนเนื่องจากเป็นร้อยละทำให้ง่ายต่อผู้ใช้ ค่า MAPE มีค่าระหว่างร้อยละศูนย์ถึงร้อยละอนันต์ หรือ Infinite มีสูตรคำนวณดังนี้ (20)

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (20)$$

ความแตกต่างของตัวประเมินวัดความแม่นยำของแบบจำลองทั้งสามแบบได้แก่ RMSE, MAE และ MAPE คือ การคำนวณค่าคลาดเคลื่อน โดยที่ RMSE จะใช้การคำนวณจะให้น้ำหนักโดยการยกกำลังสองค่าคลาดเคลื่อนสูงทำให้เห็นค่าคลาดเคลื่อนได้ชัดเจน แตกต่างกับ MAE ที่ลักษณะในการคำนวณคลาดเคลื่อนโดยใช้ค่าสัมบูรณ์ (Absolute) ส่วน MAPE สามารถวัดความคลาดเคลื่อนจากมาตราส่วน (Scale) ของข้อมูล และเปอร์เซ็นต์ร้อยละในการคำนวณค่าคลาดเคลื่อน ในการประเมินวัดความแม่นยำของแบบจำลองควรใช้แบบจำลองมากกว่า 1-2 ตัวขึ้นไปเพื่อให้สามารถเปรียบเทียบความแม่นยำของแบบจำลองได้ดียิ่งขึ้น

2.8 งานวิจัยที่เกี่ยวข้อง

การทบทวนวรรณกรรมของงานวิจัยนี้ได้ทำการศึกษางานวิจัยที่เกี่ยวข้องกับการพยากรณ์ค่าความเข้มข้นฝุ่นละออง PM2.5 โดยงานวิจัยที่เกี่ยวข้องมีรายละเอียดดังต่อไปนี้

(1) บทความวิจัยเรื่อง Air pollution forecasting using RNN with LSTM (Tsai, Zeng, & Chang, 2018)

งานวิจัยได้ใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ข้อมูลที่ใช้เป็นข้อมูลจากได้หวัน ตั้งแต่ ปี ค.ศ. 2012 ถึง ค.ศ. 2016 โดยมีจำนวน Feature ที่ใช้ 20 คอลัมน์ ได้แก่ ค่าความเข้มข้นฝุ่นละออง PM2.5, ค่าความเข้มข้น PM10, SO2 (ซัลเฟอร์ไดออกไซด์), O3 (โอโซน), NOX (ไนโตรเจนออกไซด์), NO2 (ไนโตรเจนไดออกไซด์), NO (ไนตริกออกไซด์), CO (คาร์บอนมอนอกไซด์), Temperature (อุณหภูมิ), Humidity (ความชื้นสัมพัทธ์), Wind Speed Hour (ความเร็วลมรายชั่วโมง), Wind Direct (ทิศทางลมรายชั่วโมง), Rainfall (ปริมาณน้ำฝน), Wind Speed [last 10 min/hrs.] (ความเร็วลม 10 นาทีสุดท้ายต่อชั่วโมง), Wind Direct [last 10

min/hrs.] (ทิศทางลม 10 นาทีสุดท้ายต่อชั่วโมง), Month (เดือน), Weekday (สัปดาห์), Hour (ชั่วโมง), %PM2.5 และ %PM10 งานวิจัยได้ทำการทดลองประเมินค่าความเข้มข้นฝุ่นละออง PM2.5 ใน 4 ชั่วโมงถัดไป โดยใช้ 66 สถานีรอบไต้หวันแบ่งออกเป็น 4 กลุ่ม คือ Northern Taiwan, Middle Taiwan, Southern central Taiwan และ Southern Taiwan

ผลวิจัยพบว่าแบบจำลองสามารถพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 ได้อย่างมีประสิทธิภาพ โดยมีการประเมินวัดความแม่นยำของแบบจำลอง โดยแบ่งออกในการพยากรณ์ 4 ชั่วโมงถัดไปทุกพื้นที่ ตั้งแต่ 1–4 ชั่วโมง พบว่าค่าที่น้อยที่สุดในชั่วโมงที่ 1 โดยมีค่า RMSE เฉลี่ยทุกพื้นที่อยู่ที่ $4.72 \mu\text{g}/\text{m}^3$ และค่า MAE เฉลี่ยทุกพื้นที่อยู่ที่ $3.41 \mu\text{g}/\text{m}^3$

(2) บทความวิจัยเรื่อง Prediction of 2.5 Concentration Based on CEEMD-LSTM Model (Li & Shen, 2019)

งานวิจัยได้ใช้แบบจำลองแบบ Hybrid ของ CEEMD-LSTM (Complementary Ensemble Empirical Model Decomposition Long Short-Term Memory) เพื่อพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 โดยข้อมูลที่ใช้เป็นข้อมูลสถานทูตสหรัฐอเมริกาในกรุงปักกิ่ง และตัวแปรมลพิษทางอากาศที่สำคัญอื่น ๆ ถูกนำมาใช้ใน input ของแบบจำลอง โดยเปรียบเทียบ CEEMD-LSTM กับอัลกอริทึม ELM, SVR และ LSTM

ผลวิจัยพบว่าสามารถพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 โดยประเมินวัดความแม่นยำของแบบจำลองโดยการเปรียบเทียบค่า MAE, RMSE และค่า MAPE สรุปผลได้ว่าแบบจำลอง CEEMD-LSTM ให้ค่า MAE เป็น $10.787 \mu\text{g}/\text{m}^3$, RMSE เป็น $14.637 \mu\text{g}/\text{m}^3$ และ MAPE เป็น $19.071 \mu\text{g}/\text{m}^3$ ออกมาดีที่สุด เมื่อเปรียบเทียบกับอัลกอริทึมตัวอื่น

(3) บทความวิจัยเรื่อง PM10 Density Forecast Model Using Long Short-Term Memory (Park et al., 2017)

งานวิจัยได้เปรียบเทียบแบบจำลอง LSTM (RMSprop), LSTM (Adagrad), Linear Regression และ RNN เพื่อใช้สำหรับพยากรณ์ความเข้มข้น PM10 โดยข้อมูลที่ใช้เป็นข้อมูลจากกรุงโซล ประเทศเกาหลีใต้ ตั้งแต่เดือน มกราคม ปี ค.ศ. 2005 ถึงเดือนมีนาคม ค.ศ. 2016

ผลวิจัยพบว่าแบบจำลองสามารถพยากรณ์ฝุ่นละออง PM10 ได้อย่างมีประสิทธิภาพ โดยมีการประเมินวัดความแม่นยำของแบบจำลองเพื่อเปรียบเทียบค่า MSE และ RMSE ของแต่ละแบบจำลอง พบว่าแบบจำลอง LSTM ทำงานได้ดีกว่าแบบจำลอง Linear

Regression กับ RNN และเมื่อทำการเปรียบเทียบแบบจำลองระหว่าง LSTM (RMSprop) กับ LSTM (AdaGrad) ผลที่ได้ คือ แบบจำลอง LSTM (RMSprop) ทำงานได้ผลที่ดีที่สุด

(4) บทความวิจัยเรื่อง Time Series Forecasting using Sequence-to-Sequence Deep Learning Framework (Du, Li, & Horng, 2018)

งานวิจัยได้ใช้แบบจำลอง Shallow ประกอบไปด้วยแบบจำลอง SVR-POLY, SVR-RBF, SVR-LINEAR เปรียบเทียบกับแบบจำลอง Deep Learning ประกอบไปด้วย LSTM, GRU, RNN และเปรียบเทียบกับแบบจำลอง TSDLF โดยชุดข้อมูลคุณภาพอากาศที่เผยแพร่โดยสถานทูตอเมริกาในกรุงปักกิ่ง โดยข้อมูลจะถูกรวบรวมเป็นรายชั่วโมง ตั้งแต่วันที่ 1 มกราคม ค.ศ. 2010 ถึง 31 ธันวาคม ค.ศ. 2017 ซึ่งมี 43824 แถว

ผลวิจัยพบว่าสามารถทำนายปริมาณฝุ่นละออง PM2.5 โดยมีการเปรียบเทียบค่า RMSE สรุปผลได้ว่าแบบจำลอง TSDLF ทำงานได้ดีกว่าแบบจำลอง Shallow และ Deep Learning

(5) บทความวิจัยเรื่อง Deep Air Quality Forecasting Using Hybrid Deep Learning (Du, Li, Yang, & Horng, 2019)

งานวิจัยได้ใช้แบบจำลอง DAQFF (Deep Air Quality Forecasting Framework) เป็นแบบจำลองแบบ Hybrid ที่รวมกันของ CNN และ Bi-directional LSTM เปรียบเทียบกับอัลกอริทึมประเภทอื่น โดยใช้ชุดข้อมูล 2 ชุด คือ ชุดข้อมูลฝุ่นละออง PM2.5 กรุงปักกิ่ง (Beijing PM2.5) ข้อมูลระหว่างวันที่ 1 มกราคม ค.ศ. 2010 ถึง 31 ธันวาคม ค.ศ. 2014 จำนวน 43,824 แถว 1 สถานี และชุดข้อมูลคุณภาพอากาศในตัวเมือง (Urban Air Quality) ข้อมูลระหว่างวันที่ 1 พฤษภาคม ค.ศ. 2014 ถึง 30 เมษายน ค.ศ. 2015 จำนวน 278,023 แถว 36 สถานี

ผลวิจัยพบว่าแบบจำลองสามารถพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 โดยมีการเปรียบเทียบค่า RMSE และ MAE ในการเปรียบเทียบอัลกอริทึมทั้งหมด พบว่าแบบจำลอง DAQFF เป็นแบบจำลองที่ทำงานได้ดีที่สุดจากชุดข้อมูลทั้ง 2 ประเภท

(6) บทความวิจัยเรื่อง System for Detecting and Forecasting PM2.5 Concentration Levels using Long Short-Term Memory and LoRa (Song, Han, Xie, Gao, & Song, 2019)

งานวิจัยได้ใช้แบบจำลอง Long Short-Term Memory ในการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 โดยชุดข้อมูลจากเทคโนโลยีการสื่อสาร LoRa และระบบคลาวด์ (Cloud)

ที่รวบรวมข้อมูลฝุ่นละออง PM2.5 ที่มีประสิทธิภาพ ซึ่งชุดข้อมูลรวบรวมมาจากประเทศจีนตอนเหนือ รายชั่วโมง

ผลวิจัยพบว่าสามารถพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 โดยมีการเปรียบเทียบค่า MAE และ RMSE ของ Time step 3 ช่วง คือ 12 ชั่วโมง, 24 ชั่วโมง และ 48 ชั่วโมง พบว่าจากการเปรียบเทียบค่า MAE และ RMSE Time step 48 ชั่วโมงให้ประสิทธิภาพที่ดีที่สุด

(7) บทความวิจัยเรื่อง Evaluation of a Recurrent Neural Network LSTM for the Detection of Exceedances of Particles PM10 (Montañez, Fernandez, Arriaga, Arreguin, & Calderon, 2019)

งานวิจัยได้ใช้งานวิจัยนี้ได้ใช้แบบจำลอง Long Short-Term Memory ในการพยากรณ์ความเข้มข้นฝุ่น PM10 โดยชุดข้อมูลที่ใช้เป็นกรุงโซล ประเทศเกาหลีใต้ ระหว่างเดือนมกราคม ค.ศ. 2015 ถึง มีนาคม ค.ศ. 2016

ผลวิจัยพบว่าสามารถพยากรณ์ความเข้มข้นฝุ่นละออง PM10 โดยมีการเปรียบเทียบค่า MAE และ RMSE ของ Time step 3 ช่วง คือ 24 ชั่วโมง, 48 ชั่วโมง และ 72 ชั่วโมง พบว่าจากการเปรียบเทียบค่า MAE และ RMSE Time step 72 ชั่วโมงให้ประสิทธิภาพที่ดีที่สุด

(8) บทความวิจัยเรื่อง Long Short-Term Memory Deep Neural Network model for PM2.5 Forecasting in the Bangkok urban area (Thaweephon & Wiwatwattana, 2019)

งานวิจัยได้เปรียบเทียบแบบจำลอง LSTM กับ SARIMAX ในการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 ใน 24 ชั่วโมงล่วงหน้า โดยใช้ข้อมูลสถานีตำรวจนครบาลโชคชัย เป็นตัวแทนจังหวัดกรุงเทพมหานครชั้นใน โดยในแบบจำลอง LSTM ได้ใช้ Optimizer เป็น RMSprop

ผลวิจัยพบว่าแบบจำลอง LSTM (RMSprop) ให้ค่า RMSE และ MAE ได้ดีกว่าแบบจำลอง SARIMAX โดยการพยากรณ์ล่วงหน้า 1 ชั่วโมงมีค่า RMSE เท่ากับ $4.45 \mu\text{g}/\text{m}^3$ และค่า MAE เท่ากับ $3.37 \mu\text{g}/\text{m}^3$

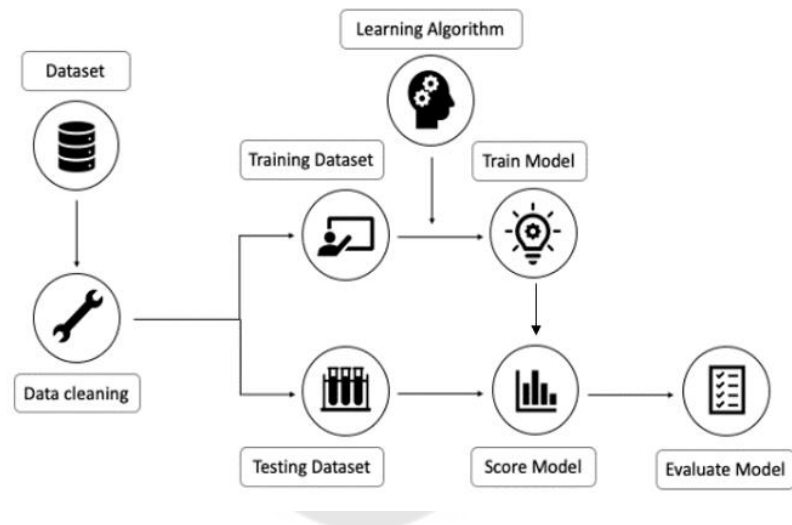
บทที่ 3

วิธีการดำเนินการวิจัย

วิธีการและขั้นตอนในการดำเนินงานวิจัย ผู้วิจัยได้แบ่งขั้นตอนออกไป 5 ส่วน ดังนี้

1. การสร้างเครื่องมือที่ใช้ในการวิจัย
2. การสำรวจข้อมูล (Exploratory Data Analysis) ก่อนทำความสะอาดข้อมูล
3. การทำความสะอาดข้อมูล (Data Cleaning)
4. การสำรวจข้อมูล (Exploratory Data Analysis) หลังทำความสะอาดข้อมูล
5. การสร้างแบบจำลอง LSTM

3.1 การสร้างเครื่องมือที่ใช้การวิจัย



ภาพประกอบ 15 แสดงขั้นตอนการสร้างแบบจำลองการเรียนรู้ของเครื่อง

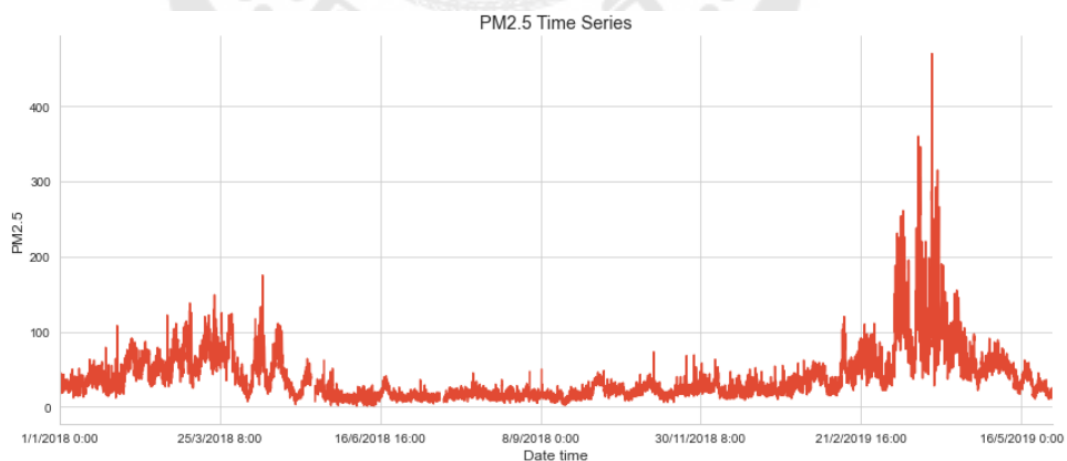
งานวิจัยมีขั้นตอนการสร้างแบบจำลองโดยเริ่มจากการนำชุดข้อมูล (Dataset) เข้าแบบจำลอง, การทำความสะอาดข้อมูล (Data Cleaning), แบ่งข้อมูลออกเป็น 2 ส่วน คือ ชุดการเรียนรู้ (Training Dataset) และชุดการทดสอบ (Testing Dataset), สร้างแบบจำลอง LSTM โดยนำข้อมูลชุดการเรียนรู้เข้าสู่แบบจำลองเพื่อเรียนรู้แบบจำลองนั้นทำการทดสอบแบบจำลองเพื่อให้ทราบผลลัพธ์ของแบบจำลอง และทำการประเมินผลแบบจำลอง (Evaluate Model) ดังแสดงในภาพประกอบ 15

3.2 การสำรวจข้อมูล (Exploratory Data Analysis) ก่อนทำความสะอาดข้อมูล

	NO	NOX	NO2	SO2	O3
Date time					
31/5/2019 14:00	1.0	4.0	3.0	0.0	40.0
31/5/2019 15:00	1.0	5.0	3.0	0.0	34.0
31/5/2019 16:00	3.0	10.0	7.0	1.0	24.0
31/5/2019 17:00	4.0	14.0	11.0	0.0	19.0
31/5/2019 18:00	NaN	NaN	NaN	NaN	NaN
31/5/2019 19:00	2.0	15.0	13.0	0.0	14.0
31/5/2019 20:00	2.0	14.0	13.0	0.0	9.0

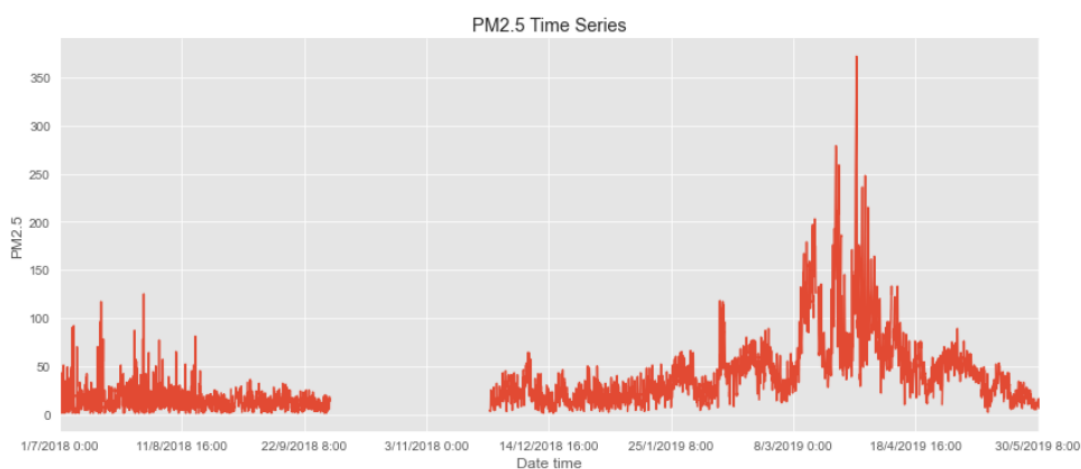
ภาพประกอบ 16 แสดงค่า NaN

จากภาพประกอบ 16 เป็นภาพตัวอย่างข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ที่ทำการสำรวจข้อมูลเบื้องต้นพบว่าในชุดข้อมูลจะมีค่า NaN หรือค่า Missing value อยู่ภายในชุดข้อมูล ดังนั้นผู้วิจัยจึงต้องทำความสะอาดข้อมูล Data Cleaning เพื่อข้อมูลมีความถูกต้องมีประสิทธิภาพมากขึ้นสำหรับการนำไปใช้ในแบบจำลองหน่วยความจำระยะสั้นแบบยาว โดยผู้วิจัยต้องทำความสะอาดข้อมูลให้เรียบร้อยสำหรับข้อมูลทั้ง 2 ชุด ทั้งศูนย์ราชการจังหวัดเชียงใหม่และโรงเรียนยุพราชวิทยาลัย



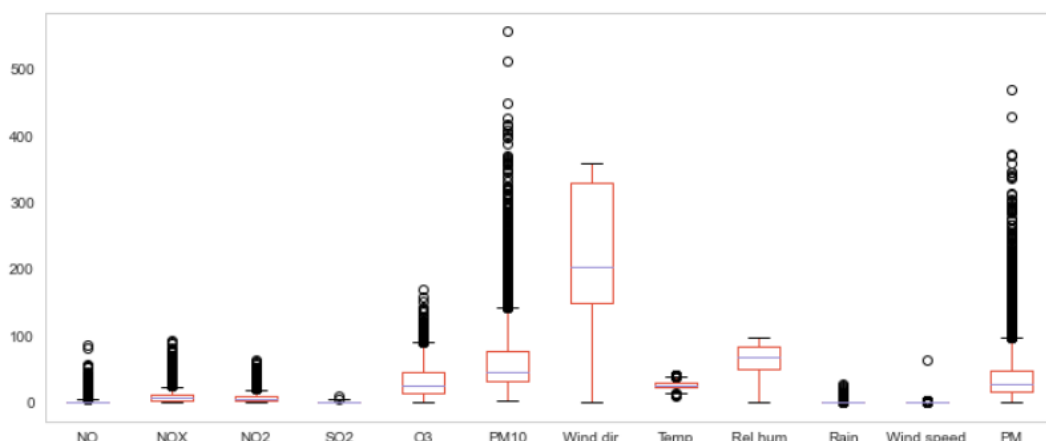
ภาพประกอบ 17 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (ก่อนทำการทำความสะอาดข้อมูล)

จากภาพประกอบ 17 ได้ทำการวิเคราะห์แนวโน้มที่เกิดขึ้นของความเข้มข้นฝุ่นละออง PM2.5 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (ก่อนทำการทำความสะอาดข้อมูล) โดยเทียบข้อมูล PM2.5 กับช่วง Date time ระหว่างวันที่ 1 มกราคม 2018 เวลา 0:00 ถึง 31 พฤษภาคม 2019 เวลา 23:00 จะเห็นได้ว่าข้อมูลชุดความเข้มข้นฝุ่นละออง PM2.5 มีค่า Missing value ค่อนข้างน้อยและข้อมูลมีความต่อเนื่องที่ดี โดยจะมีค่าฝุ่นละออง PM2.5 ที่สูงในช่วงต้นปีเดือนกุมภาพันธ์ ถึงเมษายน ค.ศ. 2019



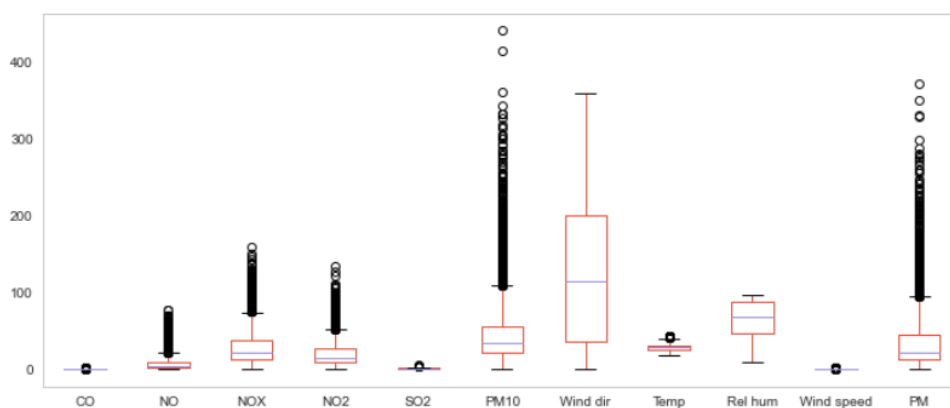
ภาพประกอบ 18 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย (ก่อนทำการทำความสะอาดข้อมูล)

จากภาพประกอบ 18 ได้ทำการวิเคราะห์แนวโน้มที่เกิดขึ้นของความเข้มข้นฝุ่นละออง PM2.5 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย (ก่อนทำการทำความสะอาดข้อมูล) โดยเทียบข้อมูล PM2.5 กับช่วง Date time วันที่ 1 กรกฎาคม 2018 เวลา 0:00 ถึง 31 พฤษภาคม 2019 เวลา 23:00 จะเห็นได้ว่าข้อมูลชุดความเข้มข้นฝุ่นละออง PM2.5 มีค่า Missing value ค่อนข้างมากและมีข้อมูลที่ขาดหายไปช่วงตุลาคม ค.ศ. 2018 ถึงต้นเดือนธันวาคม ค.ศ. 2018 โดยจะมีค่าฝุ่นละออง PM2.5 ที่สูงในช่วงต้นปีเดือนกุมภาพันธ์ถึงเมษายน ค.ศ. 2019 แต่จะมีช่วงที่ข้อมูลความเข้มข้นฝุ่นละออง PM2.5 เป็นค่า Missing value ในช่วงเดือนตุลาคม ถึงต้นเดือน ธันวาคม ค.ศ. 2018



ภาพประกอบ 19 แสดงกราฟ Boxplot (ก่อนทำความสะอาดข้อมูล) ของชุดข้อมูลศูนย์ราชการ จังหวัดเชียงใหม่

จากภาพประกอบ 19 แสดงกราฟ Boxplot (ก่อนทำความสะอาดข้อมูล) แสดงให้เห็นถึงการแจกแจงของข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ โดยจะเห็นได้ว่าตัวแปร PM10 และ PM (PM2.5) มี Outlier ค่อนข้างมากเมื่อเทียบกับตัวแปรประเภทอื่น NO, NOX, NO2, SO2 ค่อนข้างเกาะกลุ่ม การกระจายของข้อมูลค่อนข้างแคบ แต่ O3, PM2.5, PM10 การกระจายของข้อมูลค่อนข้างกว้างกว่าข้อมูลโรงเรียนยุพราชวิทยาลัย



ภาพประกอบ 20 แสดงกราฟ Boxplot (ก่อนทำความสะอาดข้อมูล) ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 20 แสดงกราฟ Boxplot แสดงให้เห็นถึงการแจกแจงของข้อมูลโรงเรียนยุพราชวิทยาลัย โดยจะเห็นได้ว่า Feature PM10 และ PM (PM2.5) มี Outlier ค่อนข้างมากเมื่อเทียบกับตัวแปรประเภทอื่น แตกต่างกันว่า NOX และ NO2 การกระจายของข้อมูลค่อนข้างกว้างกว่า NOX และ NO2 ของข้อมูลสถานีศูนย์ราชการจังหวัดเชียงใหม่

3.3 การทำความสะอาดข้อมูล (Data Cleaning)

ในการเตรียมข้อมูลเพื่อสร้างแบบจำลองผู้วิจัยได้ทำการทำความสะอาดชุดข้อมูลของทั้ง 2 ชุด คือ ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ และชุดข้อมูลโรงเรียนยุพราชวิทยาลัย โดยการแทนค่า Missing Value หรือค่าที่เป็น NaN ด้วยค่าเฉลี่ย (Average) เพื่อให้ข้อมูลที่ลำดับมากขึ้น

แบบจำลอง LSTM เป็นแบบจำลองที่ต้องใช้กับข้อมูลมีลักษณะเป็นลำดับ (Continuous) ผู้วิจัยจึงใช้วิธีการแทนค่าเฉลี่ย (Average) ของแต่ละ Feature แทนค่าที่ NaN ที่เป็นช่องว่าง ซึ่งหากตัดข้อมูลทิ้งไปทำให้ข้อมูลไม่ต่อเนื่องได้

	NO	NOX	NO2	SO2	O3
Date time					
31/5/2019 17:00	4.000000	14.000000	11.00000	0.000000	19.00000
31/5/2019 18:00	1.941253	9.649514	7.67216	0.876299	32.02589
31/5/2019 19:00	2.000000	15.000000	13.00000	0.000000	14.00000
31/5/2019 20:00	2.000000	14.000000	13.00000	0.000000	9.00000
31/5/2019 21:00	1.000000	6.000000	5.00000	0.000000	12.00000
31/5/2019 22:00	1.000000	4.000000	3.00000	0.000000	16.00000
31/5/2019 23:00	1.000000	2.000000	2.00000	0.000000	20.00000

ภาพประกอบ 21 แสดงตัวอย่างการแทนค่า NaN ด้วยค่าเฉลี่ย

3.4 การสำรวจข้อมูล (Exploratory Data Analysis) หลังทำความสะอาดข้อมูล

การสำรวจข้อมูลสถิติเบื้องต้นของข้อมูลทั้ง 2 ชุด คือ ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ และชุดข้อมูลโรงเรียนยุพราชวิทยาลัย โดยใช้ฟังก์ชัน `df.describe()`

จากภาพประกอบ 22, 23 และ 24 จะเป็นการสำรวจข้อมูลของชุดข้อมูลทั้ง 2 ชุด ทำให้เห็นค่าสถิติเบื้องต้นของแต่ละคอลัมน์โดยจะแสดงจำนวนข้อมูล (count), ค่าเฉลี่ย (mean), ค่าส่วนเบี่ยงเบนมาตรฐาน (std), ค่าต่ำสุด(min), Quartiles 25%, 50%, 75% และค่าสูงสุด(max) เพื่อให้ทราบถึงการกระจายตัวของชุดข้อมูล

	NO	NOX	NO2	SO2	O3	PM10	Wind dir	Temp	Rel hum	Rain
count	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000
mean	1.941253	9.649514	7.67216	0.876299	32.025890	60.414032	221.668535	26.088361	66.669971	0.075486
std	3.745597	9.189198	7.10341	0.905685	22.371689	41.703899	113.823416	4.886454	21.181902	0.753198
min	0.000000	0.000000	0.000000	0.000000	0.000000	4.000000	0.000000	11.100000	0.000000	0.000000
25%	0.000000	4.000000	3.000000	0.000000	15.000000	33.000000	150.000000	23.000000	51.000000	0.000000
50%	1.000000	8.000000	6.000000	1.000000	28.000000	47.000000	206.000000	25.500000	69.000000	0.000000
75%	2.000000	12.000000	10.000000	1.000000	43.000000	77.000000	331.000000	29.300000	85.000000	0.000000
max	87.000000	94.000000	65.000000	11.000000	171.000000	557.000000	360.000000	40.900000	99.000000	29.000000

ภาพประกอบ 22 แสดงการสำรวจข้อมูลสถิติเบื้องต้นข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

	NO2	SO2	O3	PM10	Wind dir	Temp	Rel hum	Rain	Wind speed	PM
12384.00000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000	12384.000000
7.67216	0.876299	32.025890	60.414032	221.668535	26.088361	66.669971	0.075486	0.550041	37.889518	
7.10341	0.905685	22.371689	41.703899	113.823416	4.886454	21.181902	0.753198	0.722371	33.388866	
0.00000	0.000000	0.000000	4.000000	0.000000	11.100000	0.000000	0.000000	0.000000	1.000000	
3.00000	0.000000	15.000000	33.000000	150.000000	23.000000	51.000000	0.000000	0.200000	17.000000	
6.00000	1.000000	28.000000	47.000000	206.000000	25.500000	69.000000	0.000000	0.400000	27.000000	
10.00000	1.000000	43.000000	77.000000	331.000000	29.300000	85.000000	0.000000	0.800000	48.000000	
65.00000	11.000000	171.000000	557.000000	360.000000	40.900000	99.000000	29.000000	63.000000	470.000000	

ภาพประกอบ 23 แสดงการสำรวจข้อมูลสถิติเบื้องต้นข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (ต่อ)

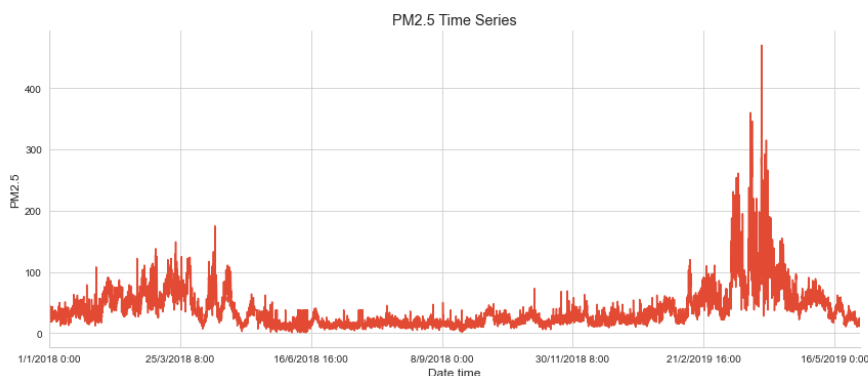
	CO	NO	NOX	NO2	SO2	PM10	Wind dir	Temp	Rel hum	Wind speed	PM
count	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000	8016.000000
mean	0.843849	8.117860	28.907410	21.035410	1.356488	45.565398	126.131411	29.379800	66.555875	0.409471	34.692320
std	0.344509	8.750932	20.571597	16.293275	0.546424	38.101080	102.074713	4.799599	23.765272	0.267653	31.380857
min	0.300000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	18.000000	10.000000	0.100000	1.000000
25%	0.600000	2.000000	14.000000	10.000000	1.000000	22.000000	36.000000	26.000000	48.000000	0.200000	15.000000
50%	0.800000	6.000000	24.000000	16.000000	1.000000	36.000000	116.000000	29.000000	68.000000	0.400000	32.000000
75%	0.900000	10.000000	37.000000	26.000000	2.000000	57.000000	200.000000	32.000000	89.000000	0.500000	40.000000
max	3.400000	78.000000	159.000000	134.000000	6.000000	441.000000	360.000000	44.400000	98.000000	2.100000	372.000000

ภาพประกอบ 24 แสดงการสำรวจข้อมูลสถิติเบื้องต้นข้อมูลโรงเรียนยุพราชวิทยาลัย

ตาราง 2 แสดงตารางเปรียบเทียบค่าเฉลี่ยของแต่ละตัวแปรของข้อมูลศูนย์ราชการจังหวัด
เชียงใหม่และโรงเรียนยุพราชวิทยาลัย

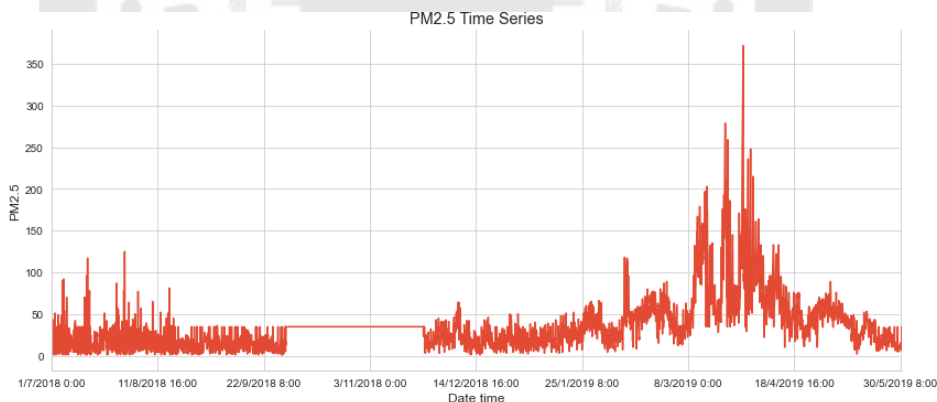
ตัวแปร	ศูนย์ราชการจังหวัดเชียงใหม่ (ค่าเฉลี่ย)	โรงเรียนยุพราชวิทยาลัย (ค่าเฉลี่ย)
CO	-	0.99
NO	1.94	7.97
NOX	9.64	26.03
NO2	7.67	18.47
SO2	0.87	3.42
O3	0.82	-
PM10	60.41	44.72
Wind dir	221.66	134.56
Temp	26.09	28.81
Rel hum	66.67	67.14
Rain	0.08	-
Wind speed	0.55	0.32
PM2.5	37.89	25.82

จากตารางที่ 2 ตารางเปรียบเทียบค่าเฉลี่ยของแต่ละตัวแปรของข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ และโรงเรียนยุพราชวิทยาลัย จากทั้งหมด 13 ตัวแปร จะมี 10 ตัวแปรที่มีเหมือนกันทั้งสองชุดข้อมูล โดยที่ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ พบว่ามีตัวแปร PM10, Wind dir, Wind speed และ PM2.5 ที่มีค่าเฉลี่ยที่สูงกว่าชุดข้อมูลโรงเรียนยุพราชวิทยาลัย ในส่วนชุดข้อมูลโรงเรียนยุพราชวิทยาลัย พบว่ามีตัวแปร NO, NOX, NO2, SO2, Temp และ Rel hum ที่มีค่าเฉลี่ยที่สูงกว่าข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ เนื่องจากข้อมูล PM10 และ PM2.5 ของศูนย์ราชการจังหวัดเชียงใหม่มีค่าเฉลี่ยที่สูงกว่าข้อมูลโรงเรียนยุพราชวิทยาลัย ดังนั้น ตัวแปร Wind dir และ Wind speed อาจมีส่วนทำให้ PM10 และ PM2.5 เพิ่มขึ้น



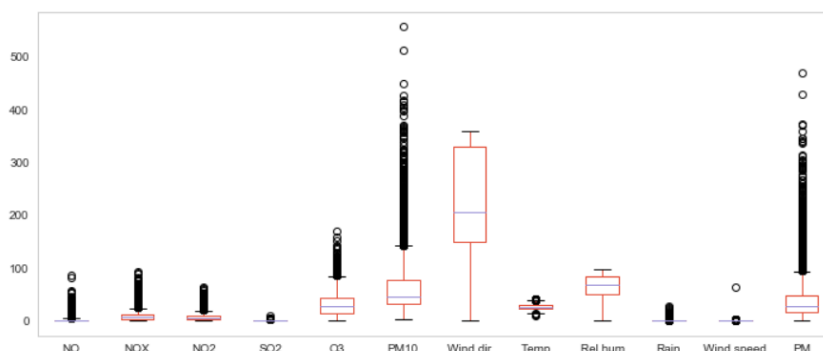
ภาพประกอบ 25 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ (หลังทำความสะอาดข้อมูล)

จากภาพประกอบ 25 ได้ทำการวิเคราะห์แนวโน้มที่เกิดขึ้นของความเข้มข้นฝุ่นละออง PM2.5 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่หลังจากทำความสะอาดข้อมูลเนื่องจากจำนวน Missing value ของข้อมูลศูนย์ราชการจังหวัดเชียงใหม่มีไม่ได้มากนัก ทำให้ลักษณะกราฟไม่แตกต่างไปจากเดิม



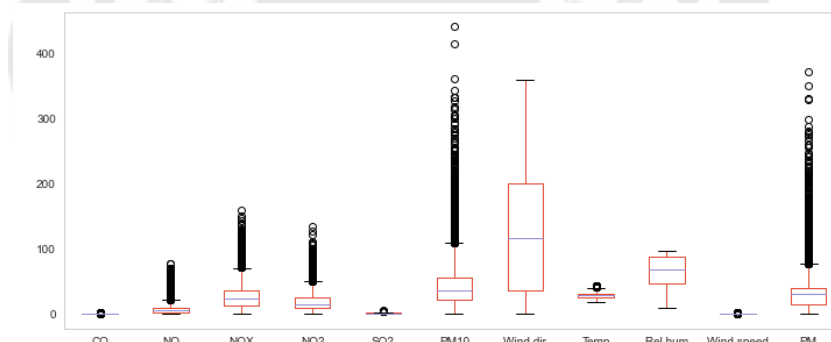
ภาพประกอบ 26 แสดงการสำรวจข้อมูลแนวโน้ม PM2.5 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย (หลังทำความสะอาดข้อมูล)

จากภาพประกอบ 26 ได้ทำการวิเคราะห์แนวโน้มที่เกิดขึ้นของความเข้มข้นฝุ่นละออง PM2.5 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย หลังจากทำการทำความสะอาดข้อมูลเรียบร้อยแล้ว พบว่าในช่วงเดือนตุลาคม ถึงพฤศจิกายน ค.ศ. 2018 ที่มีจำนวน Missing values ค่อนข้างมากได้ถูกแทนค่าด้วยค่าเฉลี่ยตลอดช่วง ทำให้ลักษณะกราฟมีความแตกต่างไปจากเดิม



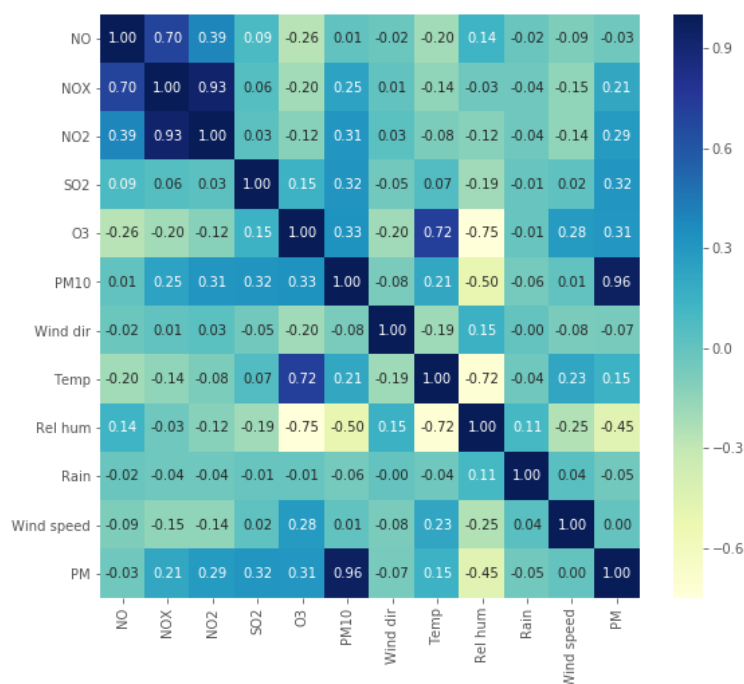
ภาพประกอบ 27 แสดงกราฟ Boxplot ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่
(หลังทำความสะอาดข้อมูล)

จากภาพประกอบ 27 ลักษณะของกราฟ Boxplot หลังทำความสะอาดข้อมูล เนื่องจากจำนวน Missing value ของข้อมูลศูนย์ราชการจังหวัดเชียงใหม่มีไม่ได้มากนัก ทำให้ลักษณะกราฟไม่แตกต่างไปจากเดิม



ภาพประกอบ 28 แสดงกราฟ Boxplot ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย
(หลังทำความสะอาดข้อมูล)

จากภาพประกอบ 28 กราฟ Boxplot ของข้อมูลโรงเรียนยุพราชวิทยาลัยหลังทำความสะอาดข้อมูล ลักษณะกราฟ Boxplot ไม่ได้แตกต่างจากเดิมมากนัก จะมีเปลี่ยนไปที่บริเวณของค่าเฉลี่ยของ PM2.5 ที่จะเปลี่ยนได้เห็นได้ชัดเนื่องจากมีค่า Missing value ค่อนข้างมาก ส่วนตัวแปรตัวอื่นถูกปรับค่าเฉลี่ยเพิ่มขึ้นลงน้อย แต่ลักษณะกราฟไม่ได้แตกต่างจากเดิมมากนัก

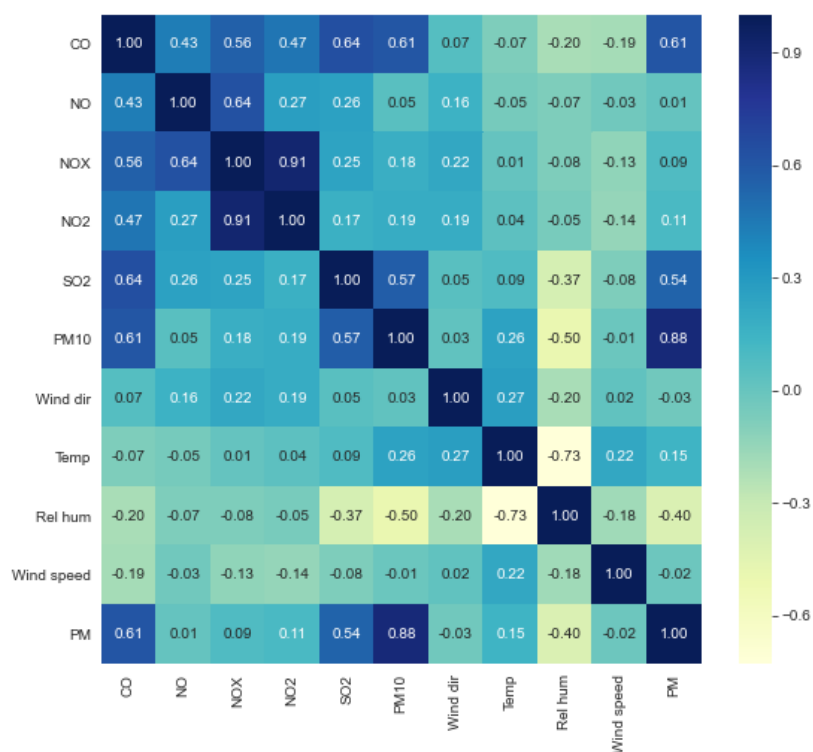


ภาพประกอบ 29 แสดงค่า Correlation ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

จากภาพประกอบ 29 แสดงค่า Correlation ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ แสดงค่า Correlation เพื่อหาความสัมพันธ์ระหว่างตัวแปรค่าความเข้มข้น PM2.5 เปรียบเทียบกับตัวแปรอื่น ๆ พบว่าความสัมพันธ์ในทางบวก และความสัมพันธ์ในทางลบ ได้แก่

- ความสัมพันธ์ในทางบวก พบว่า PM10 มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ 0.96 รองลงมา คือ SO2 มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ 0.32 และ O3 มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ 0.31

- ความสัมพันธ์ในทางลบ พบว่า Rel hum มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ -0.45 รองลงมา คือ Wind dir มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ -0.07 และ Rain มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ -0.05



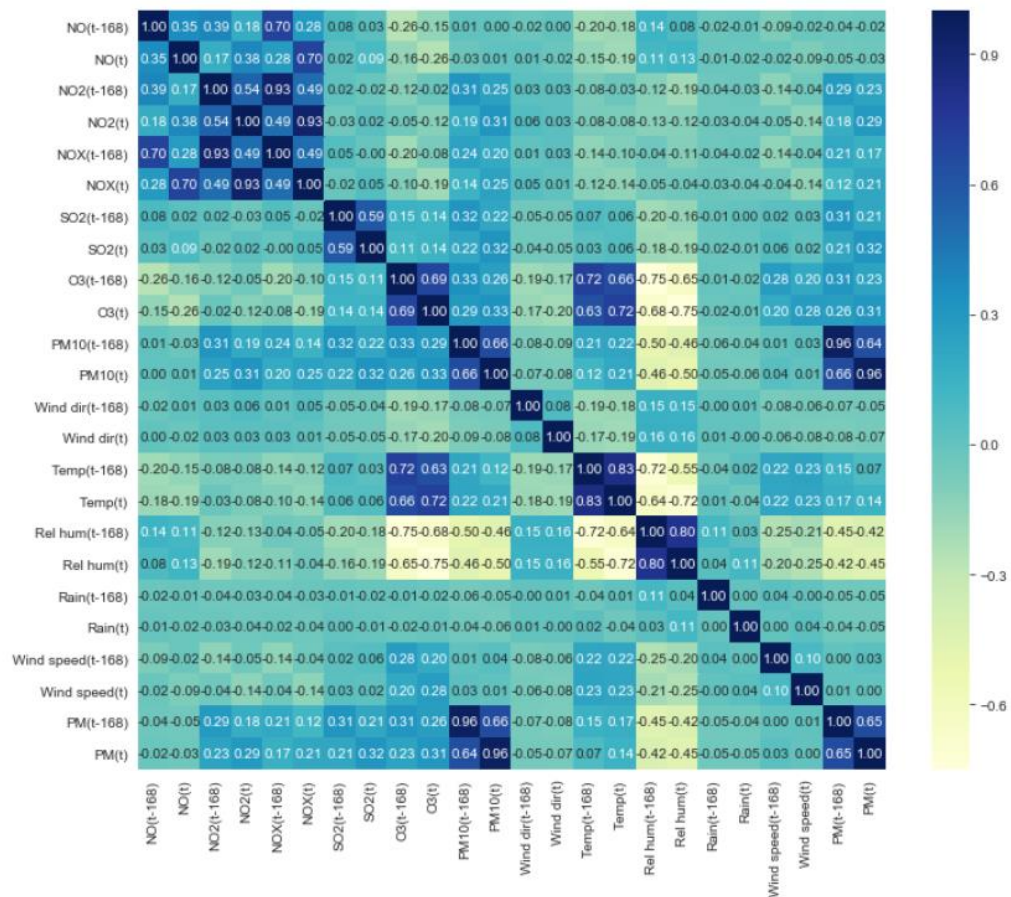
ภาพประกอบ 30 แสดงค่า Correlation ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 30 แสดงค่า Correlation ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย แสดงค่า Correlation เพื่อหาความสัมพันธ์ระหว่างตัวแปรค่าความเข้มข้น PM2.5 เปรียบเทียบกับตัวแปรอื่น ๆ พบว่าความสัมพันธ์ในทางบวก และความสัมพันธ์ในทางลบ ได้แก่

- ความสัมพันธ์ในทางบวก พบว่า PM10 มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ 0.88 รองลงมา คือ CO มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ 0.61 และ SO2 มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ 0.54

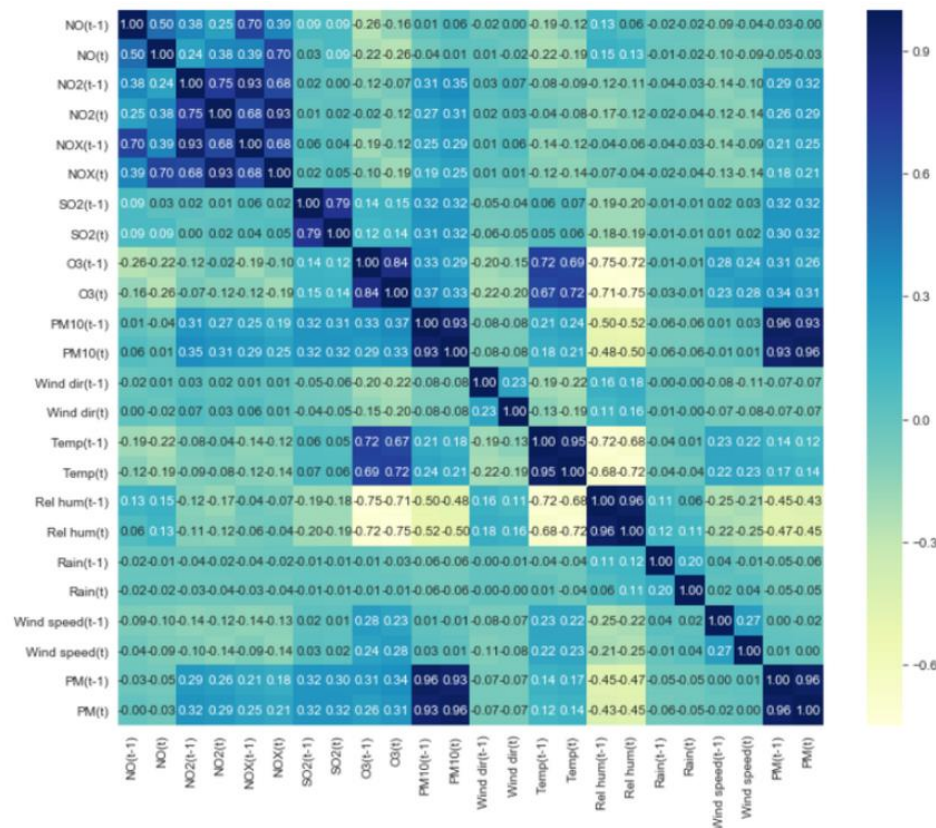
- ความสัมพันธ์ในทางลบ พบว่า Rel hum มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ -0.40 รองลงมา คือ Wind dir มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ -0.03 และ Wind speed มีค่า Correlation สัมพันธ์กับ PM2.5 อยู่ที่ -0.02

จากข้อมูลทั้งสองชุดเป็นที่ชัดเจนว่า PM10 มีความสัมพันธ์ในทางบวกมากที่สุดกับ PM2.5 เนื่องจากเป็นข้อมูลความเข้มข้นฝุ่นละอองเหมือนกัน รองลงมาเป็น SO2 มีความสัมพันธ์ในทางบวกทั้งสองชุดข้อมูลเหมือนกัน และ Rel hum มีความสัมพันธ์ในทางลบทั้งสองชุดเหมือนกัน



ภาพประกอบ 31 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ t-168 ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

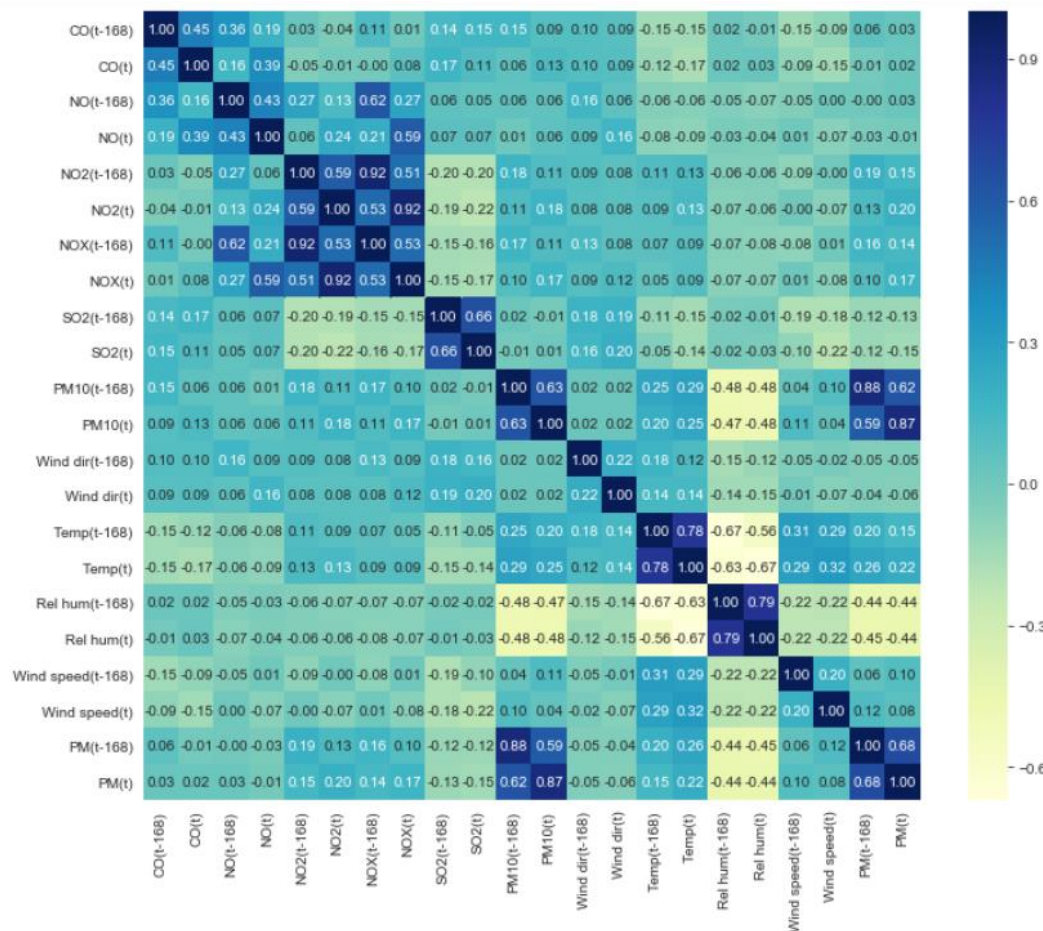
จากภาพประกอบ 31 แสดงค่า Correlation ของแต่ละตัวแปรระหว่างช่วงเวลา t กับช่วงเวลา t-168 ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ เพื่อทำการสำรวจข้อมูลของตัวแปร PM2.5 ในช่วงระยะเวลายาวโดยที่ระยะเวลา t กับ t-168 ตัวแปร PM2.5 จะมีความสัมพันธ์กันอย่างไร พบว่า PM2.5 (t-168) มีค่า Correlation สัมพันธ์กับ PM2.5 (t) อยู่ที่ 0.65 ดังนั้น แสดงว่าตัวแปร PM2.5 ในแต่ละช่วงเวลามีความสัมพันธ์กันในทางบวกเป็นไปในทิศทางเดียวกัน



ภาพประกอบ 32 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ $t-1$
ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

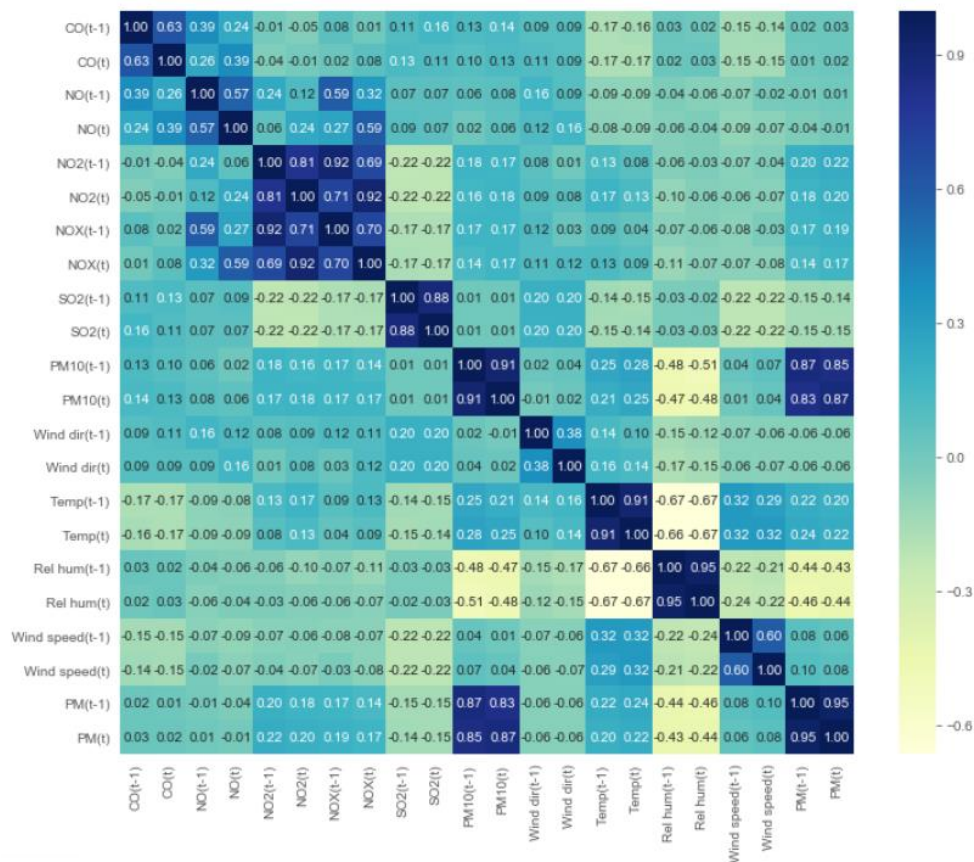
จากภาพประกอบ 32 แสดงค่า Correlation ของแต่ละตัวแปรระหว่างช่วงเวลา t กับช่วงเวลา $t-1$ ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ เพื่อทำการสำรวจข้อมูลของตัวแปร PM2.5 ในช่วงระยะเวลาสั้นโดยที่ระยะเวลา t กับ $t-1$ ตัวแปร PM2.5 จะมีความสัมพันธ์กันอย่างไรพบว่า PM2.5 ($t-1$) มีค่า Correlation สัมพันธ์กับ PM2.5 (t) อยู่ที่ 0.96 ดังนั้น แสดงว่าตัวแปร PM2.5 ในแต่ละช่วงเวลามีความสัมพันธ์กันในทางบวกเป็นไปในทิศทางเดียวกัน

สำหรับข้อมูลศูนย์ราชการจังหวัดเชียงใหม่เป็นที่ชัดเจนว่า PM2.5 (t) มีความสัมพันธ์ไปในทางบวกกับ PM2.5 ($t-1$) และ PM2.5 ($t-168$) โดยที่มีความแตกต่างกันที่ PM2.5 ($t-1$) มีค่า Correlation อยู่ที่ 0.96 ส่วน PM2.5 ($t-168$) มีค่า Correlation อยู่ที่ 0.65 ทำให้สรุปได้ว่าช่วงระยะเวลาสั้น ($t-1$) ค่าความเข้มข้นฝุ่นละออง PM2.5 จะมีความสัมพันธ์มากกว่าช่วงระยะเวลาที่ยาว ($t-168$)



ภาพประกอบ 33 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ t-168 ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 33 แสดงค่า Correlation ของแต่ละตัวแปรระหว่างช่วงเวลา t กับช่วงเวลา t-168 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย พบว่า PM2.5 (t-168) มีค่า Correlation สัมพันธ์กับ PM2.5 (t) อยู่ที่ 0.68 ดังนั้น แสดงว่าตัวแปร PM2.5 ในแต่ละช่วงเวลามีความสัมพันธ์กันในทางบวกเป็นไปในทิศทางเดียวกัน



ภาพประกอบ 34 แสดงค่า Correlation ของแต่ละ feature ระหว่างช่วงเวลา t กับ t-1 ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 34 แสดงค่า Correlation ของแต่ละตัวแปรระหว่างช่วงเวลา t กับช่วงเวลา t-1 ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย พบว่า PM2.5 (t-1) มีค่า Correlation สัมพันธ์กับ PM2.5 (t) อยู่ที่ 0.95 ดังนั้น แสดงว่าตัวแปร PM2.5 ในแต่ละช่วงเวลามีความสัมพันธ์กันในทางบวกเป็นไปในทิศทางเดียวกัน

สำหรับข้อมูลโรงเรียนยุพราชวิทยาลัยเป็นที่ชัดเจนว่า PM2.5 (t) มีความสัมพันธ์ไปในทางบวกกับ PM2.5 (t-1) และ PM2.5 (t-168) โดยที่มีความแตกต่างกันที่ PM2.5 (t-1) มีค่า Correlation อยู่ที่ 0.95 ส่วน PM2.5 (t-168) มีค่า Correlation อยู่ที่ 0.68 ทำให้สรุปได้ว่าช่วงระยะเวลาสั้น (t-1) ค่าความเข้มข้นฝุ่นละออง PM2.5 จะมีความสัมพันธ์มากกว่าช่วงระยะเวลาที่ยาว (t-168) สรุปได้ว่าทั้งสองชุดข้อมูลทั้งข้อมูลศูนย์ราชการจังหวัดเชียงใหม่และข้อมูลโรงเรียนยุพราชวิทยาลัย มีค่า Correlation ในทิศทางเดียวกันโดยที่ความสัมพันธ์ในระยะใกล้ที่ช่วงเวลา t กับช่วงเวลา t-1 มีความสัมพันธ์มากกว่าช่วงเวลาที่ระยะที่ไกลกัน คือ ช่วงเวลา t กับ t-168

3.5 การสร้างแบบจำลอง Long Short-Term Memory (LSTM)

3.5.1 การ Shift ข้อมูลย้อนหลัง

การ Shift ข้อมูลย้อนหลังเป็นการเลื่อนค่าของ Feature กับ Target (PM2.5) เพื่อให้ข้อมูลย้อนหลัง 7 วัน สำหรับทำนายข้อมูลการทำนายข้อมูล PM2.5 ใน 24 ชั่วโมงล่วงหน้า โดยทุกตัวแปรถูก Shift ย้อนหลัง 7 วัน หรือ lag = 7 วัน หรือ 168 ชั่วโมง

Date time	NO(t-168)	NOX(t-168)	NO2(t-168)	SO2(t-168)	O3(t-168)	PM10(t-168)	Wind dirt(t-168)	Temp(t-168)	Rel hum(t-168)	Rain(t-168)	...	NO2(t+23)	SO2(t+23)	O3(t+23)	PM10(t+23)
8/1/2018 0:00	0.117021	0.117021	0.153846	0.090909	0.052632	0.113924	0.913889	0.328859	0.868687	0.0	...	0.261538	0.090909	0.011696	0.137432
8/1/2018 1:00	0.127660	0.127660	0.169231	0.090909	0.046784	0.065099	0.866667	0.355705	0.868687	0.0	...	0.076923	0.090909	0.040936	0.061483
8/1/2018 2:00	0.148936	0.148936	0.169231	0.090909	0.017544	0.077758	0.850000	0.322148	0.919192	0.0	...	0.138462	0.000000	0.023392	0.054250
8/1/2018 3:00	0.095745	0.095745	0.138462	0.090909	0.023392	0.083183	0.311111	0.312081	0.929293	0.0	...	0.123077	0.000000	0.011696	0.048825
8/1/2018 4:00	0.117021	0.117021	0.138462	0.000000	0.011696	0.072333	0.863889	0.308725	0.939394	0.0	...	0.046154	0.000000	0.023392	0.054250

ภาพประกอบ 35 ตัวอย่างการ Shift ย้อนหลัง lag 7 วัน หรือ 168 ชั่วโมง

3.5.2 การแบ่งข้อมูล (Split data)

ทำการแบ่งข้อมูลออกเป็น 2 ชุด ได้แก่ ชุดสำหรับในการฝึกฝน (Train dataset) 90 เปอร์เซ็นต์ และชุดทดสอบ (Test dataset) 10 เปอร์เซ็นต์

1. ชุดข้อมูลสำหรับในการฝึกฝน ใช้สำหรับนำเข้าแบบจำลองเพื่อฝึกฝนในแบบจำลอง (Train Model)

2. ชุดข้อมูลสำหรับทดสอบ ใช้สำหรับการทดสอบว่าแบบจำลองทำงานได้ดีแค่ไหนกับข้อมูลที่ไม่เคยเจอมาก่อน ไม่เคยผ่านการอบรมในแบบจำลอง

โดยชุดแรกข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ แบ่งเป็นชุดสำหรับฝึกฝน ตั้งแต่วันที่ 1 มกราคม 2018 เวลา 0:00 ถึง 10 เมษายน 2019 เวลา 3:00 และชุดทดสอบ ตั้งแต่วันที่ 10 เมษายน 2019 เวลา 4:00 ถึง 31 พฤษภาคม 2019 เวลา 23:00 8,016 record

ข้อมูลชุดที่สองข้อมูลโรงเรียนยุพราชวิทยาลัย แบ่งเป็นชุดสำหรับฝึกฝน ตั้งแต่วันที่ 1 กรกฎาคม 2018 เวลา 0:00 ถึง 28 เมษายน 2019 เวลา 9:00 และชุดทดสอบ ตั้งแต่วันที่ 28 เมษายน 2019 เวลา 10:00 ถึง 31 พฤษภาคม 2019 เวลา 23:00 12,384 record

3.5.3 การแปลงข้อมูลโดยใช้ Min-Max Scaler

ทำการแปลงข้อมูล หรือลดขนาดข้อมูลให้เป็นมาตรฐาน โดยใช้ Min-Max Scaler เป็นวิธีที่ง่ายที่สุด ที่จะปรับช่วงของข้อมูล เพื่อให้ข้อมูลอยู่ใน Scale ที่เหมาะสมในช่วง $[0,1]$ ตามค่าต่ำสุดและสูงสุดของชุดข้อมูล ดังสมการ ดังนี้ (21)

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (21)$$

โดยที่

x' คือ ค่าปกติ

x คือ ค่าดั้งเดิม

3.5.4 การปรับพารามิเตอร์ (Optimizing LSTM)

ในแบบจำลอง LSTM ประกอบด้วยค่าปรับอัลกอริทึม Adagrad, RMSprop และ Adam โดยใช้ทั้ง 3 อัลกอริทึมให้ค่า Epochs, Neuron และค่า Batch size แต่ละตัวมีค่าเท่ากันเพื่อใช้ในการเปรียบเทียบพารามิเตอร์ของแบบจำลอง

- Epoch คือ การที่นำข้อมูลเข้าฝึกอบรม (train model) ข้อมูลที่เข้าแบบจำลองโดยใช้ข้อมูลใน dataset ที่มีอยู่จนครบทั้งหมดทุกตัวเป็นจำนวน 1 รอบ
- Neuron คือ จำนวนโหนดที่ใช้ในแบบจำลอง
- Batch size คือ การแบ่งข้อมูลให้เป็นกลุ่มที่เล็กลง เนื่องจากแบบจำลองไม่สามารถนำข้อมูลทั้งหมดที่ใส่เข้าไปเพื่อฝึกอบรมแบบจำลอง (Train model) จึงจำเป็นต้องแบ่งออกเป็นกลุ่มเล็ก ซึ่งสามารถกำหนดได้ว่าการให้ Batch size ขนาดเท่าไร

บทที่ 4

ผลการศึกษา

ในการวิจัยการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 ใน 24 ชั่วโมงข้างหน้า ซึ่งใช้ข้อมูลฝุ่นละออง PM2.5 ข้อมูลทางอุตุนิยมวิทยาบริเวณพื้นที่สถานีศูนย์ราชการจังหวัดเชียงใหม่ และสถานีโรงเรียนยุพราชจังหวัดเชียงใหม่ โดยใช้เทคนิคการเรียนรู้เชิงลึก ผู้วิจัยได้ดำเนินการวิจัยโดยการศึกษาตามขอบเขตและขั้นตอนต่าง ๆ ตลอดจนการวัดประสิทธิภาพ เพื่อให้บรรลุจุดประสงค์ของการวิจัยที่ได้กำหนดไว้ ได้ดังนี้

1. ผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีศูนย์ราชการจังหวัดเชียงใหม่
2. ผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีโรงเรียนยุพราชจังหวัดเชียงใหม่

4.1 ผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีศูนย์ราชการจังหวัดเชียงใหม่

จากการสร้างแบบจำลอง LSTM โดยใช้ข้อมูลกรมอุตุนิยมวิทยาระหว่างสถานีศูนย์ราชการจังหวัดเชียงใหม่ ใช้ข้อมูลย้อนหลัง 7 วัน เพื่อใช้ในการพยากรณ์ฝุ่นละออง PM2.5 ใน 24 ชั่วโมงล่วงหน้า โดยใช้ตัววัดประสิทธิภาพแบบจำลอง LSTM คือ ค่า RMSE, MAE และ MAPE แบบจำลองได้กำหนดให้ใช้เปรียบเทียบ Optimizer ทั้ง 3 ตัว ได้แก่ Adam, RMSprop และ AdaGrad โดยที่กำหนดให้ Epochs = 100 และ Batch size = 72 ผลลัพธ์ของการพยากรณ์พบว่าในการพยากรณ์ระยะสั้น 1 ชั่วโมงข้างหน้า Optimizer ทั้ง 3 แบบมีประสิทธิภาพที่ไม่แตกต่างกัน แต่ในการพยากรณ์ล่วงหน้า 24 ชั่วโมง พบว่า Adam ทำงานได้ดีที่สุด รองลงมาเป็น RMSprop และ AdaGrad ตามลำดับ ดังตัวอย่างตารางผลลัพธ์ของการเปรียบเทียบการปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีศูนย์ราชการจังหวัดเชียงใหม่

ตาราง 3 แสดงผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีสุนัขราชการจังหวัดเชียงใหม่

Time Step	AdaGrad			RMSprop			Adam		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
t+1	7.89	5.69	13.02%	7.11	5.26	13.04%	7.31	5.42	11.43%
t+2	8.92	6.56	14.58%	8.66	6.42	15.66%	8.99	6.74	14.03%
t+3	10.23	7.68	16.13%	9.09	6.74	16.74%	9.93	7.47	16.36%
t+4	11.04	8.33	17.53%	10.25	7.67	18.74%	11.49	8.73	17.60%
t+5	11.40	8.50	18.91%	11.36	8.52	18.10%	12.3	9.46	18.80%
t+6	12.18	9.25	18.69%	11.67	8.70	19.03%	12.85	9.92	18.99%
t+7	12.82	9.74	20.09%	11.40	8.59	20.31%	12.6	9.59	19.75%
t+8	12.70	9.70	19.58%	12.24	9.27	20.92%	13.15	10.13	20.38%
t+9	12.94	9.89	19.98%	12.45	9.31	21.90%	14.21	11.01	21.10%
t+10	13.25	10.14	20.32%	12.49	9.41	22.35%	13.64	10.36	22.17%
t+11	12.99	9.92	20.27%	12.00	9.12	22.15%	13.76	10.4	21.58%
t+12	13.37	10.21	20.47%	12.45	9.37	23.42%	13.69	10.35	22.64%
t+13	13.32	10.11	21.25%	12.67	9.53	22.64%	13.44	10.22	22.83%
t+14	12.83	9.82	21.13%	12.74	9.72	22.87%	12.87	9.67	23.59%
t+15	13.28	10.13	20.68%	12.69	9.65	23.91%	13.06	9.95	24.62%
t+16	12.81	9.82	21.71%	12.94	9.91	23.84%	12.67	9.53	24.29%
t+17	13.05	9.87	22.61%	12.47	9.42	23.46%	12.6	9.47	25.32%
t+18	12.58	9.62	21.37%	12.35	9.35	25.78%	12.43	9.43	24.65%
t+19	12.35	9.53	20.88%	11.90	9.19	25.75%	12.15	9.22	23.69%
t+20	11.76	8.99	20.85%	12.08	9.07	25.44%	12.13	9.14	23.79%
t+21	11.96	9.13	20.09%	12.05	9.23	25.71%	12.21	9.18	21.70%
t+22	11.75	8.94	20.47%	12.10	9.16	25.72%	11.91	8.99	21.63%
t+23	11.34	8.61	20.03%	12.11	9.21	25.49%	11.47	8.73	20.99%
t+24	11.83	9.05	20.42%	12.11	9.16	23.26%	11.48	8.78	21.65%
ค่าเฉลี่ย	12.02	9.13	19.63%	11.64	8.79	21.93%	12.18	9.25	20.98%

จากตาราง 3 โดยใช้ข้อมูลสถานีศูนย์ราชการจังหวัดเชียงใหม่ พบว่าในการพยากรณ์ 1 ชั่วโมง Optimizer ทั้ง 3 ตัวมีประสิทธิภาพการพยากรณ์ที่ไม่แตกต่างกันอย่างมีนัยสำคัญ แต่ในการพยากรณ์ 24 ชั่วโมงล่วงหน้า Adam และ AdaGrad มีประสิทธิภาพมากกว่า RMSprop โดยที่ในการพยากรณ์ 24 ชั่วโมงล่วงหน้า Adam มีค่า RMSE เท่ากับ $11.48 \mu\text{g}/\text{m}^3$, ค่า MAE เท่ากับ $8.78 \mu\text{g}/\text{m}^3$ และค่า MAPE เท่ากับ 21.65 เปอร์เซ็นต์ ส่วน AdaGrad มีค่า RMSE เท่ากับ $11.83 \mu\text{g}/\text{m}^3$, ค่า MAE เท่ากับ $9.05 \mu\text{g}/\text{m}^3$ และค่า MAPE เท่ากับ 20.42 เปอร์เซ็นต์ มีความชัดเจนว่าสำหรับข้อมูลจากสถานีศูนย์ราชการจังหวัดเชียงใหม่นี้ แบบจำลองที่สร้างด้วย Optimizer ทั้งสามสามารถเรียนรู้การพยากรณ์ค่าความเข้มข้นในระยะสั้น 1 ชั่วโมงล่วงหน้าได้ดีกว่าการพยากรณ์ค่าความเข้มข้นในระยะยาว 24 ชั่วโมง แต่พบว่าการพยากรณ์ค่าความเข้มข้นในระยะยาว 24 ชั่วโมงล่วงหน้าได้ดีกว่าระยะกลาง

4.2 ผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีโรงเรียนยุพราชวิทยาลัย

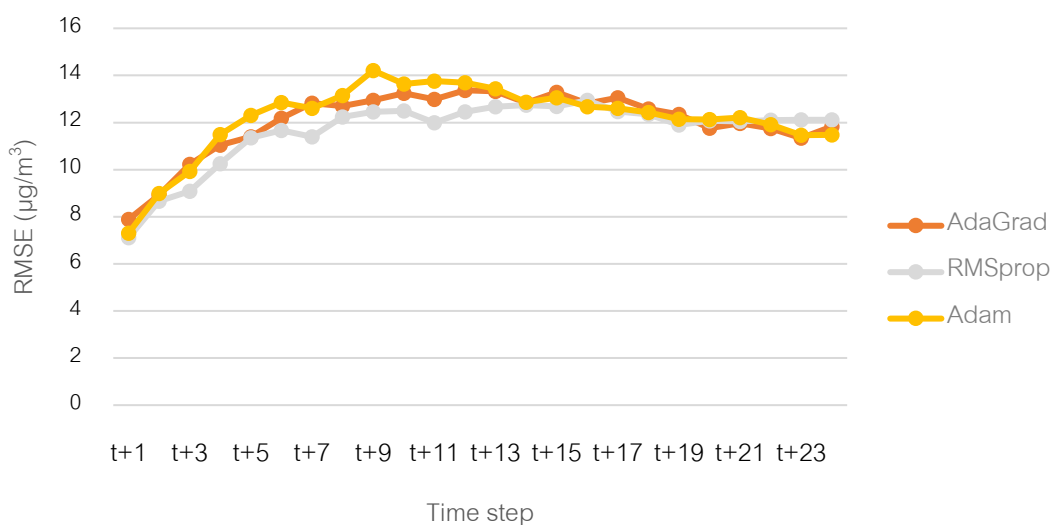
จากการสร้างแบบจำลอง LSTM โดยใช้ข้อมูลกรมอุตุนิยมวิทยาระหว่างสถานีโรงเรียนยุพราชวิทยาลัยจังหวัดเชียงใหม่ ใช้ข้อมูลย้อนหลัง 7 วัน เพื่อใช้ในการพยากรณ์ฝุ่นละออง PM2.5 ใน 24 ชั่วโมงล่วงหน้า โดยใช้ตัววัดประสิทธิภาพแบบจำลอง LSTM คือ ค่า RMSE, MAE และ MAPE แบบจำลองได้กำหนดให้ใช้เปรียบเทียบ Optimizer ทั้ง 3 ตัว ได้แก่ Adam, RMSprop และ AdaGrad โดยที่กำหนดให้ Epochs = 100 และ Batch size = 72 ผลลัพธ์ของการพยากรณ์พบว่า Optimizer ทั้ง 3 แบบมีประสิทธิภาพที่แตกต่างกัน โดย RMSprop ทำงานได้ดีที่สุด ทั้งในระยะสั้น 1 ชั่วโมงล่วงหน้า และระยะยาว 24 ชั่วโมงล่วงหน้า รองลงมาเป็น AdaGrad และ Adam ตามลำดับ ดังตัวอย่างตารางผลลัพธ์ของการเปรียบเทียบการปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีโรงเรียนยุพราชวิทยาลัยจังหวัดเชียงใหม่

ตาราง 4 แสดงผลลัพธ์การปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเลือกใช้ Optimizer ของสถานีโรงเรียนนุพราชาวิทยาลัย

Time Step	AdaGrad			RMSprop			Adam		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
t+1	6.76	4.93	26.22%	6.52	4.77	25.25%	6.82	5.17	27.73%
t+2	7.96	5.98	31.32%	7.59	5.63	24.88%	8.06	6.19	29.57%
t+3	9.15	6.92	35.10%	8.68	6.55	31.74%	9.38	7.2	34.43%
t+4	10.01	7.66	40.63%	9.05	6.84	34.00%	10.28	7.83	37.42%
t+5	10.45	8.15	42.31%	9.94	7.54	37.62%	11.07	8.52	40.45%
t+6	10.83	8.41	43.38%	10.04	7.51	34.43%	11.58	8.94	40.47%
t+7	11.24	8.87	45.76%	9.94	7.60	36.82%	12.01	9.28	43.51%
t+8	10.8	8.43	41.31%	10.66	8.16	39.06%	12.81	10.07	46.00%
t+9	10.53	8.26	41.65%	10.21	7.81	37.74%	12.5	9.83	45.79%
t+10	11.11	8.75	43.45%	10.40	7.97	38.19%	12.44	9.88	46.29%
t+11	10.53	8.28	40.58%	10.18	7.81	34.99%	12.47	9.89	45.20%
t+12	11.19	9.01	45.53%	10.11	7.90	39.43%	12.67	10.31	52.74%
t+13	10.59	8.49	44.38%	10.15	7.90	37.54%	12.08	9.8	50.27%
t+14	10.15	8.02	41.01%	9.93	7.79	39.05%	11.72	9.35	43.82%
t+15	10.93	8.76	45.76%	9.90	7.76	37.07%	11.99	9.77	49.69%
t+16	11.45	9.27	47.74%	10.57	8.35	43.01%	11.53	9.38	46.98%
t+17	11.71	9.56	50.32%	9.93	7.75	38.85%	11.71	9.59	49.84%
t+18	11.63	9.4	50.70%	10.04	7.84	40.90%	10.8	8.79	45.81%
t+19	12	9.69	50.90%	9.57	7.42	35.97%	11.31	9.28	51.54%
t+20	10.91	8.8	46.76%	10.01	7.68	40.34%	10.84	8.81	48.46%
t+21	11.42	9.26	50.08%	10.14	7.82	42.06%	10.68	8.52	46.83%
t+22	10.64	8.36	45.95%	10.38	7.93	41.42%	11.63	9.18	53.67%
t+23	11.62	9.29	51.95%	10.04	7.62	40.95%	10.53	8.17	45.59%
t+24	10.96	8.57	49.09%	10.38	8.01	46.40%	11.02	8.59	49.04%
ค่าเฉลี่ย	10.61	8.38	43.83%	9.77	7.50	37.40%	11.16	8.85	44.63%

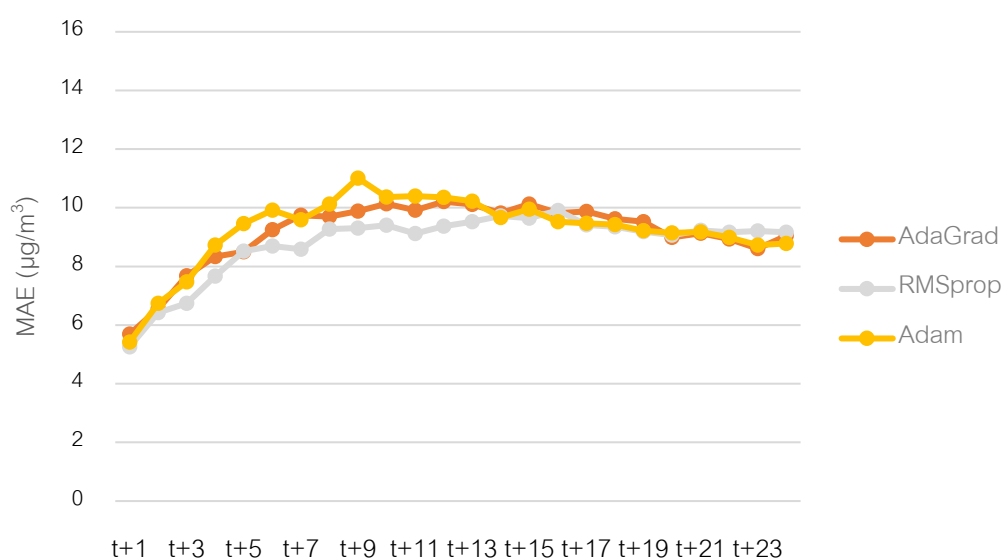
จากตาราง 4 โดยใช้ข้อมูลสถานีโรงเรียนยุพราชวิทยาลัย พบว่าในการพยากรณ์ระยะสั้น 1 ชั่วโมง Optimizer ทั้ง 3 ตัว มีประสิทธิภาพที่ไม่แตกต่างกัน แต่ในการพยากรณ์ตั้งแต่ 3 ชั่วโมงขึ้นไปพบว่า RMSprop ทำงานได้ดีที่สุด รองลงมาเป็น AdaGrad และ Adam ตามลำดับ โดยที่ RMSprop มีค่า RMSE เท่ากับ $10.38 \mu\text{g}/\text{m}^3$, ค่า MAE เท่ากับ $8.01 \mu\text{g}/\text{m}^3$ และ MAPE เท่ากับ 46.40 เปอร์เซ็นต์ มีความชัดเจนว่าสำหรับข้อมูลจากสถานียุพราชวิทยาลัยจังหวัดเชียงใหม่ RMSprop สามารถเรียนรู้แนวโน้มและการขึ้นลงของค่าความเข้มข้น PM2.5 ในรอบ 24 ชั่วโมงได้ค่อนข้างดีมากตลอดเวลา แต่ค่า MAPE ค่อนข้างสูงเนื่องจากฐานค่าจริงของค่า PM2.5 นั้นค่อนข้างต่ำ

ทำการพล็อตกราฟค่าความคลาดเคลื่อนทั้ง 3 ประเภทได้แก่ RMSE, MAE และ MAPE เพื่อเปรียบเทียบ Optimizer ทั้ง 3 ประเภทได้แก่ AdaGrad, RMSprop และ Adam โดยแบ่งเป็น 2 ชุดข้อมูล ได้แก่ ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ และชุดข้อมูลโรงเรียนยุพราชวิทยาลัย ทั้งหมด 6 กราฟ



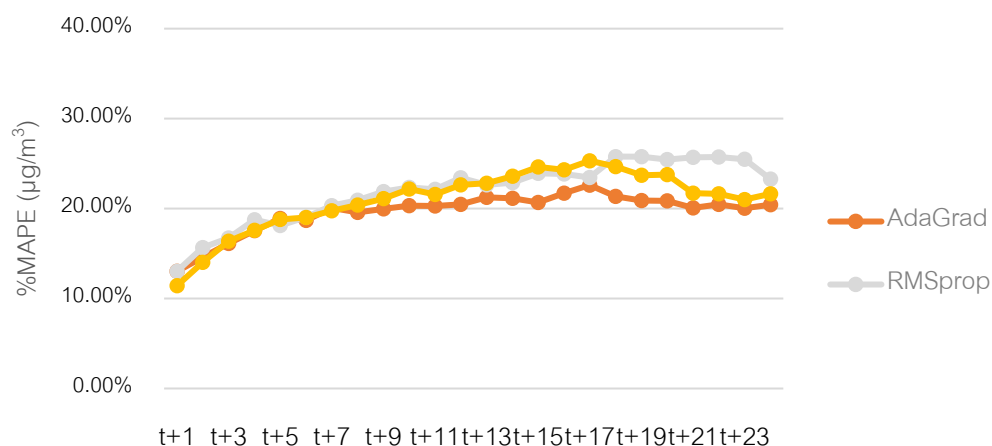
ภาพประกอบ 36 กราฟแสดงค่า RMSE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

จากภาพประกอบ 36 ในความคลาดเคลื่อน RMSE ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ จะเห็นได้ว่าในช่วงเวลาระยะสั้น $t+1$ Optimizer ทั้งสามประเภทมีค่า RMSE ที่ต่ำที่สุดที่ใกล้เคียงกัน ในช่วงระยะเวลากลาง $t+7$ และ $t+11$ จะเห็นว่า RMSprop มีค่า RMSE น้อยที่สุดเท่ากับ $11.40 \mu\text{g}/\text{m}^3$ และ $12.00 \mu\text{g}/\text{m}^3$ ที่ต่ำที่สุด แต่ในช่วงระยะเวลา $t+24$ จะเห็นว่า AdaGrad และ Adam มีค่า RMSE น้อยที่สุด คือ AdaGrad มีค่า RMSE เท่ากับ $11.83 \mu\text{g}/\text{m}^3$ และ Adam มีค่า RMSE เท่ากับ $11.48 \mu\text{g}/\text{m}^3$



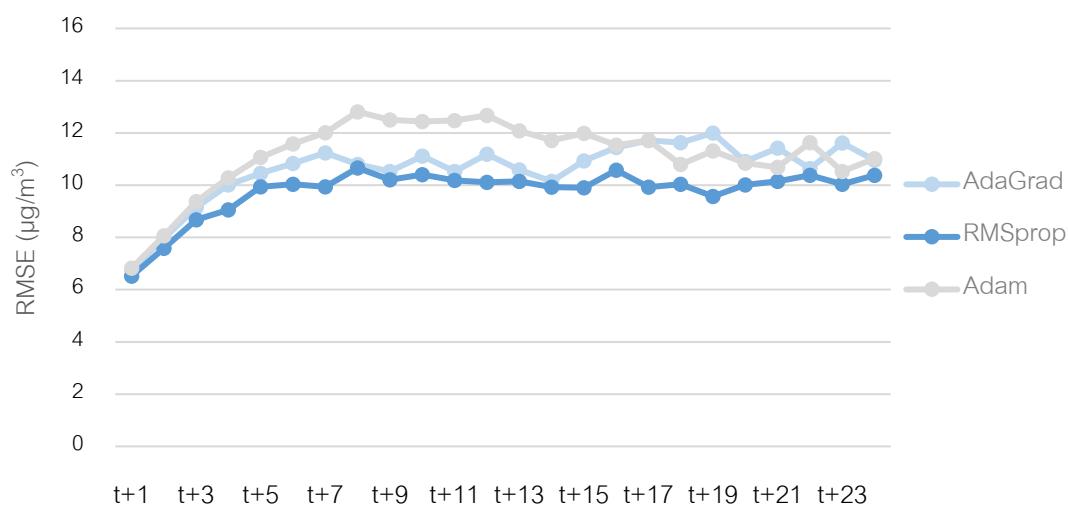
ภาพประกอบ 37 กราฟแสดงค่า MAE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

จากภาพประกอบ 37 ในความคลาดเคลื่อน MAE ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ จะเห็นได้ว่าในช่วงเวลาระยะสั้น $t+1$ Optimizer ทั้งสามประเภทมีค่า MAE ที่ต่ำที่สุดที่ใกล้เคียงกัน ในช่วงระยะเวลากลาง $t+7$ และ $t+11$ จะเห็นว่า RMSprop มีค่า MAE = $8.59 \mu\text{g}/\text{m}^3$ และ $9.12 \mu\text{g}/\text{m}^3$ ที่ต่ำที่สุด แต่ในช่วงระยะเวลา $t+24$ จะเห็นว่า AdaGrad และ RMSprop มีค่า MAE น้อยที่สุด คือ AdaGrad มีค่า MAE = $9.05 \mu\text{g}/\text{m}^3$, Adam มีค่า MAE = $8.79 \mu\text{g}/\text{m}^3$



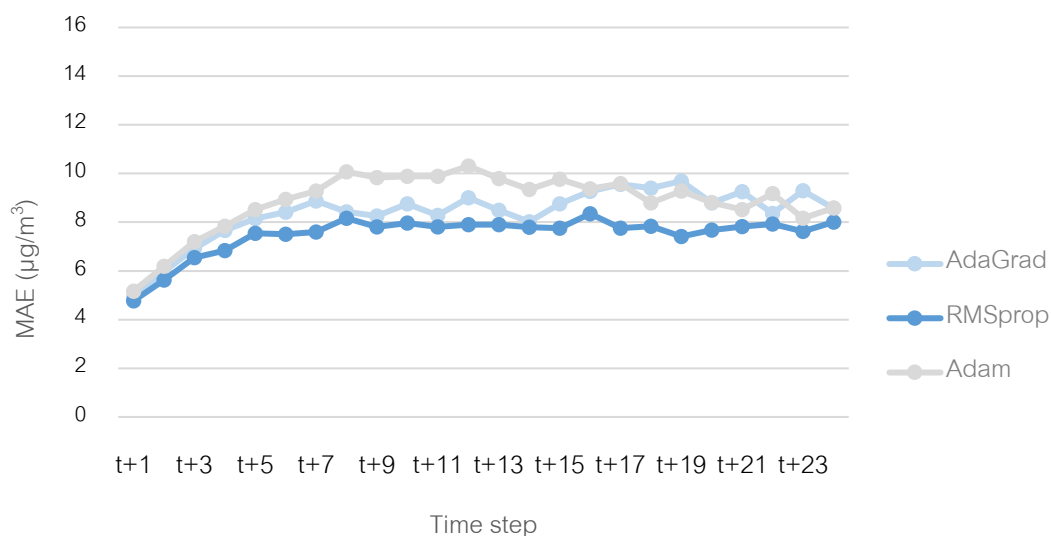
ภาพประกอบ 38 กราฟแสดงค่า MAPE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

จากภาพประกอบ 38 ในความคลาดเคลื่อน MAPE ของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ จะเห็นได้ว่าในช่วงเวลาระยะสั้น t+1 Optimizer ทั้งสามประเภทมีค่า MAPE ที่ต่ำที่สุดที่ใกล้เคียงกัน ในช่วงระยะเวลา t+8 ถึง t+24 จะเห็นว่า AdaGrad มีค่า MAPE น้อยที่สุดตลอดเวลาช่วง รองลงมาเป็น Adam ที่มีค่า MAPE น้อยที่สุด



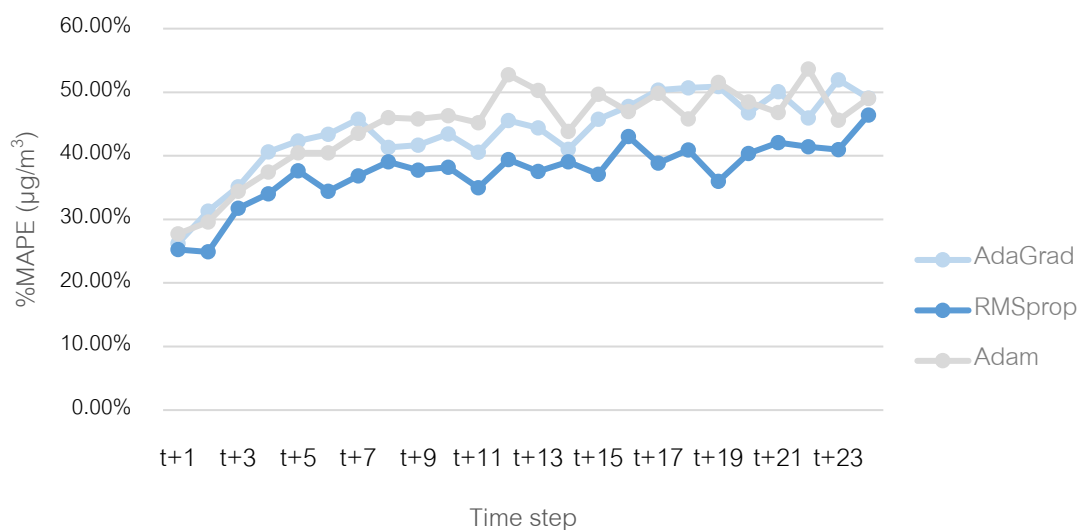
ภาพประกอบ 39 กราฟแสดงค่า RMSE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 39 ในความคลาดเคลื่อน RMSE ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย จะเห็นได้ว่าในช่วงเวลาระยะสั้น $t+1$ Optimizer ทั้งสามประเภทมีค่า RMSE ที่ต่ำที่สุดที่ใกล้เคียงกัน ในช่วงระยะเวลา $t+4$ ถึง $t+24$ จะเห็นว่า RMSprop มีค่า RMSE น้อยที่สุดตลอดช่วง เป็นที่ชัดเจนว่า Optimizer RMSprop เหมาะสมสำหรับชุดข้อมูลโรงเรียนยุพราชวิทยาลัย



ภาพประกอบ 40 กราฟแสดงค่า MAE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 40 ในความคลาดเคลื่อน MAE ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย จะเห็นได้ว่าในช่วงเวลาระยะสั้น $t+1$ Optimizer ทั้งสามประเภทมีค่า MAE ที่ต่ำที่สุดที่ใกล้เคียงกัน แต่ในช่วงระยะเวลา $t+3$ ถึง $t+24$ จะเห็นว่า RMSprop มีค่า MAE น้อยที่สุดตลอดช่วง เป็นที่ชัดเจนว่า Optimizer RMSprop เหมาะสมสำหรับชุดข้อมูลโรงเรียนยุพราชวิทยาลัย



ภาพประกอบ 41 กราฟแสดงค่า MAPE เปรียบเทียบ Optimizer ของ AdaGrad, RMSprop และ Adam ชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

จากภาพประกอบ 41 ในความคลาดเคลื่อน MAPE ของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย จะเห็นได้ว่าในช่วงเวลาระยะสั้น t+1 Optimizer ทั้งสามประเภทมีค่า MAPE ที่ต่ำที่สุดที่ใกล้เคียงกัน แต่ในช่วงระยะเวลา t+2 ถึง t+24 จะเห็นว่า RMSprop มีค่า MAPE น้อยที่สุดตลอดช่วง เป็นที่ชัดเจนว่า Optimizer RMSprop เหมาะสมสำหรับชุดข้อมูลโรงเรียนยุพราชวิทยาลัย

บทที่ 5

สรุป อภิปรายผล และข้อเสนอแนะ

ในการวิจัยการพยากรณ์ความเข้มข้นฝุ่นละออง PM2.5 ใน 24 ชั่วโมงล่วงหน้า ซึ่งใช้ข้อมูลความเข้มข้นฝุ่นละออง PM2.5 บริเวณพื้นที่สถานีศูนย์ราชการจังหวัดเชียงใหม่ และสถานีโรงเรียนยุพราชวิทยาลัย โดยใช้เทคนิคการเรียนรู้เชิงลึก แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ผู้วิจัยได้วัดความแม่นยำแบบจำลองของแต่ละสถานี และปรับปรุงประสิทธิภาพของแบบจำลองในการพยากรณ์ให้แม่นยำมากขึ้นด้วยการเปรียบเทียบการเลือก Optimizer ทั้ง 3 แบบ ได้แก่ AdaGrad, RMSprop และ Adam เพื่อให้แบบจำลองมีประสิทธิภาพสูงสุดควรเลือกใช้ Optimizer ที่เหมาะสมกับแบบจำลอง และชุดของข้อมูล

5.1 สรุปผลการวิจัย

แบบจำลอง LSTM มีประสิทธิภาพในการพยากรณ์ค่าความเข้มข้น PM2.5 ได้อย่างแม่นยำ โดยแตกต่างกันที่การเลือกใช้ Optimizer ในชุดข้อมูลสถานีศูนย์ราชการจังหวัดเชียงใหม่ เป็นชุดข้อมูลที่มีค่าข้อมูลสูญหายไม่มาก ควรเลือกใช้ Optimizer เป็น Adam หรือ AdaGrad จะทำให้สามารถพยากรณ์ล่วงหน้า 24 ชั่วโมงได้ดีที่สุด Adam มีค่า RMSE เท่ากับ $11.48 \mu\text{g}/\text{m}^3$, ค่า MAE เท่ากับ $8.78 \mu\text{g}/\text{m}^3$ และค่า MAPE เท่ากับ 21.65 เปอร์เซ็นต์ ส่วน AdaGrad มีค่า RMSE เท่ากับ $11.83 \mu\text{g}/\text{m}^3$, ค่า MAE เท่ากับ $9.05 \mu\text{g}/\text{m}^3$ และค่า MAPE เท่ากับ 20.42 เปอร์เซ็นต์ แต่ในส่วนชุดข้อมูลสถานีโรงเรียนยุพราชวิทยาลัยเป็นชุดข้อมูลที่มีค่าข้อมูลสูญหายมาก หรือร้อยละ 18.8 แบบจำลองควรเลือก Optimizer เป็น RMSprop จะทำให้สามารถพยากรณ์ล่วงหน้า 24 ชั่วโมงได้ดีที่สุด RMSprop มีค่า RMSE เท่ากับ $10.38 \mu\text{g}/\text{m}^3$, ค่า MAE เท่ากับ $8.01 \mu\text{g}/\text{m}^3$ และ MAPE เท่ากับ 46.40 เปอร์เซ็นต์

แบบจำลอง LSTM มีประสิทธิภาพการทำงานได้ดีในระยะยาว 24 ชั่วโมงล่วงหน้า มากกว่าในระยะกลาง เนื่องจากความคลาดเคลื่อนทั้งสามประเภท ได้แก่ RMSE, MAE และ MAPE มีค่าคลาดเคลื่อนค่อนข้างสูงบริเวณระยะกลาง ซึ่งมีความแตกต่างกับประสิทธิภาพการทำงานในระยะยาว ($t+24$) เนื่องจากในเวลา 24 ชั่วโมงเป็นช่วงเวลาที่มีความเข้มข้นฝุ่นละออง PM2.5 มีรูปแบบของรอบการขึ้นลงค่อนข้างชัดเจน ทำให้แบบจำลองสามารถพยากรณ์ได้ดีในระยะ 24 ชั่วโมงล่วงหน้า

ตาราง 5 เปรียบเทียบการเลือก Optimizer ตามช่วงเวลาของข้อมูลศูนย์ราชการจังหวัดเชียงใหม่

Time Step	RMSE	MAE	MAPE
t+1	Optimizer ทั้ง สาม ประเภท ได้แก่ AdaGrad, RMSprop และ Adam	Optimizer ทั้ง สาม ประเภท ได้แก่ AdaGrad, RMSprop และ Adam	Optimizer ทั้ง สาม ประเภท ได้แก่ AdaGrad, RMSprop และ Adam
t+3	RMSprop	RMSprop	RMSprop
t+6	RMSprop	RMSprop	RMSprop
t+9	RMSprop	RMSprop	Adam
t+12	RMSprop	RMSprop	Adam
t+15	Optimizer ทั้ง สาม ประเภท ได้แก่ AdaGrad, RMSprop และ Adam	Optimizer ทั้ง สาม ประเภท ได้แก่ AdaGrad, RMSprop และ Adam	Adam
t+18	AdaGrad และ Adam	AdaGrad และ Adam	Adam
t+21	AdaGrad และ Adam	AdaGrad และ Adam	AdaGrad และ Adam
t+24	AdaGrad และ Adam	AdaGrad และ Adam	AdaGrad และ Adam

จากตารางที่ 5 เปรียบเทียบการเลือกใช้ Optimizer ตามช่วงเวลาของชุดข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ พบว่าในช่วงระยะเวลาด้าน t+1 สามารถเลือกใช้ Optimizer ได้ทั้งสามประเภท ได้แก่ AdaGrad, RMSprop และ Adam โดยจะสามารถเลือกใช้ตัวใดก็ได้, ในช่วงระยะกลาง t+3, t+6, t+9, t+12 ควรเลือกใช้ Optimizer RMSprop เป็น Optimizer ที่เหมาะสมที่สุด รองลงมาจะเป็น Optimizer Adam, ในช่วงระยะเวลา t+18 ถึง t+21 ควรเลือกใช้เฉพาะ Optimizer AdaGrad หรือ Adam เท่านั้น และในช่วงระยะเวลา t+24 ควรเลือกใช้เฉพาะ Optimizer AdaGrad หรือ Adam เท่านั้น

ตาราง 6 เปรียบเทียบการเลือก Optimizer ตามช่วงเวลาของข้อมูลโรงเรียนยุพราชวิทยาลัย

Time Step	RMSE	MAE	MAPE
t+1	Optimizer ที่ ๓ ส าม ป ระ เก ท ไ ต้ แ ก่ AdaGrad, RMSprop และ Adam	Optimizer ที่ ๓ ส าม ป ระ เก ท ไ ต้ แ ก่ AdaGrad, RMSprop และ Adam	Optimizer ที่ ๓ ส าม ป ระ เก ท ไ ต้ แ ก่ AdaGrad, RMSprop และ Adam
t+3	RMSprop	RMSprop	RMSprop
t+6	RMSprop	RMSprop	RMSprop
t+9	RMSprop	RMSprop	RMSprop
t+12	RMSprop	RMSprop	RMSprop
t+15	RMSprop	RMSprop	RMSprop
t+18	RMSprop	RMSprop	RMSprop
t+21	RMSprop	RMSprop	RMSprop
t+24	RMSprop	RMSprop	RMSprop

จากตารางที่ 6 เปรียบเทียบการเลือกใช้ Optimizer ตามช่วงเวลาของชุดข้อมูลโรงเรียนยุพราชวิทยาลัย พบว่าในช่วงระยะเวลาล้น t+1 สามารถเลือกใช้ Optimizer ได้ทั้งสามประเภท ได้แก่ AdaGrad, RMSprop และ Adam โดยจะสามารถเลือกใช้ตัวใดก็ได้, ในช่วงระยะกลาง t+3, t+6, t+9, t+12, t+15, t+18, t+21 และ t+24 ควรเลือกใช้ Optimizer RMSprop เท่านั้น เพราะเป็น Optimizer ที่เหมาะสมที่สุด และมีความคลาดเคลื่อนน้อยที่สุด

5.2 อภิปรายผลการวิจัย

งานวิจัยได้แสดงให้เห็นว่าการสร้างแบบจำลองพยากรณ์ค่าความเข้มข้น PM2.5 นั้น นอกจากจะต้องพิจารณาปัจจัยของช่วงเวลา การเลือกอัลกอริทึมแล้ว ยังต้องพิจารณารายละเอียดของการปรับพารามิเตอร์ของอัลกอริทึมรวมทั้งค่าการประเมินผลด้วย ผู้วิจัยหวังว่าผลลัพธ์จากงานวิจัยนี้สามารถช่วยแนะแนวทางในการสร้างแบบจำลองการพยากรณ์ค่า PM2.5 สำหรับตำแหน่งอื่น ๆ ในประเทศต่อไป และการเลือก Optimizer จะสามารถทำให้แบบจำลองมีประสิทธิภาพความแม่นยำมากขึ้น การทำเปรียบเทียบจะสามารถทำให้ผู้วิจัยสามารถเลือกใช้

Optimizer ที่เหมาะสมในแต่ละแบบจำลอง และในแต่ละชุดข้อมูลอื่นจะทำให้ได้ผลลัพธ์ที่ดีขึ้นตามสถานการณ์

การเลือกช่วงเวลาให้แบบจำลองมีประสิทธิภาพดีที่สุด ควรจะเลือกพยากรณ์ล่วงหน้าในระยะสั้น และระยะยาว หรือ 24 ชั่วโมงล่วงหน้าจะให้ผลลัพธ์ที่ดีที่สุด เนื่องจากในเวลา 24 ชั่วโมงเป็นช่วงเวลาที่มีความเข้มข้นฝุ่นละออง PM2.5 ที่มีรูปแบบของรอบการขึ้นลงค่อนข้างชัดเจน ทำให้แบบจำลองสามารถพยากรณ์ได้ดีในระยะ 24 ชั่วโมงล่วงหน้า ช่วงระยะเวลาในช่วงที่ 24 หรือระยะยาวจึงให้ผลลัพธ์ที่ดีกว่าในระยะกลาง ดังนั้นจึงควรเลือกใช้แบบจำลองในช่วงระยะสั้น และระยะยาว สำหรับในช่วงระยะกลางจะสามารถพยากรณ์ล่วงหน้าได้ แต่มีประสิทธิภาพที่ลดลงเนื่องจากมีความคลาดเคลื่อนสูงกว่าระยะอื่น

5.3 ข้อเสนอแนะ

1. เนื่องจากในการวิจัยนี้ใช้ข้อมูลจาก 2 สถานีในการเปรียบเทียบผลลัพธ์ของการพยากรณ์ โดยพบว่า สถานีศูนย์ราชการจังหวัดเชียงใหม่ และสถานีโรงเรียนยุพราชวิทยาลัย จังหวัดเชียงใหม่ให้ผลลัพธ์ที่แตกต่างกัน เกิดจากค่า Missing value ของสถานีโรงเรียนยุพราชวิทยาลัยจังหวัดเชียงใหม่ ทำให้ค่าพยากรณ์ให้ผลที่ออกมาแตกต่างกัน ดังนั้น หากทดลองเลือกใช้ข้อมูลจังหวัดที่มีค่า Missing value ที่ไม่มากจะทำให้ผลการพยากรณ์ดียิ่งขึ้น

2. เนื่องจากในการวิจัยได้ใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาวในการพยากรณ์ฝุ่นละออง PM2.5 ล่วงหน้า 24 ชั่วโมง ซึ่งเป็นแบบจำลองในลักษณะอนุกรมเวลาจำเป็นต้องใช้ลักษณะของฤดูกาลเพื่อหารูปแบบของชุดข้อมูล แต่เนื่องจากข้อมูลชุดเป็นข้อมูลที่มีระยะเวลาประมาณ 1-1.5 ปี หากมีจำนวนข้อมูลที่มากขึ้นหลายปีอาจทำให้ส่งผลถึงประสิทธิภาพการพยากรณ์ที่ดีขึ้น

3. เนื่องจากในการวิจัยได้ใช้ข้อมูลศูนย์ราชการจังหวัดเชียงใหม่ และข้อมูลโรงเรียนยุพราชวิทยาลัย ที่มีตัวแปรบางตัวที่มีลักษณะที่แตกต่างกัน เช่น CO (คาร์บอนมอนอกไซด์), O3 (โอโซน) และ Rain (ค่าปริมาณน้ำฝน) อีกทั้งยังมีจำนวนที่ข้อมูลที่ไม่เท่ากันอาจส่งผลกระทบต่อพยากรณ์ หากเลือกตัวแปรที่เหมือนกันจะทำให้เปรียบเทียบได้ง่ายขึ้น

4. เนื่องจากในการวิจัยได้ใช้ข้อมูลสถานียุพราชวิทยาลัยที่มีค่า Missing value ของค่าปริมาณฝุ่นละออง PM2.5 จำนวนมาก โดยวิธีการเตรียมขั้นตอนใช้วิธีแทนค่าเฉลี่ยลงในค่านั้น ซึ่งทำให้ข้อมูลที่ได้นั้นไม่เป็นไปตามฤดูกาล ส่งผลถึงการพยากรณ์ และหากสามารถทำการ clean data โดยการแทนค่า Missing value เป็นค่าเฉลี่ยรายชั่วโมงจะทำให้ข้อมูลเป็นไปตามฤดูกาลมากขึ้น ทำให้การพยากรณ์มีผลลัพธ์ที่ดีขึ้นโดยพื้นฐาน

5. เนื่องจากในการวิจัยนี้ใช้ข้อมูลจากจังหวัดเชียงใหม่เพียงจังหวัดเดียว หากใช้ข้อมูลที่ได้จากสถานี่อื่นในประเทศไทย จะทำให้เปรียบเทียบประสิทธิภาพของการพยากรณ์ได้ดียิ่งขึ้น

6. เนื่องจากในการวิจัยได้ใช้ Optimizer 3 ประเภท ได้แก่ AdaGrad, RMSprop และ Adam ในการเปรียบเทียบประสิทธิภาพของแบบจำลองหากเลือกใช้ Optimizer ประเภทอื่นมาเปรียบเทียบกับอาจส่งผลต่อการพยากรณ์

7. เนื่องจากในงานวิจัยได้ใช้ Epoch, Neuron และ Batch size ที่เท่ากันในการเปรียบเทียบ Optimizer ทั้งสามประเภท หากทำการ Tuning Parameter ให้เหมาะสมกับแต่ละ Optimizer จะทำให้แบบจำลองมีประสิทธิภาพได้ดีขึ้น

8. เนื่องจากในงานวิจัยได้มีการประเมินวัดผลแบบจำลอง 3 ประเภท ได้แก่ RMSE, MAE และ MAPE หากเลือกใช้ตัววัดผลประเภทอื่นเข้ามาเพิ่มทำให้งานวิจัยสามารถวิเคราะห์ความคลาดเคลื่อนได้หลากหลายรูปแบบมากขึ้น

9. เนื่องจากในงานวิจัยได้มีการพูดถึงการทำงานของแบบจำลองที่มีค่าข้อมูลสูญหายเพื่อนำมาใช้สำหรับการพยากรณ์ความเข้มข้นฝุ่นละอองล่วงหน้า 24 ชั่วโมง ซึ่งในชีวิตประจำวันชุดข้อมูลที่ได้รับมามากจะได้รับมาในรูปแบบที่ไม่สมบูรณ์อยู่แล้ว หากทำการศึกษาลงรายละเอียดถึงเทคนิคในการจัดการข้อมูล การทำความสะอาดข้อมูลในรูปแบบต่าง ๆ สำหรับข้อมูลที่ผิดปกติจะทำให้สามารถช่วยแก้ปัญหาและปรับใช้กับงานวิจัยอื่น ๆ ได้ดี และเป็นประโยชน์ต่อการวิเคราะห์ข้อมูลเป็นอย่างมาก

บรรณานุกรม

- Du, S., Li, T., & Horng, S. (2018, 26-28 Dec. 2018). *Time Series Forecasting Using Sequence-to-Sequence Deep Learning Framework*. Paper presented at the 2018 9th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP).
- Du, S., Li, T., Yang, Y., & Horng, S. (2019). Deep Air Quality Forecasting Using Hybrid Deep Learning Framework. *IEEE Transactions on Knowledge and Data Engineering*, 1-1.
- Fister, D., & Jagric, T. (2019). Online Long Short-Term Memory Network for Stock Trading (pp. 5-8).
- Li, J., & Shen, J. (2019, 27-30 July 2019). *Prediction of PM2.5 Concentration Based on CEEMD-LSTM Model*. Paper presented at the 2019 Chinese Control Conference (CCC).
- Medium. (2016a). MAE and RMSE – Which Metric is Better? Retrieved from <https://medium.com/human-in-a-machine-world/mae-and-rmse-which-metric-is-better-e60ac3bde13d>
- Medium. (2016b). Understanding LSTM and its diagrams. Retrieved from <https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714>
- Montañez, J. A. R., Fernandez, M. A. A., Arriaga, S. T., Arreguin, J. M. R., & Calderon, G. A. S. (2019, 11-13 Sept. 2019). *Evaluation of a Recurrent Neural Network LSTM for the Detection of Exceedances of Particles PM10*. Paper presented at the 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE).
- Park, J., Yoo, S., Kim, K., Gu, Y., Lee, K., & Son, U. (2017, 4-7 July 2017). *PM10 density forecast model using long short term memory*. Paper presented at the 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN).
- Song, W., Han, J., Xie, J., Gao, Y., & Song, L. (2019, 14-17 July 2019). *System for Detecting and Forecasting PM2.5 Concentration Levels Using Long Short-Term*

Memory and LoRa. Paper presented at the 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData).

Standard, T. (2561). PM2.5 ฝุ่นละอองขนาดเล็กในอากาศ กับวิกฤตสุขภาพที่คนไทยจะต้องแลก.

Retrieved from <https://thestandard.co/pm-2-5-environmental-nano-pollutants/>

Tanty, R., & Desmukh, T. S. (2015). Application of Artificial Neural Network in Hydrology- A Review. *International journal of engineering research and technology*, 4.

Thaweephol, K., & Wiwatwattana, N. (2019, 20-22 Nov. 2019). *Long Short-Term Memory Deep Neural Network Model for PM2.5 Forecasting in the Bangkok Urban Area*. Paper presented at the 2019 17th International Conference on ICT and Knowledge Engineering (ICT&KE).

Tsai, Y., Zeng, Y., & Chang, Y. (2018, 12-15 Aug. 2018). *Air Pollution Forecasting Using RNN with LSTM*. Paper presented at the 2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech).

กรมควบคุมมลพิษ, ก. (2563). ข้อมูลดัชนีคุณภาพอากาศ. Retrieved from

http://air4thai.pcd.go.th/webV2/aqi_info.php

ประกอบพล, ธ. (2552). โครงข่ายประสาทเทียม Artificial Neural Networks. วารสาร มจร.วิชาการ, 24, 73-87.

ประชาชาติธุรกิจ. (2562). สถิติผู้เยี่ยมเยือนจังหวัดเชียงใหม่เดือนมกราคม-กรกฎาคมปี 2562.

Retrieved from <https://www.prachachat.net/tourism/news-374725>

สงวนสัจย์, ป. (2562). *Artificial Intelligence with Machine Learning, AI สร้างได้ด้วยแมชชีน เลิร์นนิ่ง*. นนทบุรี: ไอดีซี พรีเมียร์.

ประวัติผู้เขียน

ชื่อ-สกุล นายปรัชญา สิงหรวงศ์
วัน เดือน ปี เกิด 10 กรกฎาคม 2536
สถานที่เกิด นครราชสีมา
วุฒิการศึกษา พ.ศ. 2554
วิทยาศาสตรบัณฑิต สาขาวิชาสถิติ
จาก มหาวิทยาลัยศรีนครินทรวิโรฒ
พ.ศ. 2561
วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
จาก มหาวิทยาลัยศรีนครินทรวิโรฒ

