ผลกระทบของการทำการสกัดคุณลักษณะของข้อมูลและการแทนที่ข้อมูล กับแบบจำลองพยากรณ์ฝุ่นละออง PM2.5 สำหรับบริเวณกรุงเทพมหานคร

THE IMPACT OF FEATURE EXTRACTION AND DATA IMPUTATION ON PM2.5 FORECASTING MODEL FOR BANGKOK AREA

ภานุพงษ์ ร่องอ้อ

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2562

ผลกระทบของการทำการสกัดคุณลักษณะของข้อมูลและการแทนที่ข้อมูล กับแบบจำลองพยากรณ์ฝุ่นละออง PM2.5 สำหรับบริเวณกรุงเทพมหานคร

ภานุพงษ์ ร่องอ้อ

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2562

THE IMPACT OF FEATURE EXTRACTION AND DATA IMPUTATION ON PM2.5 FORECASTING MODEL FOR BANGKOK AREA

PANUPONG RONG-O

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of MASTER OF SCIENCE

(Information Technology)

Faculty of Science, Srinakharinwirot University

2019

ปริญญานิพนธ์

เรื่อง

ผลกระทบของการทำการสกัดคุณลักษณะของข้อมูลและการแทนที่ข้อมูล

กับแบบจำลองพยากรณ์ฝุ่นละออง PM2.5

สำหรับบริเวณกรุงเทพมหานคร

ของ

ภานุพงษ์ ร่องอ้อ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

.....

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

คณบดีบัณฑิตวิทยาลัย

.....

คณะกรรมการสอบปากเปล่าปริญญานิพนธ์

..... ที่ปรึกษาหลัก

(ผู้ช่วยศาสตราจารย์ ดร.นุวีย์ วิวัฒนวัฒนา)

..... ประธาน

(ดร.สุทธิพงศ์ ธัชยพงษ์)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.จันตรี ผลประเสริฐ)

ชื่อเรื่อง	ผลกระทบของการทำการสกัดคุณลักษณะของข้อมูลและการแทนที่ข้อมูล กับแบบจำลองพยากรณ์ฝุ่นละออง PM2.5 สำหรับบริเวณกรุงเทพมหานคร
ผู้วิจัย	ภานุพงษ์ ร่องอ้อ
ปริญญา	วิทยาศาสตรมหาบัณฑิต
ปีการศึกษา	2562
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. นุวีย์ วิวัฒนวัฒนา

ไม่กี่ปีที่ผ่านมาเกิดเหตุการณ์มลภาวะทางอากาศฝุ่นละอองขนาดเล็ก PM2.5 เพิ่มขึ้นอย่างรวดเร็วในเมืองใหญ่ของประเทศไทย มีหลายงานวิจัยพยายามพัฒนาเทคนิคความจำระยะสั้นแบบยาว (Long Short-Term Memory : LSTM) ที่เป็นแบบจำลองแบบโครงข่ายประสาทเชิงลึก (Deep Neural Network) สำหรับพยากรณ์ค่าความหนาแน่นของ PM2.5 อย่างไรก็ตามยังมีการศึกษาผลกระทบของการทำการแทนที่ข้อมูลสูญหายและการสกัดคุณลักษณะที่มีผลกับประสิทธิภาพของ LSTM ในการพยากรณ์ค่า PM2.5 อยู่ค่อนข้างน้อย ในงานวิจัยนี้ผู้วิจัยแทนที่ข้อมูลสูญหายด้วย Kalman Smoothing และ Linearly Weighted Moving Average และใช้ LSTM Autoencoder (LSTM AE) สำหรับการสกัดคุณลักษณะ ชุดข้อมูลที่ใช้พยากรณ์ PM2.5 ได้มาจากการตรวจวัดบริเวณสถานีตำรวจนครบาลโชคชัย โดยงานวิจัยนี้ทำการพยากรณ์ค่า PM2.5 ล่วงหน้า 24 ชั่วโมง ผู้วิจัยพิสูจน์ให้เห็นว่าประสิทธิภาพของแบบจำลอง LSTM เพิ่มขึ้นโดยรวมประมาณ 7 เปอร์เซ็นต์เมื่อใช้ Root Mean Square Error (RMSE) ประเมิน และประสิทธิภาพเพิ่มขึ้นมากกว่า 10 เปอร์เซ็นต์เมื่อใช้ Mean Abosolute Error (MAE) ประเมิน ส่วน LSTM AE เพิ่มประสิทธิภาพในการพยากรณ์ในแต่ละช่วงเวลาแตกต่างกันไปตามจำนวนชั่วโมงโดยเฉพาะการพยากรณ์ล่วงหน้า 22 ถึง 24 ชั่วโมงมีแนวโน้มจะได้ประสิทธิภาพที่ดี

คำสำคัญ : การแทนที่ข้อมูลสูญหาย, ความจำระยะสั้นแบบยาว, ตัวเข้ารหัสแบบความจำระยะสั้นแบบยาว, การพยากรณ์ PM2.5, การสกัดคุณลักษณะ

Title	THE IMPACT OF FEATURE EXTRACTION AND DATA IMPUTATION ON PM2.5 FORECASTING MODEL FOR BANGKOK AREA
Author	PANUPONG RONG-O
Degree	MASTER OF SCIENCE
Academic Year	2019
Thesis Advisor	Assistant Professor Dr. Nuwee Wiwatwattana

The last few years have seen a dramatic increase in PM2.5 air pollution in Thailand's major cities. Various works have tried to develop efficient Long Short-Term Memory (LSTM) deep neural network models for PM2.5 concentration forecasting. However, little has been studied about the impact of data imputation and feature extraction on the model performance in this context. In this reserch, we imputed missing values using Kalman Smoothing and Linearly Weighted Moving Average. We utilized the LSTM Autoencoder (LSTM AE) for feature extraction. Using the Chokchai Police station in Bangkok as a case study to predict PM2.5 in the next 24 hours, we demonstrated that the performance gain from training LSTM models with imputed data is more than 7 percent overall with respect to the root mean square error (RMSE) and more than 10 percent overall with respect to the mean absolute error (MAE). Improvement with LSTM AE varies according to time steps. Forecasting 22 to 24 hours ahead tends to favor the use of LSTM AE.

Keyword : Data imputation, LSTM, LSTM autoencoder, PM2.5 forecasting, Feature extraction

กิตติกรรมประกาศ

ปริญญานิพนธ์นี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์จาก ผศ.ดร.นุวีย์ วิวัฒนวัฒนา อาจารย์ที่ปรึกษา ที่ให้คำปรึกษา คำแนะนำในการทำปริญญานิพนธ์ ตลอดจนสนับสนุนข้อมูลทางวิชาการ และข้อมูลสำหรับทำปริญญานิพนธ์นี้

กราบขอบพระคุณกรมควบคุมมลพิษ ที่ให้ความอนุเคราะห์ข้อมูลฝุ่นละออง PM2.5 และข้อมูลอุตุนิยมวิทยาที่ถูกจัดเก็บในพื้นที่กรุงเทพมหานครชั้นในมาใช้ในการทำปริญญานิพนธ์นี้

ขอกราบขอบพระคุณคณะกรรมการสอบปริญญานิพนธ์ ที่ได้ให้คำแนะนำและข้อเสนอแนะสำหรับการปรับปรุงปริญญานิพนธ์ และการใช้เทคนิคการแทนที่ข้อมูล (Data Imputation) และการสกัดคุณลักษณะ (Feature Extraction) ร่วมกับการเรียนรู้ของเครื่อง โดยใช้แบบจำลองโครงข่ายประสาทเทียมเชิงลึก เพื่อเพิ่มประสิทธิภาพของการพยากรณ์ PM2.5 ต่อไป

ขอกราบขอบพระคุณบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ สำหรับทุนสนับสนุนการนำเสนอผลงานวิจัยของนิสิตบัณฑิตศึกษา ทำให้ได้รับประสบการณ์นำเสนอปริญญานิพนธ์ รวมถึงได้แลกเปลี่ยนความรู้ทางด้านเทคโนโลยีต่าง ๆ

สุดท้ายนี้ ขอกล่าวขอบพระคุณ น.ส.วารี ทัพน้อย มีศักดิ์เป็นป้าของผู้วิจัย ผู้ซึ่งเปิดโอกาสให้ได้รับการศึกษาเล่าเรียน ตลอดจนคอยช่วยเหลือและให้กำลังใจผู้วิจัยเสมอมาจนสำเร็จการศึกษา

ภานุพงษ์ ร่องอ้อ

สารบัญ

หน้า

สารบัญตาราง

หน้า

สารบัญรูปภาพ

หน้า

บทที่1
บทนำ

1.1 ที่มาและความสำคัญ

ที่ผ่านมาประเทศไทยได้ประสบปัญหาหมอกควันปกคลุมเมืองใหญ่หลายเมืองมาอย่างต่อเนื่อง ปัจจัยที่ทำให้เกิดเหตุการณ์หมอกควันเกิดขึ้นจากหลายปัจจัยตามที่ World Health Organization ได้ให้ข้อมูลไว้ได้แก่ เชื้อเพลิงที่ใช้ในยานพาหนะ ระบบสร้างความร้อนภายในบ้าน การเผาในภาคการเกษตรกรรม การเผาในภาคอุตสาหกรรม แม้กระทั่งการการหุงหาอาหารที่ทำให้เกิดควันจากการเผาไหม้ สิ่งที่ทำให้หมอกควันนั้นอันตรายเพราะในหมอกควันประกอบด้วยฝุ่นละอองขนาดเล็กกว่า 2.5 ไมครอน หรืออีกชื่อหนึ่งคือ PM2.5 ด้วยขนาดที่เล็กกว่าขนจมูกมนุษย์ถึง 10 เท่าส่งผลให้กลไกการหายใจของมนุษย์ไม่สามารถกรอง PM2.5 ได้ ก่อเกิดผลเสียต่อคุณภาพชีวิตของประชาชนทุกคน

ที่ผ่านมาการพยากรณ์ค่าความหนาแน่น PM2.5 ใช้พื้นฐานจากข้อมูลมลพิษและข้อมูลทางอุตุนิยมวิทยา เช่น อุณหภูมิ ความกดอากาศ หรือค่าความชื้นสัมพัทธ์ แต่ในปัจจุบันความพยายามในการพยากรณ์ PM2.5 มุ่งเน้นไปในทิศทางการสร้างเครื่องมือพยากรณ์โดยใช้การเรียนรู้ของเครื่องจักร (Machine Learning) ด้วยการเรียนรู้เครื่องจักรแบบการเรียนรู้เชิงลึก (Deep Learning) และ ข่ายงานแบบเวียนเกิด (Recurrent Neural Network) สามารถสร้างผลลัพธ์การพยากรณ์ที่มีประสิทธิภาพ หน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory : LSTM) เป็นแบบจำลอง RNN ที่ได้รับการพิสูจน์ในวงกว้างแล้วว่ามีประสิทธิภาพที่ดีกว่าวิธีการทางสถิติ Autoregressive Integrated Moving Average (ARIMA) อย่างไรก็ตามพบว่าการศึกษาวิธีการจัดการข้อมูลสูญหายเพื่อลดข้อมูลรบกวน (Noise) ในบริบทของการพยากรณ์มลพิษทางอากาศมีจำนวนที่ไม่มาก และค่าPM2.5 ที่อยู่ในชุดข้อมูลนั้นมีผลต่อประสิทธิภาพแบบจำลอง LSTM โดยตรง

ดังนั้นผู้วิจัยจึงเห็นความสำคัญและทำการศึกษาการแทนที่ข้อมูล (Data Imputation) เพื่อลดการเกิด Noise ของชุดข้อมูล ร่วมกับการสกัดคุณลักษณะ (Feature Extraction) ข้อมูล PM2.5 เพื่อเป็นการขจัด Noise ของ PM2.5 ในการวิจัยนี้ ได้ทดลองกับชุดข้อมูล PM2.5 ที่ถูกตรวจวัดบริเวณสถานีตำรวจโชคชัย ซึ่งชุดข้อมูลมีค่าสูญหาย (Missing Value) จำนวนหนึ่ง โดยการวัดประสิทธิภาพของแบบจำลองจากการพยากรณ์ ค่าฝุ่นละออง PM2.5 ล่วงหน้าเป็นระยะเวลา 24 ชั่วโมง ซึ่งผู้วิจัยคาดหวังว่าแบบจำลอง LSTM ที่สร้างจากข้อมูลที่แทนที่ข้อมูลสูญ

หายร่วมกับการสกัดคุณลักษณะจะสามารถพยากรณ์ค่าความหนาแน่นของฝุ่นละออง PM25 ได้อย่างแม่นยำมากขึ้น นำไปสู่การแจ้งเตือนและการเตรียมความพร้อมรับมือเหตุการณ์หมอกควันที่จะเกิดขึ้นกับประชาชนต่อไป

1.2 วัตถุประสงค์

1. ศึกษาและประยุกต์ใช้แบบจำลองเครือข่ายประสาทเทียมประเภทหน่วยความจำระยะสั้นแบบยาว (LSTM) กับการพยากรณ์ฝุ่นละออง PM2.5
2. ศึกษาวิธีการแทนที่ข้อมูล (Data Imputation) ด้วยวิธี Kalman Smoothing และ Linearly Weighted Moving Average
3. ศึกษาวิธีการสกัดคุณลักษณะ (Feature Extraction) ด้วย LSTM Autoencoder
4. ศึกษาวิธีประยุกต์ใช้การแทนที่ข้อมูลและการสกัดคุณลักษณะกับแบบจำลองเครือข่ายประสาทเทียมประเภทหน่วยความจำระยะสั้นแบบยาว (LSTM) และเปรียบเทียบประสิทธิภาพของแบบจำลองสำหรับพยากรณ์ฝุ่นละออง PM2.5

1.3 ขอบเขตของการวิจัย

งานวิจัยนี้พัฒนาแบบจำลองการพยากรณ์ PM2.5 โดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ที่ทำงานในรูปแบบของโปรแกรมคอมพิวเตอร์ โดยใช้ข้อมูลที่ถูกเก็บมาจากสถานีวัดคุณภาพอากาศบริเวณหน้าสถานีตำรวจนครบาลโชคชัย ริมถนนลาดพร้าว ซึ่งอยู่ในพื้นที่ของกรุงเทพมหานคร ข้อมูลที่ใช้ในการพยากรณ์ประกอบด้วยข้อมูลฝุ่นละออง PM2.5 ฝุ่นละออง PM10 ปี (year) เดือน (month) วันที่ (day) ชั่วโมง (hour) คาร์บอนมอนอกไซด์ (CO) ไนโตรเจนมอนอกไซด์ (NO) ออกไซด์ของไนโตรเจน (NOX) ไนโตรเจนไดออกไซด์ (NO2) ซัลเฟอร์ไดออกไซด์ (SO2) โอโซน (O3) ความเร็วลม (Wind Speed) ทิศทางลม (Wind Direction) อุณหภูมิ (Temp) ความชื้นสัมพัทธ์ (Rel hum) ความกดอากาศ (Pressure) และปริมาณน้ำฝน (Rain)

เทคนิคการเรียนรู้ของเครื่องที่จะประยุกต์ใช้ในงานวิจัยนี้ประกอบด้วยแบบจำลองเครือข่ายประสาทเทียมประเภทหน่วยความจำระยะสั้นแบบยาว (LSTM) ร่วมกับเทคนิค การแทนที่ข้อมูล (Data Imputation) ด้วยวิธี Kalman Smoothing และ Linearly Weighted Moving Average และ การสกัดคุณลักษณะ (Feature Extraction) ด้วย LSTM Autoencoder โดยศึกษาว่าเทคนิคเหล่านี้ส่งผลให้ LSTM มีประสิทธิภาพในการทำงานดีขึ้นหรือไม่ โดยเครื่องมือที่ใช้วัดประสิทธิภาพของแบบจำลองได้แก่ Root Mean Square Error (RMSE) และ Mean Absolute Error (MAE)

1.4 วิธีดำเนินการวิจัย

1. ศึกษาหลักการทฤษฎีและทบทวนวรรณกรรม
2. ทำการแทนที่ข้อมูล (Data Imputation) ที่สูญหายด้วยวิธี Kalman Smoothing และ Linearly Weighted Moving Average
3. ทำการสกัดคุณลักษณะ (Feature Extraction) ด้วย LSTM Autoencoder กับชุดข้อมูลที่ผ่านการทำการแทนที่ข้อมูล
4. ทดสอบประสิทธิภาพแบบจำลอง LSTM ที่สร้างจากข้อมูลที่ผ่านการแทนที่ข้อมูล (Data Imputation) และการสกัดคุณลักษณะ (Feature Extraction)
5. สรุปผลการทดลอง

1.5 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

1. เพื่อศึกษาเทคนิคที่เหมาะสมในการพยากรณ์ฝุ่นละออง PM2.5
2. ได้แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ที่มีประสิทธิภาพ
3. เพื่อเพิ่มโอกาสการนำแบบจำลองไปใช้ในสถาณการณ์นำไปสู่การแจ้งเตือนประชาชนอย่างทันท่วงที

บทที่ 2
แนวคิด ทฤษฎี เทคโนโลยี และระบบงานที่เกี่ยวข้อง

2.1 การแทนที่ข้อมูล (Data Imputation)

การรวบรวมข้อมูลเป็นขั้นสำคัญสำหรับกระบวนการเรียนรู้ของเครื่องจักร (Machine Learning) แต่ชุดข้อมูลส่วนใหญ่มักจะมีข้อมูลสูญหาย (Missing Value) รวมอยู่ด้วยไม่มากก็น้อย การสร้างแบบจำลองการเรียนรู้ของเครื่องจักรด้วยชุดข้อมูลสูญหายนั้นทำให้เกิดผลกระทบต่อประสิทธิภาพของแบบจำลองเนื่องจากรูปแบบ (Pattern) ของข้อมูลที่ได้สูญหายไป การแทนที่ข้อมูลด้วยวิธีการที่เหมาะสมเพื่อให้ชุดข้อมูลยังคงรูปแบบนำไปสู่การสร้างแบบจำลองการเรียนรู้ของเครื่องได้อย่างมีประสิทธิภาพ

การได้มาของชุดข้อมูลฝุ่น PM2.5 ในการวิจัยนี้นั้นมาจากเครื่องจักรจากสถานีตรวจวัดบริเวณสถานีตำรวจนครบาลโชคชัย เนื่องจากในบางเวลาเครื่องจักรอาจมีการชำรุดทำให้ชุดข้อมูลที่ได้เกิดการสูญหายของข้อมูล ผู้วิจัยได้ทำการศึกษาวิธีการแทนที่ข้อมูลฝุ่น PM2.5 มีงานวิจัยได้ทำการลบข้อมูลทั้งแถวเมื่อเกิดข้อมูลสูญหายเกิดขึ้น [1] บางงานวิจัยได้เลือกวิธีการแทนค่าคงที่ให้กับข้อมูลที่สูญหาย [2] และมีงานวิจัยได้ใช้วิธีการที่ชื่อว่า Moving Average เป็นการคำนวณข้อมูลใกล้เคียงเพื่อแทนที่ข้อมูลสูญหาย [3]

จากการสังเกตของผู้วิจัยพบว่าการใช้ค่าคงที่บางค่า เช่น แทนที่ด้วย 0 หรือค่าเฉลี่ยของข้อมูลนั้นอาจเป็นการสร้างข้อมูลรบกวน (Noise) ให้กับชุดข้อมูลถ้าหากชุดข้อมูลนั้นเกิดเหตุการณ์เปลี่ยนแปลงข้อมูลบ่อย อย่างเช่นข้อมูลฝุ่นละออง PM2.5 ดังนั้นผู้วิจัยจึงได้ทำการศึกษาและพบว่ามีวิธีการแทนที่ข้อมูลแบบอนุกรมเวลา (Time Series) ที่ใช้ข้อมูลรอบข้างข้อมูลที่สูญหาย เช่น ข้อมูลในอดีต หรือ ข้อมูลที่เกิดขึ้นแล้วในอนาคต มาใช้ในการแทนที่ข้อมูลที่สูญหาย มีวิธีการแทนที่ข้อมูลที่ชื่อว่า Moving Average เคยถูกใช้ร่วมกับการพยากรณ์ฝุ่นละออง [3] และ วิธีการชื่อ Kalman Smoothing สามารถพยากรณ์แนวโน้มของข้อมูลอนุกรมเวลาได้อย่างมีประสิทธิภาพ [4]

ในงานวิจัยนี้ผู้วิจัยได้เลือกใช้วิธีการแทนที่ข้อมูลโดยคำนวณจากข้อมูลข้างเคียง เพื่อคงไว้ซึ่งรูปแบบของข้อมูล โดยวิธีที่ได้เลือกใช้ในงานวิจัยนี้ได้แก่

- Kalman Smoothing (KS)
- Linearly Weighted Moving Average (LWMA)

2.1.1 Kalman Smoothing (KS)

Kalman Smoothing (KS) [5] ถูกนำเสนอในปี 1960 เทคนิคนี้ได้รับความสนใจเป็นอย่างมากและถูกประยุกต์ใช้ในงานหลายด้านเช่น ด้านการนำทาง ด้านการประมวลข้อมูลจากเซ็นเซอร์ภายใต้สัญญาณรบกวน (Noise) หรือด้านการแทนที่ข้อมูลที่สูญหาย ผลลัพธ์การประมวลผลที่มีประสิทธิภาพของ KS นั้นสืบเนื่องมาจากวิธีการทำงานในรูปแบบของการป้อนกลับของสัญญาณ (Feedback Control) รวมกับขั้นตอนการคำนวณแบบการเรียกซ้ำ (Recursive) โดยรวม KS ทำนายผลลัพธ์ข้อมูลสูญหายขณะเวลาใดเวลาหนึ่งหลังจากนั้นจึงนำค่าที่คำนวณได้ใหม่ป้อนคืนกลับในรูปแบบการวัดค่าสัญญาณรบกวน

สำหรับการคำนวณข้อมูลสูญหายด้วย KS แบ่งออกเป็น 2 ขั้นตอน

- ขั้นตอนการทำนาย (Prediction)
- ขั้นตอนการปรับแก้ (Correction)

2.1.1.1 ขั้นตอนการทำนาย (Prediction)

ขั้นตอนการทำนายเป็นการคำนวณ $\hat{x}_k^-$ คือค่าของข้อมูลสูญหายขณะเวลา k และค่า P_k^- คือค่า error covariance ที่รอการปรับแก้ด้วยค่าอัตราขยายคาลมาน (Kalman Gain) สามารถแสดงได้ดังสมการที่ (1) และ (2) ตามลำดับ

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \tag{1}$$

โดยที่

$\hat{x}_k^-$ คือค่าข้อมูลสูญหายที่ทำนายก่อนการปรับแก้ด้วย Kalman Gain ณ เวลา k

$\hat{x}_{k-1}$ คือค่าข้อมูลสูญหายที่ทำนาย ณ เวลา $k-1$

u_{k-1} คือเวกเตอร์ควบคุม ณ เวลา $k-1$

$$P_k^- = AP_{k-1}A^T + Q \tag{2}$$

โดยที่

P_k^- คือค่า error covariance ก่อนการปรับแก้ด้วย Kalman Gain ณ เวลา k

P_{k-1} คือค่า error covariance ณ เวลา $k-1$

Q คือค่า covariance ของ สัญญาณรบกวนจากระบบ (Process noise)

เมื่อได้ $\hat{x}_k^-$ และ P_k^- จากการคำนวณแล้วนำไปสู่ขั้นตอนต่อไป

2.1.1.2 ขั้นตอนการปรับแก้ (Correction)

หลังจากได้ $\hat{x}_k^-$ และ P_k^- มาจากขั้นตอนที่ 2.1.1.1 KS ได้ทำการคำนวณค่าอัตราขยายคาลมาน (Kalman gain) ดังสมการที่ 3

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (3)$$

โดยที่

K_k คือค่าอัตราขยายคาลมาน ณ เวลา k

P_k^- คือค่า error covariance ก่อนการปรับแก้ ณ เวลา k

R คือ covariance ของ สัญญาณรบกวนจากการวัด (Measurement noise)

เมื่อคำนวณค่าอัตราขยายคาลมานแล้วขั้นต่อไปคือการปรับแก้ค่าข้อมูลสูญหายและค่า error covariance ที่เวลาขณะที่ k โดยใช้ค่าอัตราขยายคาลมาน ดังสมการที่ (4) และ (5)

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (4)$$

โดยที่

$\hat{x}_k$ ค่าข้อมูลสูญหาย ณ เวลา k

$\hat{x}_k^-$ คือค่าข้อมูลสูญหายก่อนการปรับแก้ด้วย Kalman Gain ณ เวลา k

K_k คือค่าอัตราขยายคาลมาน ณ เวลา k

z_k คือค่าจริงที่ได้จากการวัด (ถ้าค่าไม่สูญหาย) ณ เวลา k

$$P_k = (I - K_k H) P_k^- \quad (5)$$

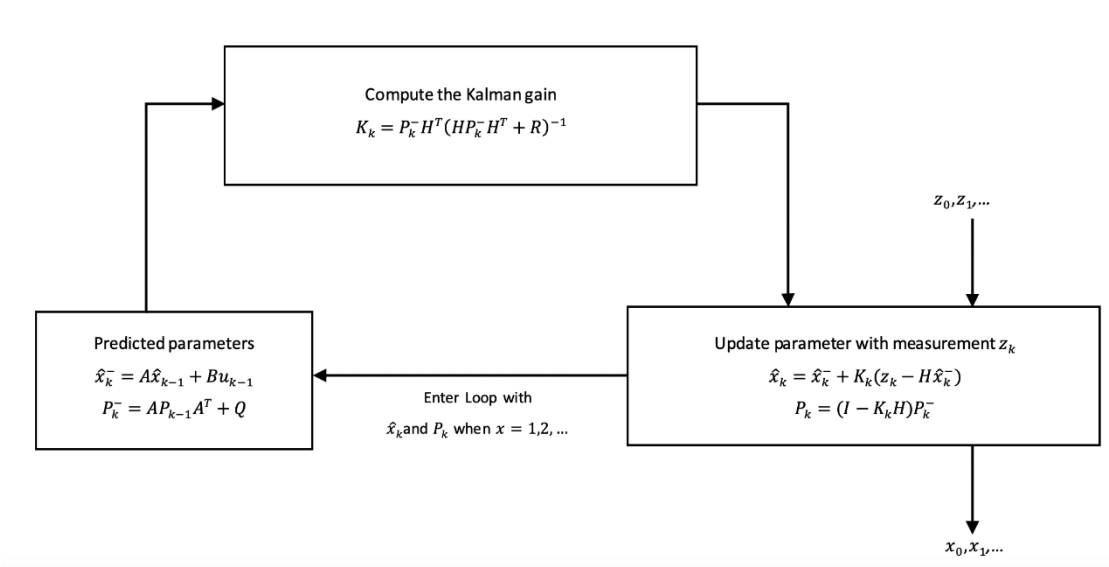
โดยที่

P_k คือค่า error covariance ที่ได้การปรับแก้ขณะเวลา k

K_k คือค่าอัตราขยายคาลมานขณะเวลา k

P_k^- คือค่า error covariance ก่อนการปรับแก้ขณะเวลา k

หลังจากการทำขั้นตอนการปรับแก้จบ กระบวนการทั้งหมดจะถูกทำซ้ำโดยข้อมูลที่ได้ถูกปรับแก้แล้วจะกลายเป็นสภาวะระบบอ้างอิงในรูปถัดไปสามารถแสดงการทำงานดังภาพประกอบ 1



ภาพประกอบ 1 แผนผังการแทนที่ข้อมูลสูญหายด้วย Kalman Smoothing

2.1.2 Linearly Weighted Moving Average (LWMA)

Moving average (MA) ถูกใช้กับข้อมูลประเภทอนุกรมเวลาเพื่อหาแนวโน้มการผันผวนของชุดข้อมูล MA ถูกพบว่าใช้กับงานที่เกี่ยวกับการเงินเป็นจำนวนมาก แต่มีงานวิจัยเกี่ยวกับการพยากรณ์ค่ามลพิษ [3] ได้นำ MA เข้ามาช่วยในการทำการแทนที่ข้อมูลสูญหาย ผู้วิจัยจึงได้ศึกษา MAเพิ่มเติมและได้พบว่า LWMA เป็นทางเลือกที่ดีในการทำการแทนที่ข้อมูลสูญหายในงานวิจัยนี้ เพราะ LWMA ได้คำนวณค่าเฉลี่ยพร้อมทั้งกำหนดค่าน้ำหนัก (Weighted) ตามระยะความใกล้จากจุดที่ต้องการแทนที่ข้อมูลสูญหาย โดยใน LWMA ในงานวิจัยชิ้นนี้จะทำการสำรวจค่าไปข้างหน้าและข้างหลังเพื่อคำนวณค่าเพื่อนำไปแทนที่ข้อมูลสูญหาย

$$x_n = \frac{\left(x_{n-\frac{k}{2}} * W_1\right) + \cdots + \left(x_{n+\frac{k}{2}} * W_k\right)}{\sum W} \quad (6)$$

โดยที่

x_n คือข้อมูลสูญหายขณะเวลา n

n คือตำแหน่งของเวลาที่ข้อมูลสูญหาย

k คือจำนวนช่วงเวลาที่ทำการคำนวณค่า LWMA

W คือน้ำหนักของข้อมูลตามตำแหน่งเวลา

2.2 การสกัดคุณลักษณะ (Feature Extraction)

การสกัดคุณลักษณะเป็นขั้นตอนทั่วไปของการเรียนรู้เครื่องจักรโดยลักษณะการทำงานของการสกัดลักษณะนั้นเป็นการค้นหาคุณลักษณะเฉพาะ (Characteristic) ของชุดข้อมูลขนาดใหญ่ โดยส่วนใหญ่การสกัดคุณลักษณะมักทำการเลือกตัวแปรหรือผสมตัวแปรตามแต่ละอัลกอริทึมที่ถูกออกแบบมาเพื่อได้ชุดข้อมูลที่มีขนาดที่ลดลงและยังคงคุณลักษณะเฉพาะไว้ทำให้สามารถสร้างแบบพยากรณ์ที่มีความแม่นยำแต่ใช้เวลาในการสร้างแบบจำลองที่ลดลง จากการศึกษาของผู้วิจัยพบว่ามีการใช้การสกัดคุณลักษณะเพื่อเพิ่มประสิทธิภาพแบบจำลอง LSTM เช่น งานวิจัย [6, 7] ได้ใช้เทคนิค Principle Component Analysist (PCA) ทำการสกัดคุณลักษณะ หรืองานวิจัย [8, 9] ได้ทำการสกัดคุณลักษณะเพื่อลด Noise ในชุดข้อมูลด้วยเทคนิค Stacked Autoencoder และผู้วิจัยได้พบการวิจัยที่น่าสนใจ [10, 11] ที่ใช้เทคนิค LSTM Autoencoder เป็นเครื่องมือสำหรับการสกัดคุณลักษณะซึ่งส่งผลให้แบบจำลอง LSTM มีประสิทธิภาพเพิ่มขึ้น

ขั้นตอนการสร้าง LSTM Autoencoder แบ่งออกเป็น 2 ขั้นตอนหลัก ได้แก่ ขั้นตอนการเข้ารหัส และ ขั้นตอนการถอดรหัส คุณลักษณะใหม่จะถูกสกัดจากข้อมูลนำเข้าที่ผ่านไปยังชั้นการเข้ารหัส สามารถแสดงได้ดังสมการ (7)

$$f_x = enc(x_{tp}) \quad (7)$$

โดยที่

f_x คือคุณลักษณะที่ถูกสกัดออกมาจากข้อมูลนำเข้า

enc คือชั้นการเข้ารหัสข้อมูลด้วย LSTM

x_{tp} คือข้อมูลนำเข้า

และหลังจากนั้นคุณลักษณะใหม่จะถูกแปลงโครงสร้างผ่านชั้นการถอดรหัสกลับเป็นข้อมูลนำเข้า ดังสมการ (8)

$$\hat{x}_{tp} = dec(f_x) \tag{8}$$

โดยที่

$\hat{x}_{tp}$ คือข้อมูลนำเข้าที่ถูกแปลงจากชั้นถอดรหัสข้อมูลด้วย LSTM

enc คือชั้นการถอดรหัสข้อมูลด้วย LSTM

f_x คือคุณลักษณะที่ถูกสกัดออกมา

เนื่องจาก LSTM AE เป็นแบบจำลองแบบการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เพื่อที่จะสร้างคุณลักษณะใหม่ที่ยังคงรูปแบบของข้อมูลดั้งเดิมจำเป็นต้องสร้างแบบจำลอง LSTM AE ที่ให้ผลลัพธ์ $\hat{x}_{tp}$ ใกล้เคียงกับข้อมูลนำเข้า x_{tp} โดยใช้ Mean Square Error (MSE) เป็น loss function

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(x_{tp} - \hat{x}_{tp})^2 \tag{9}$$

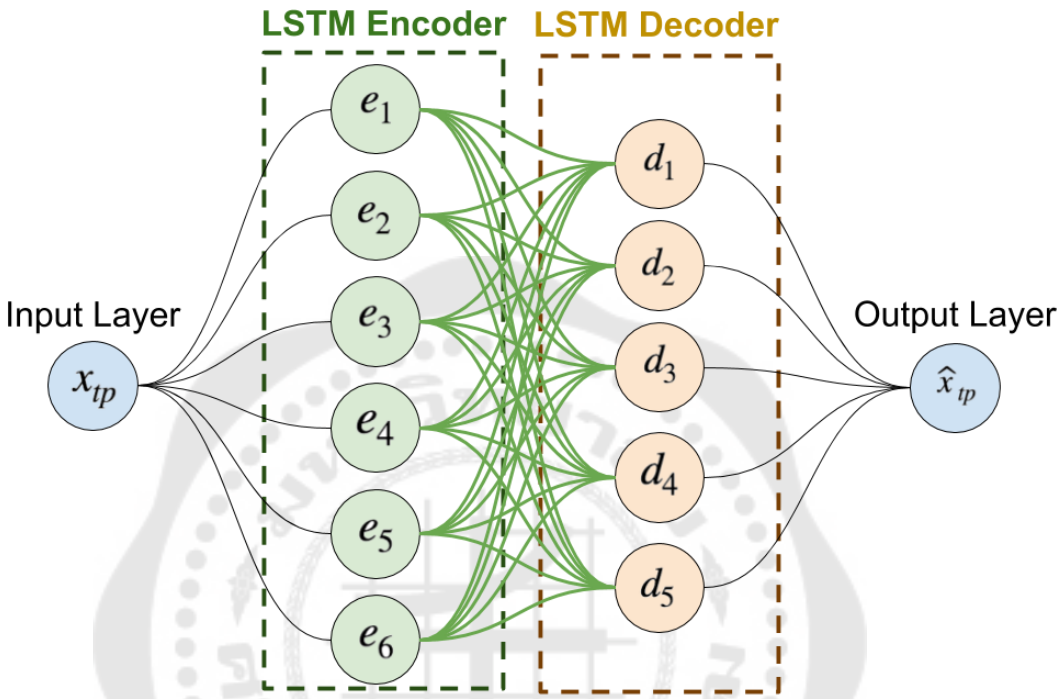
โดยที่

n คือจำนวนมิติของข้อมูลนำเข้า

x_{tp} คือข้อมูลนำเข้า

$\hat{x}_{tp}$ คือข้อมูลนำเข้าที่ถูกแปลงจากชั้นถอดรหัสข้อมูลด้วย LSTM

สามารถแสดงการสกัดคุณลักษณะด้วย LSTM Autoencoder ได้ดังภาพประกอบ 2



ภาพประกอบ 2 แสดงการสกัดคุณลักษณะด้วย LSTM Autoencoder

2.3 ทฤษฎีพื้นฐานเกี่ยวกับ แบบจำลองหน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory)

หน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory) [3] คือ โครงข่ายประสาทเทียมชนิดหนึ่งที่ถูกนำเสนอโดย Sepp Hocheriter และ Jurgen Schmidhuber ในปี 1997 LSTM ถูกพัฒนามาจากโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) ใช้กับข้อมูลที่มีความต่อเนื่องกันเช่น ข้อมูลเสียง ข้อมูลข้อความ ข้อมูลรูปภาพ หรือข้อมูลอนุกรมเวลา แต่เนื่องด้วยปัญหาของ RNN เมื่อข้อมูลขนาดยาวขึ้น ค่า Gradient ของ RNN ยิ่งมีค่าน้อยลงหรือเรียกอีกชื่อหนึ่งว่า Vanishing Gradient Problem ซึ่งมาจาก Activation Function ของ RNN เองที่ใช้ Hyperbolic Tangent (tanh) ผู้สร้าง LSTM จึงได้นำ RNN มาปรับปรุงใหม่โดยเพิ่ม Function เพิ่มเติมจาก tanh และประกอบกับประตูที่มี 3 ประตู ได้แก่ ประตูทางเข้า (Input Gate) ประตูลืม (Forgot Gate) และ ประตูทางออก (Output Gate) พร้อมกับสถานะหน่วยอัพเดท (Update Cell State)

ประตูลืม (Forgot Gate)

Forgot Gate คือประตูที่ถูกออกแบบมาเพื่อพิจารณาข้อมูลนำเข้าเวลา t ตัวแปร x_t นั้นสมควรนำมาใช้พิจารณาภายในหน่วยประมวลผลของ LSTM หรือไม่ โดยวิธีการพิจารณานั้นใช้ผลลัพธ์ในเวลาก่อน (h_{t-1}) เข้ากับข้อมูล ณ เวลา t (x_t) รวมกับ sigmoid function สามารถแสดงได้สมการที่ (10)

$$f_t = \sigma\left(w_f \cdot [h_{t-1}, \mathrm{x_t}] + b_f\right) \quad (10)$$

โดยที่

f_t คือผลลัพธ์ของ Forgot gate ณ เวลา t

σ คือ sigmoid function

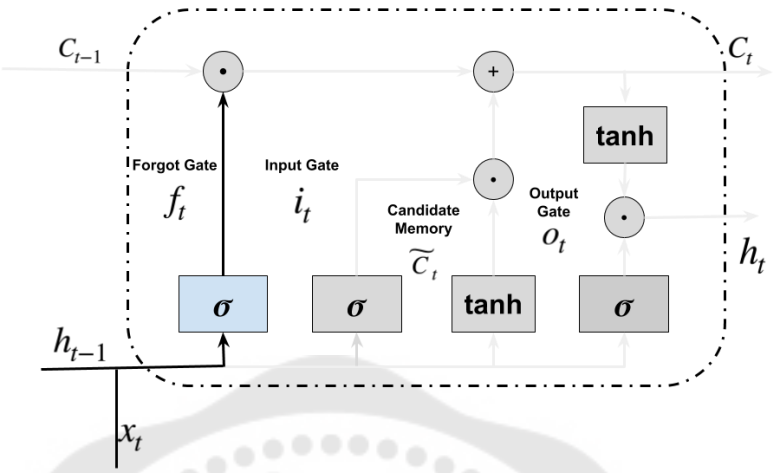
w_f คือค่า weight ของ Forgot gate

h_{t-1} คือคำตอบของเวลา t – 1

$\mathrm{x_t}$ คือข้อมูลนำเข้าเวลา t

b_f คือค่า bias ของ Forgot gate

การทำงานดังภาพประกอบ 3



ภาพประกอบ 3 การทำงานของประตูลืม (Forgot Gate) ที่มา [12]

ผลลัพธ์ได้จากขั้นตอนนี้จะมีค่าอยู่ระหว่าง 0 กับ 1 และ ถ้าผลลัพธ์มีค่าเท่ากับ 0 หมายถึง ข้อมูลนั้นจะไม่ถูกนำไปใช้ในขั้นตอนอื่น ๆ และผลลัพธ์ 1 หมายถึงข้อมูลนั้นถูกนำไปใช้ในส่วนอื่นต่อไป

ประตูทางเข้า (Input Gate)

เป็นประตูที่ออกแบบมาเพื่อเปิดรับข้อมูล ณ เวลา t (x_t) แล้วนำข้อมูลมาบันทึกในหน่วยประมวลผลย่อยของ LSTM หรือมีอีกชื่อหนึ่งคือ การเขียนข้อมูล “write” ซึ่งการคำนวณของประตูทางเข้ามีการคำนวณเหมือนกับประตูลืมซึ่งแสดงได้ดังสมการ (11)

$$i_t = \sigma\left(w_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{11}$$

โดยที่

i_t คือผลลัพธ์ของ input gate ณ เวลา t

σ คือ sigmoid function

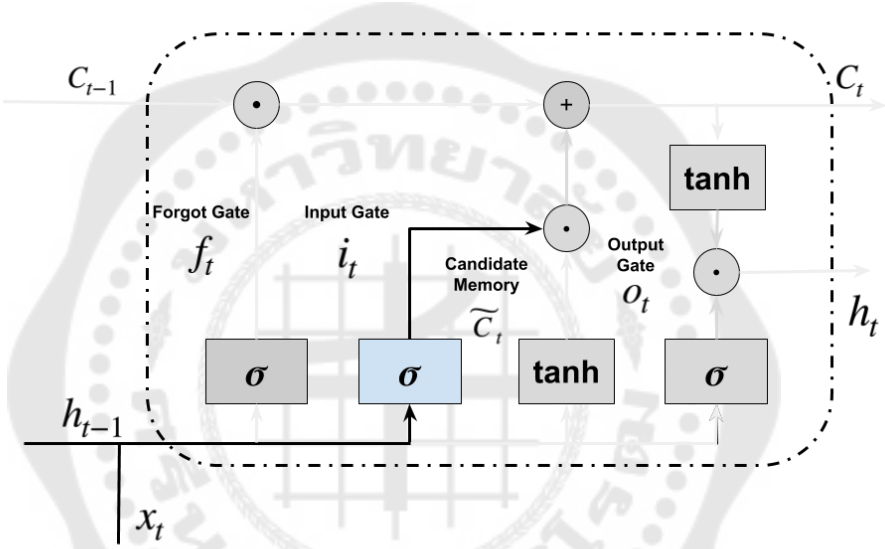
w_i คือค่า weightของ input gate

h_{t-1} คือคำตอบของเวลา t – 1

x_t คือข้อมูลนำเข้าเวลา t

b_i คือค่า bias ของ input gate

การทำงานดังภาพประกอบ 4



ภาพประกอบ 4 การทำงานของประตูทางเข้า (Input Gate) ที่มา [12]

สถานะหน่วยอัพเดท (Update Cell State)

ข้อมูล ณ เวลา t ($\mathbf{X_t}$) จะถูกนำมาแปลงโดยใช้ Hyperbolic Tangent เพื่อแปลงข้อมูลเป็นค่า Context ทางเลือก ($C_t^{\sim}$) เพื่อใช้เปรียบเทียบกับ Context ของเวลาก่อนหน้า t – 1 (C_{t-1}) เพื่อคำนวณ Context ของเวลา t (C_t) สามารถแสดงได้ดังสมการ (12) และ (13)

$$C_t^{\sim} = \tanh(w_c \cdot [h_{t-1}, x_t] + b_c) \quad (12)$$

โดยที่

$\widetilde{C_t}$ คือค่า Context ทางเลือก ณ เวลา t

σ คือ sigmoid function

w_c คือค่า weight ของ Update Cell State

h_{t-1} คือคำตอบของเวลา t – 1

x_t คือข้อมูลนำเข้าเวลา t

b_c คือค่า bias ของ Update Cell State

$$C_t = (f_t \cdot C_{t-1}) + (\widetilde{C_t} \cdot i_t) \quad (13)$$

โดยที่

C_t คือค่า context ณ เวลา t

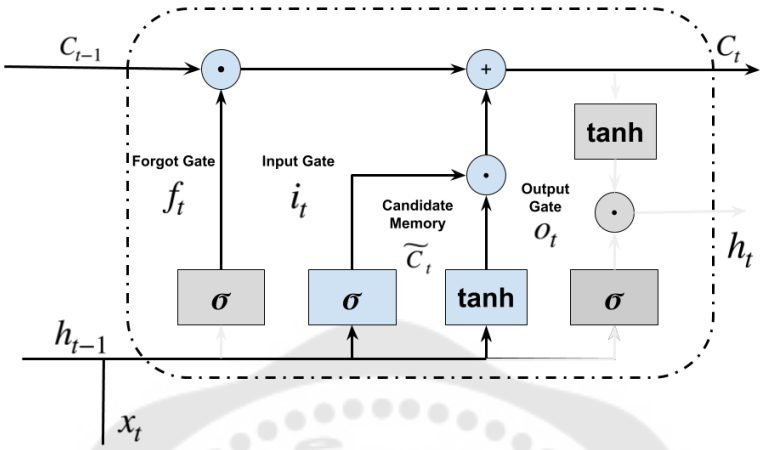
f_t คือผลลัพธ์ของ Forgot gate ณ เวลา t

C_{t-1} คือค่า context ณ เวลา t -1

$\widetilde{C_t}$ คือค่า context ทางเลือก ณ เวลา t

i_t คือผลลัพธ์ของ input gate ณ เวลา t

ทำงานดังภาพประกอบ 5



ภาพประกอบ 5 สถานะหน่วยอัพเดท (Update Cell State) ที่มา [12]

ประตูทางออก (Output Gate)

เป็นประตูที่ใช้ในการเตรียม ค่า h_t ซึ่งเป็นผลลัพธ์ของหน่วย โดยเริ่มจากการนำผลลัพธ์ของหน่วยก่อนหน้า (h_{t-1}) ผ่าน sigmoid function จะได้ค่าของประตูทางออก ณ เวลา t (o_t) หลังจากนั้นนำค่า o_t มากระทำการ pointwise กับค่า Context ณ เวลา t ที่ได้จากผลลัพธ์ของสถานะหน่วยอัพเดทนำไปสู่ผลลัพธ์ ณ เวลา t (h_t) ดังสมการ (14) และ (15)

$$o_t = \sigma\left(w_o \cdot [h_{t-1}, x_t] + b_0\right) \tag{14}$$

โดยที่

o_t	คือผลลัพธ์ของ output gate ณ เวลา t
σ	คือ sigmoid function
w_o	คือค่า weight ของ output gate
h_{t-1}	คือคำตอบของเวลา t – 1
x_t	คือข้อมูลนำเข้าเวลา t

b_0 คือค่า bias ของ output gate

$$h_t = o_t \cdot \tanh(C_t) \quad (15)$$

โดยที่

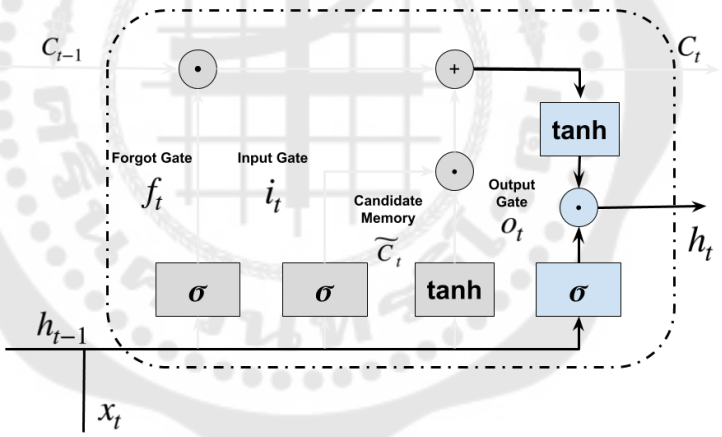
h_t คือผลลัพธ์ ณ เวลา t

o_t คือผลลัพธ์ของ output gate ณ เวลา t

$\tanh$ คือ Hyperbolic Tangent Function

C_t คือค่า context ณ เวลา t

สามารถอธิบายดังภาพประกอบ 6ภาพประกอบ 6



ภาพประกอบ 6 รูปการทำงานของประตูทางออก (Output Gate) ที่มา [12]

2.4 ทฤษฎีเกี่ยวกับการวัดประสิทธิภาพของแบบจำลอง

เนื่องจากแบบจำลองในงานวิจัยนี้เป็นการพยากรณ์ผลแบบถดถอย (Regression) เครื่องมือที่ใช้วัดความถูกต้องของแบบจำลองแบบถดถอยที่ใช้กันอย่างแพร่หลายตัวหนึ่งคือ Root Mean Square Error (RMSE) และ Mean Absolute Error (MAE)

2.4.1 Root Mean Square Error (RMSE)

Root Mean Square Error คือค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ที่คำนวณได้จากการหาค่าผลต่างของค่า PM2.5 ที่ได้จากการพยากรณ์กับค่า PM2.5 จริง แสดงตามสมการที่ (16)

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{n}} \sum_{\boldsymbol{t}=\mathbf{1}}^{\boldsymbol{n}} \left(\widehat{\boldsymbol{h}}_{\boldsymbol{t}} - \boldsymbol{h}_{\boldsymbol{t}}\right)^{\mathbf{2}}} \tag{16}$$

2.4.2 Mean Absolute Error (RMSE)

คือค่าความคลาดเคลื่อนเฉลี่ย (Average Prediction Error) ที่คำนวณได้จากการหาค่าผลต่างของค่า PM2.5 ที่ได้จากการพยากรณ์กับค่า PM2.5 จริง แสดงตามสมการที่ (18)

$$\boldsymbol{MAE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{n}} \sum_{\boldsymbol{t}=\mathbf{1}}^{\boldsymbol{n}} \left|\widehat{\boldsymbol{h}}_{\boldsymbol{t}} - \boldsymbol{h}_{\boldsymbol{t}}\right|} \tag{17}$$

ตัวแปรของ RMSE และ MAE สามารถอธิบายร่วมกันได้ดังนี้

โดยที่

n คือจำนวนของค่า PM2.5 ทั้งหมดที่นำมาทดสอบ

h_t คือค่า PM2.5 ณ เวลา t ซึ่งเป็นค่าจริง

$\widehat{h}_i$ คือค่า PM2.5 ณ เวลา t จากการพยากรณ์จากแบบจำลอง

โดยผลลัพธ์ที่ได้จากเครื่องมือ RMSE และ MAE มีหน่วยวัดเดียวกันซึ่งในการวิจัยนี้วัดค่าความคลาดเคลื่อนของฝุ่น PM2.5 ในหน่วยวัดไมโครกรัมต่อลูกบาศก์เมตร ค่าความคลาดเคลื่อนที่ต่ำของ RMSE และ MAE แสดงถึงประสิทธิภาพการพยากรณ์ที่มีประสิทธิภาพ แต่เครื่องมือ RMSE มีความไวต่อการคลาดเคลื่อนในการคำนวณมากกว่า MAE ทำให้ RMSE เหมาะกับการใช้วัดประสิทธิภาพในงานที่มีค่าความเคลื่อนเกิดขึ้นเป็นจำนวนมากดังเช่นงานวิจัยนี้

2.5 งานวิจัยที่เกี่ยวข้อง

(1) บทความเรื่อง Long Short-Term Memory Deep Neural Network Model for PM2.5 Forecasting in the Bangkok Urban Area โดย Kankamon Thaweephol และ Nuwee Wiwatwattana [1]

Kankamon Thaweephol และ Nuwee Wiwatwattana ได้ทำการศึกษาแบบจำลอง Deep learning โดยใช้ LSTM และแบบจำลองทางสถิติ ได้แก่ SARIMAX เพื่อเปรียบเทียบประสิทธิภาพในการพยากรณ์ค่าความหนาแน่นฝุ่น PM2.5 ล่วงหน้าเป็นเวลา 24 ชั่วโมงโดยชุดข้อมูลที่ใช้ทดสอบเป็นชุดข้อมูลฝุ่น PM2.5 บริเวณสถานีตำรวจนครบาลโชคชัย วิธีการจัดการข้อมูลสูญหายนั้นในบทความได้ทำการลบแถวข้อมูลนั้นทิ้งหลังจากนั้นนำไปสร้างแบบจำลอง โดยผลลัพธ์ที่ได้แบบจำลอง LSTM ให้ประสิทธิภาพในการพยากรณ์ค่าความหนาแน่นฝุ่น PM2.5 ดีกว่า SARIMAX โดยได้ค่าความคลาดเคลื่อนเมื่อวัดด้วย RMSE อยู่ที่ 6.04 ไมโครกรัมต่อลูกบาศก์เมตร และ MAE อยู่ที่ 4.86 ไมโครกรัมต่อลูกบาศก์เมตร

(2) บทความเรื่อง Time Series and Predictability Analysis of Air Pollutants in Delhi โดย Rashmi Bhardwaj และ Dimple Pruthi [13]

Rashmi Bhardwaj และ Dimple Pruthi ได้ทำการศึกษาการพยากรณ์ Air Pollutants ในเมืองเดลี ประเทศอินเดีย สถานการณ์ที่ใช้ทำการทดสอบคือช่วงก่อนการทำมาตรการณ์วันคู่วันคี่ (odd even schema) ระหว่างดำเนินมาตรการ และหลังจากการดำเนินมาตรการ ตัวแปรที่ใช้ศึกษาได้แก่ NO NO2 NOX SO2 PM2.5 โดยใช้การวิเคราะห์ทาง statistic การวิเคราะห์เชิงถดถอย Hurst Exponent Fractal Dimension และ การวิเคราะห์อนุกรมเวลา ผลลัพธ์ที่ได้คือมลพิษในกรณีนี้ไม่สามารถพยากรณ์ได้และการทำมาตรการวันคู่วันคี่ (Odd Even Schema) ไม่มีผลกระทบกับมลพิษในเดลี

(3) บทความเรื่อง PM10 Density Forecast Model Using Long Short Term Memory โดย Jing-Hwan Park และคณะ [3]

Jing-Hwan Park และคณะได้เสนอวิธีการพยากรณ์ค่าความหนาแน่นของ PM10 โดยใช้แบบจำลอง LSTM เป็นตัวแบบพยากรณ์ พร้อมทั้งนำเสนอการใช้ค่า Moving Average แทนค่า NAN เพื่อเพิ่มประสิทธิภาพของแบบจำลอง พบว่าเมื่อทดสอบเปรียบเทียบระหว่าง LSTM กับ Linear Regression พบว่า RMSE ของ LSTM ดีกว่า Linear Regression 500 เท่า

(4) บทความเรื่อง DeepAirNet:Applying Recurrent Networks for Air Quality Prediction โดย Athira V. และคณะ [14]

Athira V. และคณะได้ทำการศึกษาวิธีการการใช้การเรียนรู้เชิงลึกประเภท Recurrent network ประเภทต่างๆ ได้แก่ RNN LSTM และ GRN ทดสอบกับชุดข้อมูล AirNet ซึ่งประกอบด้วยข้อมูล PM10 PM2.5 NO2 CO O3 SO2 AQI Temperature Relative Humidity U-Wind components V-wind Component Precipitation Rate และ Total Cloud cover โครงสร้างของ Neural Network ในงานวิจัยประกอบด้วย 32 process unit ใน 1 layer และใน 1 process unit ประกอบไปด้วย 4 layer แต่ละ processing unit ทำงาน 100 epochs พบว่าโครงข่ายที่สร้างจาก GRU มีประสิทธิภาพดีที่สุดเมื่อเทียบกับ RNN และ LSTM

(5) บทความเรื่อง Assessing The Accuracy of ANFIS, EEMD-GRNN, PCR and MLR Models in Predicting PM2.5 โดย Shadi Ausati และ Jamil Amanollahi [15]

Shadi Ausati และ Jamil Amanollahi ได้นำเสอวิธีการที่ชื่อว่า EEMD-GRNN นำมาพยากรณ์เหตุการณ์ค่าฝุ่น PM2.5 เปรียบเทียบกับแบบจำลองแบบ Linear ได้แก่ MLR PCR และ ANFIS คุณสมบัติที่ใช้ ได้แก่ PM2.5 PM10 SO2 NO2 CO O3 Average minimum temperature Average maximum temperature Average atmospheric pressure Daily total precipitation Daily relative humidity level of the air daily wind speed ปรากฏว่า EEMD-GRNN ให้ผลลัพธ์ดีที่สุดเมื่อเทียบกับวิธีการอื่น

(6) บทความวิจัยเรื่อง Hybrid Speech Enhancement with Wiener Filters and Deep LSTM Denoising Autoencoders โดย Marvin Coto-Jimenez, John Goddard-Close, Leandro Do Persia และ Hugo Leonardo Rufiner [16]

งานวิจัยชิ้นนี้คณะผู้วิจัยได้นำเสอวิธีการเพิ่มประสิทธิภาพของการพูดโดยใช้ Wiener filters ร่วมกับ Deep Learning พร้อมทั้งขจัด Noise ด้วย Autoencoder โดยทดสอบกับชุดข้อมูลเสียงแบบ SLT ของ Carnegie Mellon University ผู้วิจัยได้เลือกใช้ เสียงจากเพศหญิง 1132 ประโยคแบ่งออกเป็น ชุดฝึกฝน 849 ประโยค ชุดตรวจสอบ 233 ประโยค และชุดทดสอบ 50 ประโยค ผู้วิจัยใช้เครื่องมือในการวัดสองตัวคือ Perceptual Evaluation of Speech Quality (PESQ) และ Weight-slope โดย การทดอบใช้ LSTMA และ HW-LSTMA พบว่า HW-LSTMA ให้ผลลัพธ์ดีที่สุด ที่ SNR -10 45% เมื่อเทียบกับใช้ Wiener filer เพียงอย่างเดียว

(7) บทความเรื่อง Solar Power Generation Forecast Based on LSTM โดย Jinxia Zhang และ Yuanying Chi [6]

บทความนี้ได้นำเสนอเทคนิคการพยากรณ์การสร้างพลังงานแสงอาทิตย์โดยใช้ LSTM เป็นแบบจำลองแต่เนื่องด้วยมีตัวแปรในที่ใช้เป็นข้อมูลนำเข้าจำนวนมากถึง 49 ตัว ผู้วิจัยจึงได้ใช้

เทคนิค Principal Component Analysis (PCA) เข้ามาช่วยลดมิติของข้อมูลจาก 49 เหลือ 23 ตัว จากค่า Variance ration สะสม ที่ 0.999 เมื่อนำข้อมูลที่ลดมิติแล้วไปสร้างแบบจำลองแล้วใช้ RMSE เป็นเครื่องมือวัดประสิทธิภาพพบว่า ผลลัพธ์ที่แบบจำลองพยากรณ์ได้มานั้นมีค่า RMSE อยู่ที่ 0.0254 ซึ่งมีประสิทธิภาพดีกว่าแบบจำลองที่สร้างจากข้อมูลที่ไม่ได้ลดมิติด้วย PCA

(8) บทความวิจัยเรื่อง A PCA LSTM Model For Stock Index Prediction โดย De-hua Liu และ Jing-jing Wang [7]

บทความวิจัยชิ้นนี้มีจุดประสงค์เพื่อทำการพยากรณ์หุ้นโดยนำข้อมูลมาจาก CIS INDEX 300 ตั้งแต่วันที่ 01/04/2005 ถึง 09/10/2017 มีตัวแปรได้แก่ Open Close Low High Turnover Volume และ %Cha พร้อมทั้งใช้ PCA เป็นเครื่องมือช่วยลดมิติของข้อมูล โดยเลือก 4 ตัวแปร จากค่า Variance ration สะสม ที่ 0.998 นำเข้าแบบจำลอง LSTM ผลลัพธ์จากการพยากรณ์ถูกวัดด้วยค่า MAE ได้ค่าออกมาที่ 33.1794 ซึ่งดีกว่า แบบจำลอง LSTM ซึ่งสร้างจากข้อมูลที่ไม่ได้มีการลดมิติ

(9) บทความวิจัยเรื่อง Mood Detection From Daily Conversational Speech Using Denoising Autoencoder and LSTM โดย Kun-Yi-Huang, Chung-Hsienm Hsiang Su และ Hisang-Chi Fu [8]

งานวิจัยชิ้นนี้ได้นำเสนอวิธีการตรวจจับอารมณ์โดยใช้ Autoencoder และ LSTM โดยใช้ Model ที่เปรียบเทียบคือ HMM โดยใช้ฐานข้อมูลแบ่งออกเป็น 3 ส่วน ได้แก่ MHMC คือ emotion database EP-DB คือ Emotion with Personality database และ LM-DB คือ Long-Term database วิธีการที่นำเสนอของงานวิจัยนี้แบ่งออกเป็นสามส่วนได้แก่ MHCP สร้างการทำนาย emotion โดย SVM จากนั้นทำการขจัด noise ด้วย Denosing Autoencoding หลังจากนั้น ทำ mood detect ด้วย LSTM ผลลัพธ์ที่ได้จากวิธีการที่นำเสนอเมื่อเปรียบเทียบกับ HMM พบว่า Accuracy ของวิธี LSTM ดีกว่าวิธี HMM อยู่ 5%

(10) บทความวิจัยเรื่อง Speech Emotion Recognition Using Autoencoder Bottleneck Features and LSTM โดย Kun-Yi Huang และคณะ [17]

งานวิจัยชิ้นนี้นำเสนอวิธีการใช้ Speech Recognition โดยใช้ Feature สองตัวที่ชื่อว่า Deep Scattering Spectrum (DSS) และ Low Level Descriptions (LDDs) ร่วมกับ Autoencoder เพื่อทำ extract feature เพื่อลดมิติของข้อมูลร่วมด้วยกับ LSTM โดยใช้ข้อมูลในการทดลองคือ MHMC เมื่อทดสอบ การใช้ Bottleneck เพื่อ extract feature กับ raw data พบว่า

เมื่อใช้ LLD ร่วมกับ DSS และ Autoencoder แล้วได้ผลลัพธ์ดีขึ้นกว่าเดิมประมาณ 5% เมื่อใช้เครื่องมือวัดคือ accuracy

(11) บทความวิจัยเรื่อง A Novel Approach for Automatic Novelty Detection Using A Denoising Autoencoder with Bidirectional LSTM Neural Network โดย Erik Marchi และคณะ [9]

งานวิจัยชิ้นนี้ นำเสนอวิธีการจำแนกตรวจจับสัญญาณ แบบ acoustic หรือ abnormal วิธีดำเนินการวิจัยของกลุ่มผู้วิจัย แบ่งเป็นขั้นตอนต่าง ๆ ได้แก่ ขั้นตอนการสกัดคุณลักษณะผู้วิจัยได้ใช้วิธีการที่ชื่อว่า Auditory Spectral Feature ซึ่งได้มาจากการคำนวณ Short Time Fourier Transformation (STFT) โดยใช้กรอบเวลา 30 ms และ หยุด 10 ms เพื่อสร้าง frame ถัดไป ขั้นตอนถัดมาคือขั้นตอนการขจัด noise ผู้วิจัยได้ใช้วิธีการที่ชื่อว่า Denoising Autoencoder (DEA) ในการกำจัด noise ขั้นตอนที่สามการสร้างแบบจำลองผู้วิจัยได้เลือกใช้แบบจำลอง Bidirectional Long-Short Term Memory โดยที่โครงสร้างที่ให้ผลลัพธ์ดีที่สุดของโครงข่ายคือสร้าง 6 ชั้นและมี 216 LSTM ในแต่ละ layer ทำการฝึกฝน 100 ครั้งสำหรับแต่ละ network และค่าที่ใช้ update gradient decent ในขั้นตอน backpropagation คือ Sum of Square Error (SSE) ขั้นตอนสุดท้ายคือขั้นตอนวัดผลได้ใช้เครื่องมือที่ชื่อว่า F-Measure ในการวัดผล โดยการทดสอบได้ใช้แบบจำลอง GMM HMM LSTM-CAE BLSTM-CAE LSTM-DAE BLSTM-DAE ผลลัพธ์ที่ดีที่สุดที่ได้คือ BLSTM-DAE และ ค่า F-Measure เท่ากับ 94.7

(12) บทความวิจัยเรื่อง Time-Series Extreme Event Forecasting with Neural Networks at Uber โดย Nikolay Laptev และคณะ [11]

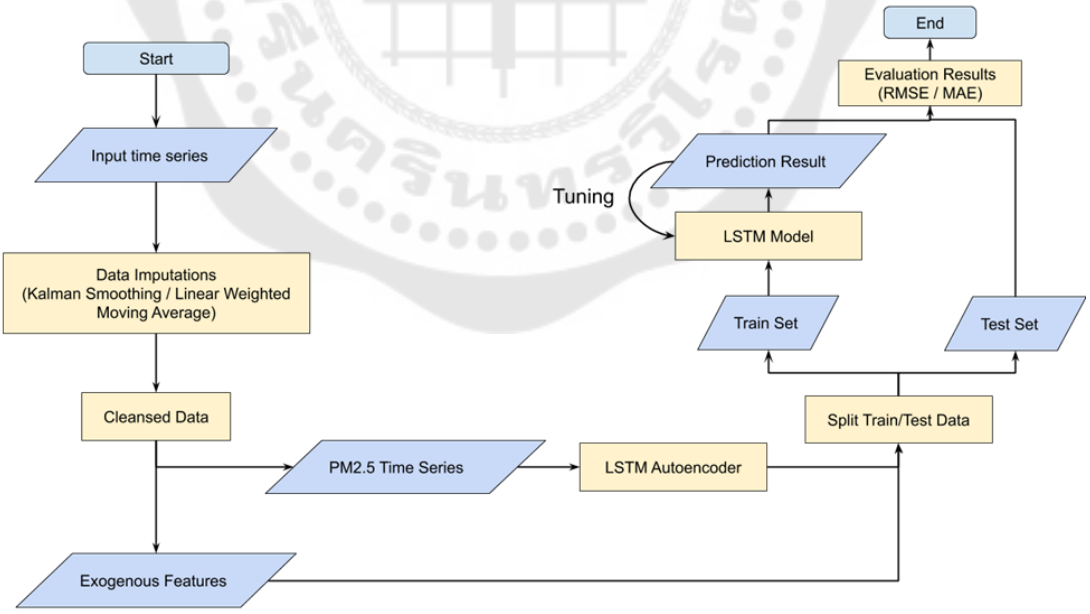
งานวิจัยชิ้นนี้ Nikolay Laptev และคณะ ได้เสนอวิธีเพื่อแก้ปัญหาการเตรียมจำนวนรถยนต์ของ Uber เพื่อที่จะสามารถให้บริการลูกค้าใช้ช่วงเวลาพิเศษของปี เช่น ปีใหม่ หรือวันหยุดพิเศษเพราะช่วงนั้นมีความต้องการการใช้บริการของ Uber ที่ผันผวนเป็นอย่างมาก ผู้เขียนงานวิจัยได้เสนอการใช้ LSTM Autoencoder เป็นเครื่องมือในการสกัดคุณลักษณะจากข้อมูล trip ย้อนหลัง และนำคุณลักษณะที่สกัดได้ไปรวมกับข้อมูลประกอบอื่น แล้วสร้างแบบจำลอง LSTM พบว่าผลลัพธ์แบบจำลอง LSTM ที่ทำการสกัดคุณลักษณะ ดีขึ้นกว่า 2%-18% มากกว่าแบบจำลองที่ไม่ได้สกัดคุณลักษณะ

บทที่ 3
วิธีดำเนินงานวิจัย

งานวิจัยนี้ได้นำเสนอการทดลองเพื่อศึกษาประสิทธิภาพแบบจำลอง LSTM ที่สร้างจากชุดข้อมูลที่ทำการแทนที่ข้อมูลสูญหาย (Data Imputation) ร่วมกับการทำการสกัดคุณลักษณะ (Feature Extraction) ในการพยากรณ์ค่าความหนาแน่นฝุ่นละออง PM2.5 เปรียบเทียบกับแบบจำลอง LSTM ที่ทำการลบแถวที่เกิดข้อมูลสูญหาย [1] ซึ่งมีขั้นตอนการดำเนินงานวิจัยดังต่อไปนี้

1. การสำรวจข้อมูล (Data Exploration)
2. การแทนที่ข้อมูลสูญหาย (Data Imputation)
3. การแปลงข้อมูล (Data Transformation)
4. การสกัดคุณลักษณะ (Feature Extraction)
5. การสร้างแบบจำลอง LSTM และวัดประสิทธิภาพ

วิธีการดำเนินการวิจัยสามารถแสดงได้ดังภาพประกอบ 7



ภาพประกอบ 7 แผนผังการดำเนินการสร้างแบบจำลอง LSTM จากชุดข้อมูลที่ทำการแทนที่ข้อมูล (Data Imputation) ร่วมกับการสกัดคุณลักษณะ (Feature Extraction)

3.1 **การสำรวจข้อมูล** (Data Exploration)

ในงานวิจัยชิ้นนี้ได้รับการอนุเคราะห์ข้อมูลจากกรมควบคุมมลพิษโดยข้อมูลฝุ่น PM2.5 จากบริเวณจุดวัดสถานีตำรวจนครบาลโชคชัย ตั้งแต่วันที่ 7 มิถุนายน 2560 (2017:06:07 00:00:00) จนถึงวันที่ 30 กรกฎาคม 2561 (2017:07:30 23:00:00) มีข้อมูลที่ใช้ในการทดลองทั้งหมด 9336 แถว คิดเป็นจำนวน 389 วัน คำอธิบายของชุดข้อมูลแสดงได้ตามตาราง 1

ตาราง 1 คำอธิบายตัวแปรที่ใช้ในการวิจัย

ชื่อตัวแปร	คำอธิบาย	หน่วยวัดและความสูงอุปกรณ์
PM2.5	ค่าความหนาแน่นของฝุ่น PM2.5	at 3 m (มคก./ลบ.ม.)
CO	ปริมาณ CO	at 3 m (ppm)
NO	ปริมาณ NO	at 3 m (ppb)
NO2	ปริมาณ NO2	at 3 m (ppb)
NOX	ปริมาณ NOX	at 3 m (ppb)
PM10	ค่าความหนาแน่นของฝุ่น PM10	at 3 m (มคก./ลบ.ม.)
Wind Speed	ความเร็วของลม	at 10 m (m/s)
Wind Direction	ทิศทางลม	at 10 m (Deg.M)
Temperature	อุณหภูมิ	at 2 m
Real Humidity	ความชื้นของอากาศ	at 2 m (%RH)
Pressure	ความกดอากาศ	at 2 m (mmhg)
Rain	ปริมาณฝน	at 3 m (mm)

สามารถแสดงข้อมูลทางสถิติได้ดังตาราง 2

ตาราง 2 ข้อมูลทางสถิติของชุดข้อมูล PM2.5 และ ข้อมูลทางอุตุนิยมวิทยาที่ใช้ในงานวิจัย

ชื่อตัวแปร	Mean	S.D.	Min	Max
PM2.5	22.209	13.712	1	138
CO	0.890	0.528	0	5.2
NO	37.278	40.336	0	311
NO2	22.764	12.992	5	91
NOX	59.670	47.091	7	366
PM10	47.243	26.359	1	196
Wind Speed	0.506	0.409	0	3
Wind Direction	157.116	94.684	0	359
Temperature	28.825	2.904	17.2	36.6
Real Humidity	66.412	14.274	23	98
Pressure	754.944	2.102	749	765
Rain	0.240	2.074	0	52.8

3.2 การแทนที่ข้อมูลสูญหาย (Data Imputation)

เนื่องจากการพยากรณ์ค่าความหนาแน่นของฝุ่นละออง PM2.5 เป็นการใช้ข้อมูลนำเข้าแบบอนุกรมเวลา (Time Series) กล่าวคือมีการใช้ Feature จากช่วงเวลาก่อนหน้าเพื่อใช้ในการพยากรณ์ค่าความหนาแน่นของฝุ่นละออง PM2.5 ล่วงหน้า 24 ชั่วโมง แต่ชุดข้อมูลที่ได้ใช้ในการวิจัยนี้มีข้อมูลสูญหายอยู่จำนวนหนึ่ง สามารถแสดงได้ดังตาราง 3

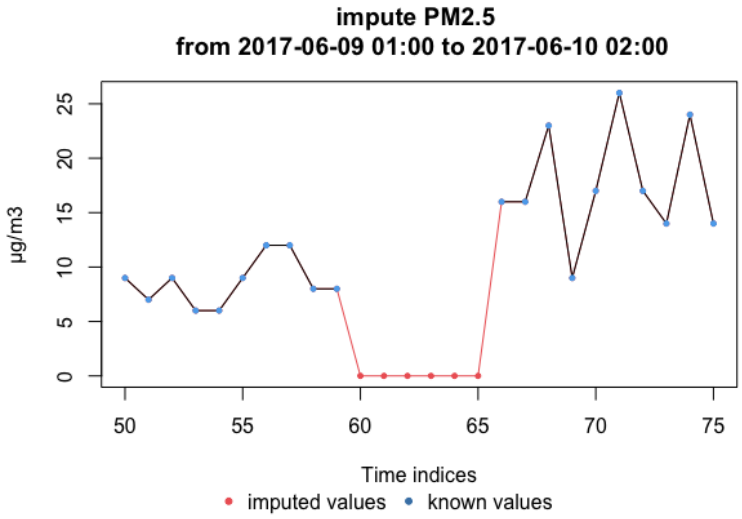
ตาราง 3 แสดงจำนวนค่าว่างของตัวแปรที่ใช้ในงานวิจัย

ชื่อตัวแปร	จำนวนค่าว่าง	% ค่าว่างของข้อมูล
PM2.5	91	0.977 %
CO	564	6.055 %
NO	471	5.057 %
NO2	471	5.057 %

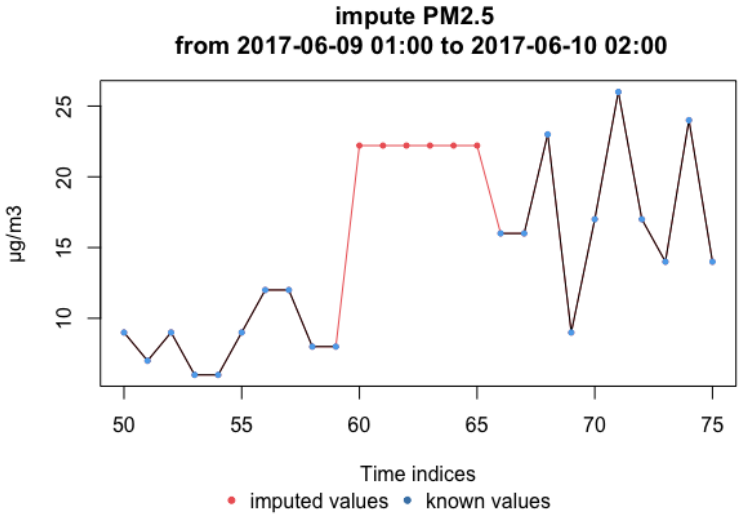
ตารางที่ 3 (ต่อ)

ชื่อตัวแปร	จำนวนค่าว่าง	% ค่าว่างของข้อมูล
NOX	471	5.057 %
PM10	75	0.805 %
Wind Direct	15	0.161 %
Wind Speed	16	0.172 %
Temperature	15	0.161 %
Real hum	15	0.161 %
Pressure	12	0.129 %
Rain	18	0.193 %
Rows	655	7.020 %

จากตาราง 3 ถ้าหากทำการลบข้อมูลที่มีค่าสูญหายออกไปพบว่ามีข้อมูลที่ถูกลบไปเป็นจำนวน 655 แถว ทำให้ข้อมูลอนุกรมเวลาเกิดความไม่ต่อเนื่องของข้อมูล ผู้วิจัยได้เห็นว่าการแทนข้อมูลด้วยค่าเดียวกับข้อมูลที่สูญหาย ได้แก่ ค่าคงที่ เช่น 0 ดังภาพประกอบ 8 หรือ การหาค่าทางสถิติอย่างง่ายเช่น การใช้ค่าเฉลี่ยดังภาพประกอบ 9 อาจก่อให้เกิดรูปแบบข้อมูลที่ผิดปกติซึ่งส่งผลให้ประสิทธิภาพของแบบจำลอง LSTM ลดลง



ภาพประกอบ 8 ตัวอย่างข้อมูลสูญหายที่ถูกแทนที่ด้วยค่าคงที่ 0



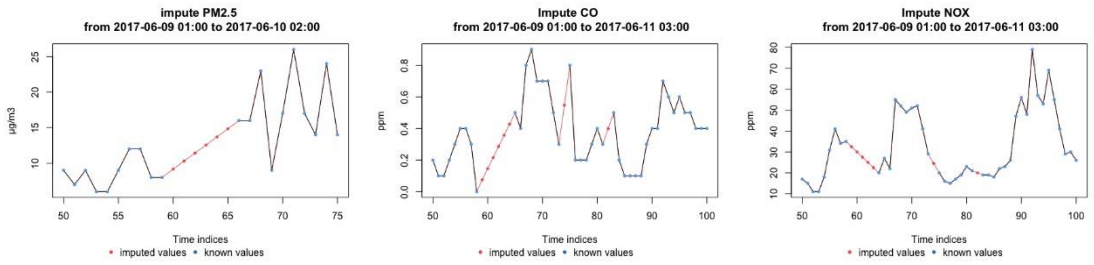
ภาพประกอบ 9 ตัวอย่างข้อมูลสูญหายที่ถูกแทนที่ด้วยค่าเฉลี่ย

ผู้วิจัยจึงเลือกใช้เทคนิค Linear Weighted Moving Average (LWMA) [3] เพราะมีการนำข้อมูลย้อนหลังและข้อมูลล่วงหน้ารวมร่วมกับการให้น้ำหนักความสำคัญของข้อมูล (Weight) มาคำนวณข้อมูลที่สูญหายไป และ Kalman Smoothing (KS) [5] ที่ใช้การคำนวณข้อมูลสูญหายด้วยเทคนิคการป้อนกลับของข้อมูลสูญหายที่รับการแก้ไขแล้วกลับไปคำนวณข้อมูลสูญหายถัดไป ผู้วิจัยจึงได้เลือกใช้ library ที่ชื่อ imputeTS [18] เป็นเครื่องมือ แบ่งออกเป็น 2 ส่วนดังนี้

1. การแทนที่ข้อมูลสูญหายด้วย KS
2. การแทนที่ข้อมูลสูญหายด้วย LWMA

3.2.1 การแทนที่ข้อมูลสูญหายด้วย Kalman Smoothing

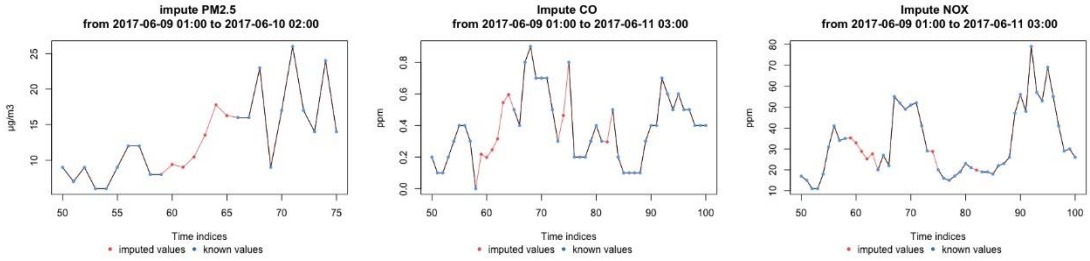
การแทนที่ข้อมูลสูญหายด้วยเทคนิค Kalman Smoothing ได้เลือกใช้ parameter เริ่มต้นของ imputeTS นั้นคือ Space Model มีค่าเท่ากับ StructTS และ smooth มีค่าเท่ากับ true สามารถแสดงตัวอย่างผลลัพธ์ข้อมูลสูญหายที่ถูกแทนที่ด้วย KS ดังภาพประกอบ 10



ภาพประกอบ 10 ตัวอย่างข้อมูลสูญหายที่ถูกแทนที่ด้วยเทคนิค KS

3.2.2 การแทนที่ข้อมูลสูญหายด้วย Linearly Weighted Moving Average

การแทนที่ข้อมูลสูญหายด้วย ImputeTS ผู้วิจัยกำหนดช่วงในการคำนวณข้อมูลสูญหายไปข้างหน้า 2 ชั่วโมงและย้อนหลัง 2 ชั่วโมง สามารถแสดงตัวอย่างผลลัพธ์การแทนที่ด้วย LWMA ได้ดังภาพประกอบ 11

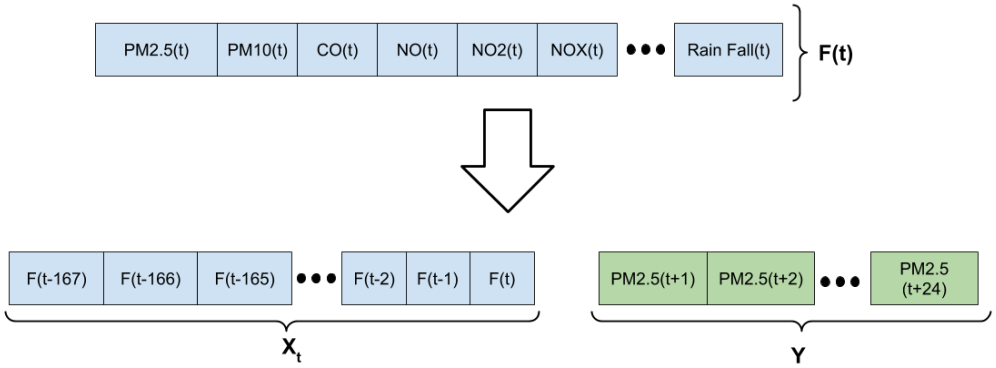


ภาพประกอบ 11 ตัวอย่างข้อมูลสูญหายที่ถูกแทนที่ด้วยเทตนิค LWMA

จากภาพประกอบ 10 การแทนที่ข้อมูลด้วยวิธี KS นั้นลักษณะข้อมูลที่ได้ออกมามีลักษณะเป็นเส้นตรง แต่จากภาพประกอบ 11 การแทนที่ข้อมูลสูญหายด้วย LWMA ได้ผลลัพธ์ที่ซับซ้อนกว่า โดยได้ค่าขึ้นลงตามแนวโน้มของข้อมูล

3.3 **การแปลงข้อมูล** (Data Transformation)

จากชุดข้อมูลที่ได้ทำการวิจัยพบว่า 1 แถวของข้อมูลประกอบไปด้วย 12 features ซึ่งรวมข้อมูลความหนาแน่นฝุ่น PM2.5 มาใช้ในการทำนายค่า PM2.5 ล่วงหน้า เพื่อที่จะสร้างข้อมูลนำเข้าที่ประกอบด้วยข้อมูลย้อนหลัง (lag observation) เป็นเวลา 7 วันหรือ 168 ชั่วโมงและสร้างข้อมูล output หรือ forecast time คือข้อมูลค่าความหนาแน่นฝุ่น PM2.5 ล่วงหน้า 24 ชั่วโมง โดยใช้เทคนิคการเลื่อนข้อมูลเวลา (Time Shift) เข้ามาช่วย สามารถแสดงได้ดังภาพประกอบ 12



ภาพประกอบ 12 การแปลงชุดข้อมูลด้วยการเลื่อนข้อมูล

หลังจากทำการเลื่อนเวลาแล้วจะได้ข้อมูลออกเป็นช่วงเวลา โดย t แทนข้อมูลชั่วโมงปัจจุบันจนไปถึง t-167แทนข้อมูลที่ย้อนหลังไป 168 ชั่วโมง ตัวแปร F แทน Feature ของแต่ละชั่วเวลา และตัวแปร PM2.5 แทนข้อมูลความหนาแน่นของ PM2.5 ล่วงหน้าตั้งแต่ t+1 จนถึง t+24 สามารถแสดงได้กลับชุดข้อมูลจริงดังภาพประกอบ 13

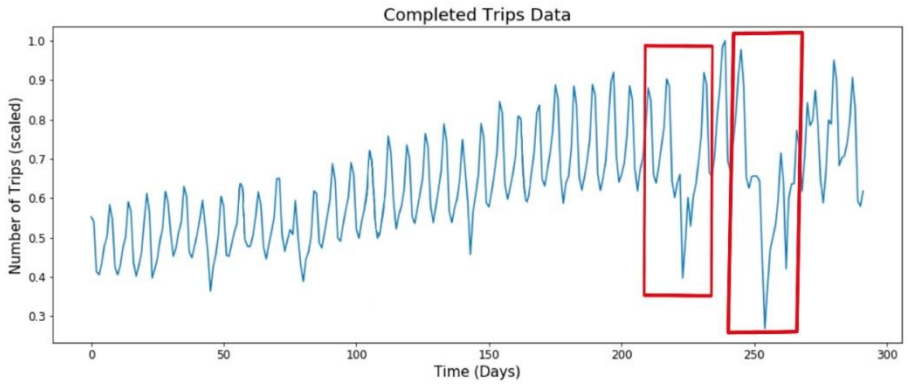
	CO(t-168)	NO2(t-168)	NOX(t-168)	PM10(t-168)	Wind speed(t-168)	Wind dir(t-168)	Temp(t-168)	Rel hum(t-168)	Pressure(t-168)	NO(t-168)
year_month_day_hour										
2017-06-14 00:00:00	0.076923	0.119565	0.089136	0.087179	0.033333	0.676880	0.546392	0.720000	0.3125	0.073955
2017-06-14 01:00:00	0.115385	0.108696	0.125348	0.107692	0.000000	0.465181	0.546392	0.760000	0.3125	0.118971
2017-06-14 02:00:00	0.134615	0.119565	0.172702	0.123077	0.066667	0.337047	0.541237	0.813333	0.2500	0.170418
2017-06-14 03:00:00	0.115385	0.119565	0.158774	0.184615	0.066667	0.259053	0.510309	0.853333	0.2500	0.154341
2017-06-14 04:00:00	0.115385	0.086957	0.100279	0.164103	0.100000	0.239554	0.494845	0.866667	0.2500	0.096463

ภาพประกอบ 13 ตัวอย่างชุดข้อมูลหลังจากทำการเลื่อนเวลา

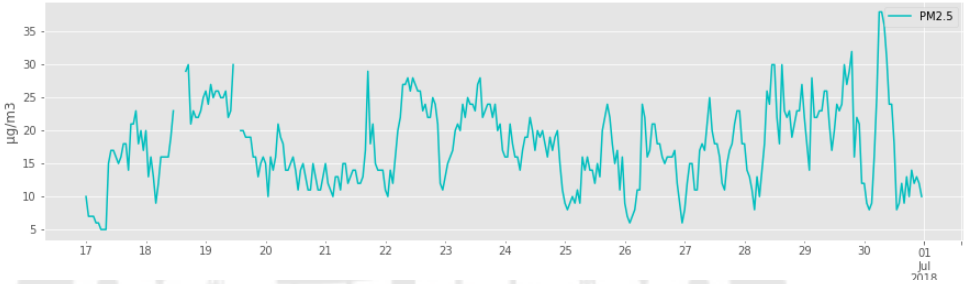
เมื่อทำการเลื่อนข้อมูลเสร็จสิ้นพบว่าข้อมูลที่ได้มี feature ทั้งหมด 2016 feature ต่อ 1 แถว เนื่องจาก lag observation ของการทดลองมีค่าเท่ากับ 168 ชั่วโมง จึงจำเป็นต้องแปลงมิติของข้อมูลให้อยู่ในรูปแบบของอาเรย์สามมิติโดยแกน X แทนจำนวนแถวทั้งหมดของชุดข้อมูล แกน Y แทน lag observation 168 มิติ และ แกน Z แทนจำนวน ของ feature จำนวน 12

3.4 **การสกัดคุณลักษณะ** (Feature Extraction)

ในการวิจัยนี้ผู้วิจัยได้เลือกใช้เทคนิค LSTM AUTENCODER [10, 11] เนื่องจาก LSTM AUTOENCODE มีความสามารถในการสกัดคุณลักษณะจากข้อมูลที่มีความเปลี่ยนแปลงอย่างรวดเร็ว ยกตัวอย่าง เช่นความต้องการของรถให้บริการของ Uber ในช่วงวันหยุดสุดสัปดาห์ แสดงได้ตามภาพประกอบ 14 เมื่อเทียบกับการเปลี่ยนแปลงความหนาแน่นของฝุ่นละออง PM2.5 สามารถแสดงได้ดังภาพประกอบ 15



ภาพประกอบ 14 ตัวอย่างข้อมูลทริปต่อวันของชุดข้อมูล Uber



ภาพประกอบ 15 ตัวอย่างข้อมูลความหนาแน่น PM2.5 บริเวณสถานีตำรวจนครบาลโชคชัย

จะเห็นได้ว่าสองเหตุการณ์มีการเปลี่ยนแปลงของข้อมูลอย่างรวดเร็ว ดังนั้น การสกัดคุณลักษณะด้วย LSTM Autoencoder อาจช่วยให้แบบจำลอง LSTM มีประสิทธิภาพเพิ่มขึ้น

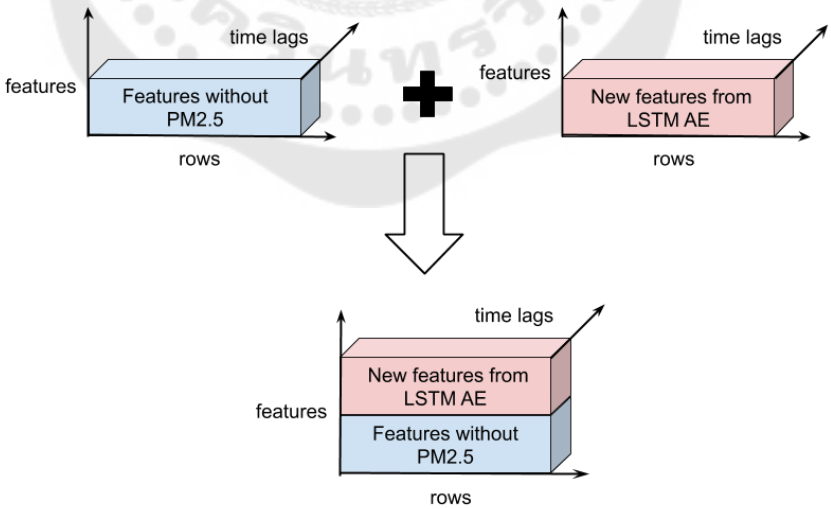
ขั้นตอน Feature Extraction เริ่มจากการแบ่ง PM2.5 จากชุดข้อมูลขั้นตอน 3.4 ซึ่งข้อมูลเป็นอาเรย์ขนาด 9145 x 168 x 12 จากนั้นแบ่งข้อมูลออกเป็นอัตราส่วน 90 ต่อ 10 เพื่อใช้ในการสร้างแบบจำลอง LSTM Autoencoder โดยเครื่องมือที่ใช้สร้างแบบจำลองคือ Keras

การสร้างแบบจำลอง LSTM Autoencoder ชั้นแรกของโครงข่ายประสาทเทียมถูกสร้างจาก ชั้น LSTM ที่รับข้อมูลนำเข้าขนาด 168 x 1 นำมาเข้ารหัสและส่งข้อมูลออกไปสู่ Layer ถัดไป โดยข้อมูลมีขนาด 168 x 128 ชั้นที่สองของโครงข่ายประสาทเทียบถูกสร้างจากชั้น LSTM ที่ทำหน้าที่ถอดรหัสข้อมูลจากข้อมูลนำเข้าเพื่อให้ได้ผลลัพธ์ขนาด 168 x 32 เพื่อส่งไปยังชั้นสุดท้ายของโครงข่ายที่จะประกอบด้วย TimeDisributed ร่วมกับ Dense Layer ผลลัพธ์ที่จะมีขนาด 168 x 1 เช่นเดียวกับข้อมูลนำเข้าในชั้นแรกเพื่อที่จะได้การสกัดคุณลักษณะข้อมูลเดิมที่ครบถ้วน Mean Square Error จึงถูกเลือกใช้เป็นเครื่องมือในการ Optimizer ของโครงข่ายประสาทเทียมนี้และ parameter อื่นแสดงได้ตาราง 4

ตาราง 4 ข้อมูล parameter ที่ใช่สร้าง LSTM Autoencoder

ชื่อ parameter	ค่า parameter
Training Set	90%
Test Set	10%
Optimizer	Adam
Activation Function	Sigmoid
Metrics	MSE
Loss Function	MSE
Batch Size	100
Dropout Rate	0.3
Training epoch	[25,50,100,200,500]

คุณลักษณะที่ถูกสกัดได้คือผลลัพธ์ของชั้นแรกของแบบจำลอง LSTM Autoencoder ที่มีขนาด 9145 x 168 เมื่อนำไปประกอบกับชุดข้อมูลจากขั้นตอนที่ 3.3 ที่ตัด feature PM2.5 เดิมทิ้งไปพบว่าจะได้ข้อมูลชุดใหม่ที่มีขนาด 9145 x 168 x 139 สามารถแสดงได้ดังภาพประกอบ 16



ภาพประกอบ 16 แสดงการผสานข้อมูลที่สกัดคุณลักษณะเข้ากับชุดข้อมูลเดิม

จากการนำทำการค้นหาแบบ Grid (Grid Search) จำนวน Epoch ของ LSTM Autoencoder พบว่าข้อมูลที่ถูกสกัดจากแบบจำลอง LSTM Autoencoder ที่ epoch เท่ากับ 50 เมื่อใช้กับแบบจำลอง LSTM เพื่อพยากรณ์ความหนาแน่นฝุ่นละออง PM2.5 ในงานวิจัยครั้งนี้

3.5 การสร้างแบบจำลอง LSTM และวัดประสิทธิภาพ

ในการสร้างและทดสอบประสิทธิภาพแบบจำลอง LSTM ได้ใช้ Root Mean Square Error (RMSE) และ Mean Absolute Error (MAE) เครื่องมือวัดความแม่นยำของการพยากรณ์ฝุ่นละออง PM2.5 โดยชุดข้อมูลที่ใช้ทำการวิจัยนั้นได้ใช้ชุดข้อมูลฝุ่นละออง PM2.5 บริเวณสถานีตำรวจนครบาลโชคชัยเก็บข้อมูลรายชั่วโมงตั้งแต่วันที่ 7 มิถุนายน 2560 จนถึงวันที่ 30 กรกฎาคม 2561 โดยแบ่งเป็นชุดฝึกฝน (Train Set) และชุดทดสอบ (Test Set)

จากการแบ่งข้อมูลออกเป็น 2 ชุดคือ Train Set และ Test Set โดยข้อมูลนำเข้าของ Train Set 8280 x 168 x 139 และ ลาเบลมีขนาด 8280 x 24 ส่วนข้อมูลนำเข้าของ Test Set 865x168x139 และ ลาเบลมีขนาด 865x24 เพื่อสร้างแบบจำลอง LSTM ที่ให้ผลลัพธ์ที่มีประสิทธิภาพที่สุดผู้วิจัยได้ใช้วิธีการ Grid Search ในการค้นหาค่า epochs โดยค่าที่ค้นหาคือ 10, 20, 30, 40, 50, 60, 70, 80, 90 parameter อื่นเช่น n_neurons = 100 จาก [1] กำหนดค่า Optimizer เป็น RMSprop จาก [3] ส่วน parameter อื่นถูกแสดงในตาราง 5

ตาราง 5 ข้อมูล parameter ที่ใช่สร้าง LSTM

ชื่อ parameter	ค่า parameter
Training Set	90%
Test Set	10%
Optimizer	RMSprop
Learning Rate	0.001
Activation Function	Sigmoid
Loss Function	MSE
Batch Size	49
Training epoch	[25,50,100,200,500]

บทที่ 4
ผลการดำเนินการวิจัย

ในการวิจัยนี้ได้ทำการทดลองนำเทคนิคการแทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะเพื่อทำนายหาความหนาแน่นของฝุ่นละออง PM2.5 ในอีก 24 ชั่วโมงข้างหน้าโดยใช้ข้อมูลนำเข้าเป็นข้อมูลช่วงเวลาก่อนหน้าเป็นเวลา 7 วัน หรือ 168 ชั่วโมง ด้วยแบบจำลองแบบโครงข่ายประสาทเทียม LSTM ชุดข้อมูลที่ใช้ทำการทดลองประกอบไปด้วยข้อมูล PM2.5 และข้อมูลทางอุตุนิยมวิทยา ชุดข้อมูลได้มาจากการเก็บข้อมูลจากบริเวณพื้นที่สถานีตำรวจนครบาลโชคชัย ผู้วิจัยได้ดำเนินการวิจัยโดยกระบวนการและขั้นตอนต่าง ๆ ที่ได้นำเสนอในบทที่ 3 และเพื่อวัดประสิทธิภาพแบบจำลองที่ผ่านการแทนที่ข้อมูลสูญหายและการทำการสกัดคุณลักษณะ ผู้วิจัยได้เลือกแบบจำลอง LSTM ที่สร้างจากการลบแถวข้อมูลสูญหาย [1] เพื่อมาเปรียบเทียบประสิทธิภาพและผลกระทบของการวิจัยนี้ที่มีต่อแบบจำลอง LSTM

การสร้างแบบจำลอง LSTM ในงานวิจัยนี้ เนื่องผู้วิจัยได้ทำการเลือกใช้วิธีการแทนที่ข้อมูลสูญหาย 2 วิถี เมื่อทำการสกัดคุณลักษณะด้วย LSTM AE จะได้แบบจำลอง 4 แบบที่ถูกสร้างขึ้นในงานวิจัยดังนี้

- แบบจำลอง LSTM ที่สร้างจากการแทนที่ข้อมูลสูญหายด้วย KS
- แบบจำลอง LSTM ที่สร้างจากการแทนที่ข้อมูลสูญหายด้วย LWMA
- แบบจำลอง LSTM ที่สร้างจากการแทนที่ข้อมูลสูญหายด้วย KS ร่วมกับการสกัดคุณลักษณะด้วย LSTM AE
- แบบจำลอง LSTM ที่สร้างจากการแทนที่ข้อมูลสูญหายด้วย LWMA ร่วมกับการสกัดคุณลักษณะด้วย LSTM AE

ในการทำการวิจัยนี้ การสร้างแบบจำลอง LSTM มาจากการ Grid Search parameter ที่ชื่อ epochs โดยจำนวน epochs ที่ทำให้แบบจำลองมีประสิทธิภาพสูงสุด สามารถแสดงได้ดังตาราง 6

ตาราง 6 ตารางแสดงจำนวน epoch ในการสร้างแบบจำลองที่ให้ประสิทธิภาพดีที่สุดในแต่ละการทดลอง

ประเภทแบบจำลอง	จำนวน epoch
KS	50
LWMA	30
KS + LSTM AE	40
LWMA + LSTM AE	20

หลังจากสร้างแบบจำลองแล้วแล้วทำการวัดผลแบบจำลองทั้ง 4 แบบ ผู้วิจัยได้แบ่งผลการดำเนินการวิจัยออกเป็น 2 ส่วนดังต่อไปนี้

1. ผลลัพธ์การพยากรณ์โดยการใช้เทคนิคแทนที่ข้อมูลด้วย KS และ การพยากรณ์โดยใช้เทคนิคแทนที่ข้อมูลด้วย KS ร่วมกับการสกัดคุณลักษณะด้วย LSTM Autoencoder
2. ผลลัพธ์การพยากรณ์โดยการใช้เทคนิคแทนที่ข้อมูลด้วย LWMA และ การพยากรณ์โดยใช้เทคนิคแทนที่ข้อมูลด้วย LWMA ร่วมกับการสกัดคุณลักษณะด้วย LSTM Autoencoder

4.1 ผลลัพธ์การพยากรณ์โดยการใช้เทคนิคแทนที่ข้อมูลด้วย KS และ การพยากรณ์โดยใช้เทคนิคแทนที่ข้อมูลด้วย KS ร่วมกับการสกัดคุณลักษณะด้วย LSTM Autoencoder

จากการแทนที่ข้อมูลสูญหายโดยเทคนิค KS กับชุดข้อมูล PM2.5 ที่ประกอบไปด้วยข้อมูลทางอุตุนิยมวิทยาจากบริเวณสถานีตำรวจนครบาลโชคชัย ผลการพยากรณ์พบว่าค่าความหนาแน่นฝุ่น PM2.5 แบบจำลอง LSTM สร้างขึ้น แล้ววัดค่าความคลาดเคลื่อนด้วย RMSE พบว่าประสิทธิภาพดีขึ้นกว่าแบบจำลองที่ลบแถวข้อมูลสูญหายทิ้ง คิดเป็นอัตราส่วน 7.12% โดยค่าความคลาดเคลื่อน RMSE เฉลี่ยอยู่ที่ 5.61 ส่วนเมื่อวัดด้วยเครื่องมือวัด MAE แบบจำลองที่ใช้เทคนิค KS แทนที่ข้อมูลสูญหาย ให้ความคลาดเคลื่อนเฉลี่ยอยู่ที่ 4.33 ซึ่งให้ผลลัพธ์ดีขึ้นจากการลบข้อมูลทิ้งอยู่ 10.91% หลังจากนั้นทำการสกัดคุณลักษณะข้อมูลด้วย LSTM Autoencoder ที่ Epochs 50 แล้วทำการสร้างแบบจำลอง LSTM เมื่อวัดประสิทธิภาพด้วยเครื่องมือ RMSE แบบจำลองที่ใช้ข้อมูลที่ผ่านการสกัดคุณลักษณะมีค่าความคลาดเคลื่อนเฉลี่ยวัดได้ 5.64 ซึ่งให้ผลลัพธ์ดีกว่าแบบจำลองที่ลบข้อมูลสูญหายอยู่ 6.62% แต่ประสิทธิภาพลดลง 0.50% เมื่อเทียบกับแบบจำลองที่สร้างขึ้นจากข้อมูลที่แทนที่ข้อมูลสูญหายเพียงอย่างเดียว เช่นเดียวกับการวัด

ประสิทธิภาพด้วยเครื่องมือ MAE แบบจำลองที่สร้างด้วยข้อมูลที่สกัดคุณลักษณะให้ค่าความคลาดเคลื่อนเฉลี่ยเท่ากับ 4.46 ซึ่งให้ผลลัพธ์ดีกว่าแบบจำลองที่ลบข้อมูลสูญหายอยู่ 8.23% แต่ประสิทธิภาพลดลง 2.68% เมื่อเทียบกับแบบจำลองที่แทนที่ข้อมูลสูญหายเพียงอย่างเดียว โดยตารางค่า RMSE แสดงในตาราง 7 และตาราง MAE แสดงในตาราง 8

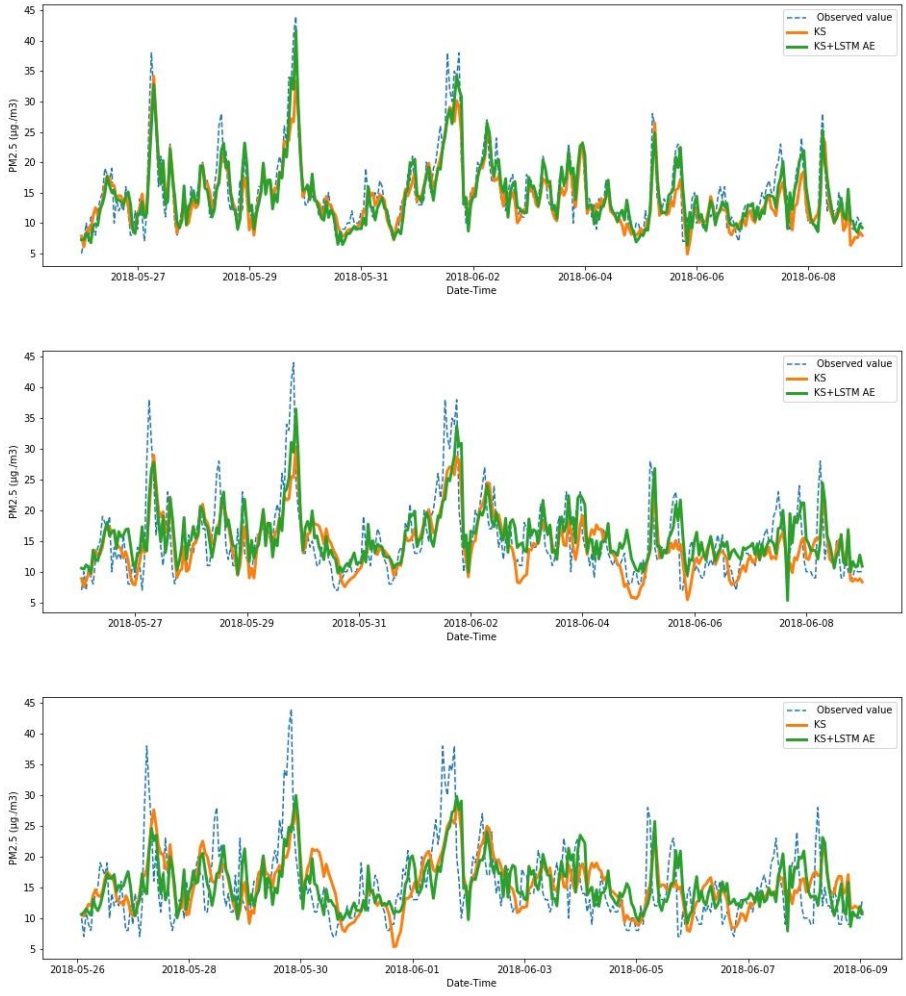
ตาราง 7 ค่า RMSE แสดงประสิทธิภาพแบบจำลอง LSTM ที่แทนที่ข้อมูลด้วย KS และสกัดคุณลักษณะด้วย LSTM Autoencoder

Time Step	ประเภทแบบจำลอง LSTM		
	Baseline	KS	KS+ LSTM AE
t+1	3.11	3.31	3.12
t+2	4.14	4.27	4.20
t+3	4.87	4.78	4.94
t+4	5.36	5.18	5.34
t+5	5.60	5.47	5.67
t+6	5.84	5.73	5.65
t+7	5.98	5.80	5.71
t+8	6.09	5.77	5.82
t+9	6.48	6.02	5.77
t+10	6.32	5.91	5.84
t+11	6.47	5.98	5.81
t+12	6.44	5.86	5.75
t+13	6.69	5.89	5.86
t+14	6.52	5.81	5.84
t+15	6.46	5.83	5.85
t+16	6.42	5.81	6.21
t+17	6.52	5.77	6.01
t+18	6.45	5.92	6.27
t+19	6.71	5.88	6.20
t+20	6.60	5.81	6.04
t+21	6.30	5.83	5.96
t+22	6.59	6.05	5.78
t+23	6.54	5.89	5.75
t+24	6.37	6.16	5.88
ค่าเฉลี่ย	6.04	5.61	5.64
อัตราประสิทธิภาพเพิ่มขึ้น (%)	-	**+7.12%**	**+6.62%**

ตาราง 8 ค่า MAE แสดงประสิทธิภาพแบบจำลอง LSTM ที่แทนที่ข้อมูลด้วย KS และสกัดคุณลักษณะด้วย LSTM Autoencoder

Time Step	ประเภทแบบจำลอง LSTM		
	Baseline	KS	KS + LSTM AE
t+1	2.36	2.54	2.37
t+2	3.16	3.30	3.22
t+3	3.74	3.66	3.86
t+4	4.11	3.94	4.17
t+5	4.36	4.24	4.47
t+6	4.55	4.30	4.36
t+7	4.70	4.36	4.38
t+8	4.80	4.33	4.55
t+9	5.29	4.66	4.41
t+10	5.07	4.48	4.66
t+11	5.24	4.65	4.62
t+12	5.20	4.56	4.51
t+13	5.56	4.62	4.75
t+14	5.25	4.49	4.65
t+15	5.26	4.56	4.67
t+16	5.27	4.45	5.14
t+17	5.37	4.41	4.80
t+18	5.24	4.46	5.19
t+19	5.54	4.51	5.06
t+20	5.46	4.46	4.87
t+21	5.08	4.57	4.85
t+22	5.50	4.84	4.49
t+23	5.44	4.63	4.48
t+24	5.20	4.96	4.56
Average	4.86	4.33	4.46
อัตราประสิทธิภาพเพิ่มขึ้น (%)	-	+10.91%	+8.23%

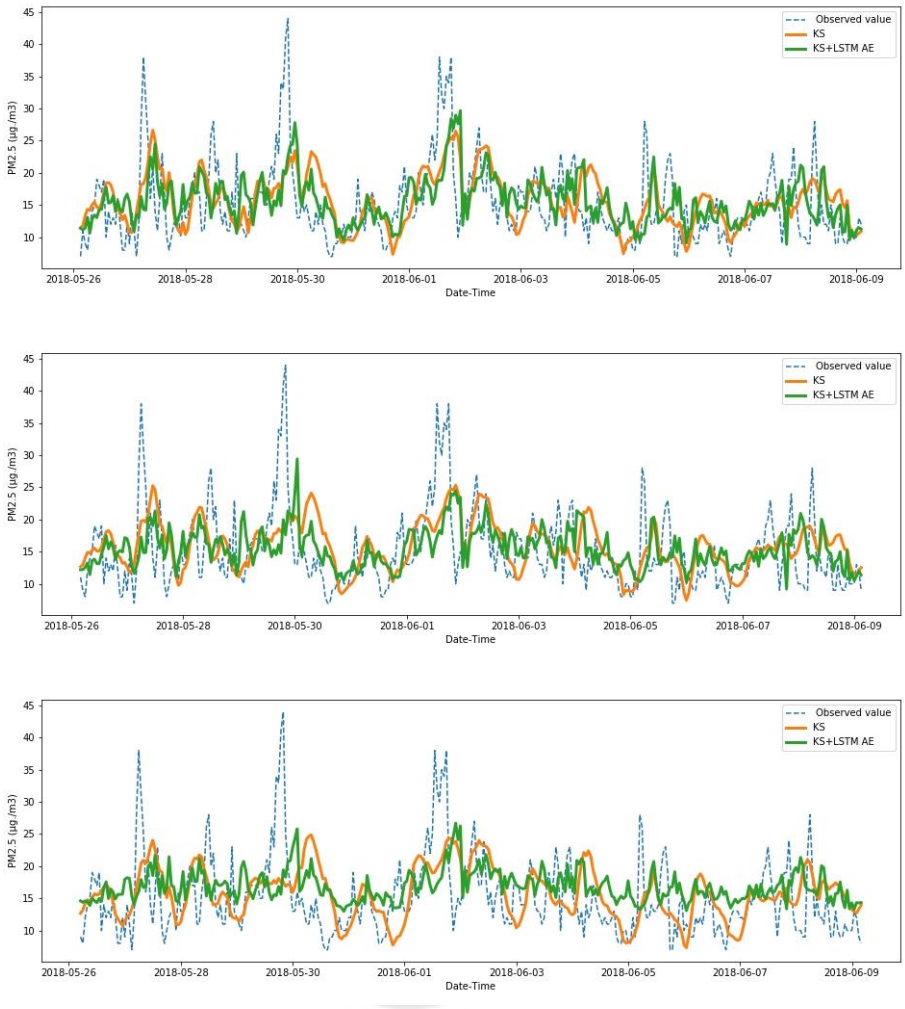
เพื่ออธิบายผลกระทบต่อประสิทธิภาพการพยากรณ์ความหนาแน่นฝุ่นละออง PM2.5 ที่ทำการแทนที่ข้อมูลด้วย KS และทำการสกัดคุณลักษณะด้วย LSTM AE ผู้วิจัยทำการนำค่าความหนาแน่นที่ได้จากการพยากรณ์ล่วงหน้าตั้งแต่ชั่วโมงที่ 1 จนถึงชั่วโมง 24 ของแบบจำลอง KS และ KS ร่วมกับ LSTM AE ที่ให้ประสิทธิภาพ RMSE ดีที่สุด นำมาพล็อตกราฟเปรียบเทียบ เพื่อให้เห็นผลกระทบที่ชัดเจนผู้วิจัยได้เลือกช่วงเวลา 336 ชั่วโมงแรกของ Test Set สามารถแสดงได้ดังนี้



ภาพประกอบ 17 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 1 ถึงชั่วโมงที่ 3

ภาพประกอบ 17 แสดงตัวอย่างแสดงผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 1 ถึงชั่วโมงที่ 3 โดยผลลัพธ์ของแบบจำลองที่แทนที่ข้อมูลสูญหายร่วมกับการสกัดคุณลักษณะมีประสิทธิภาพดีไม่ต่างกันมากกว่าแบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวที่ชั่วโมงที่ 1 และชั่วโมงที่ 2 แต่ใน

ชั่วโมงที่ 3 แบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวให้ประสิทธิภาพในการพยากรณ์มากกว่าแบบจำลองที่ทำการแทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะ

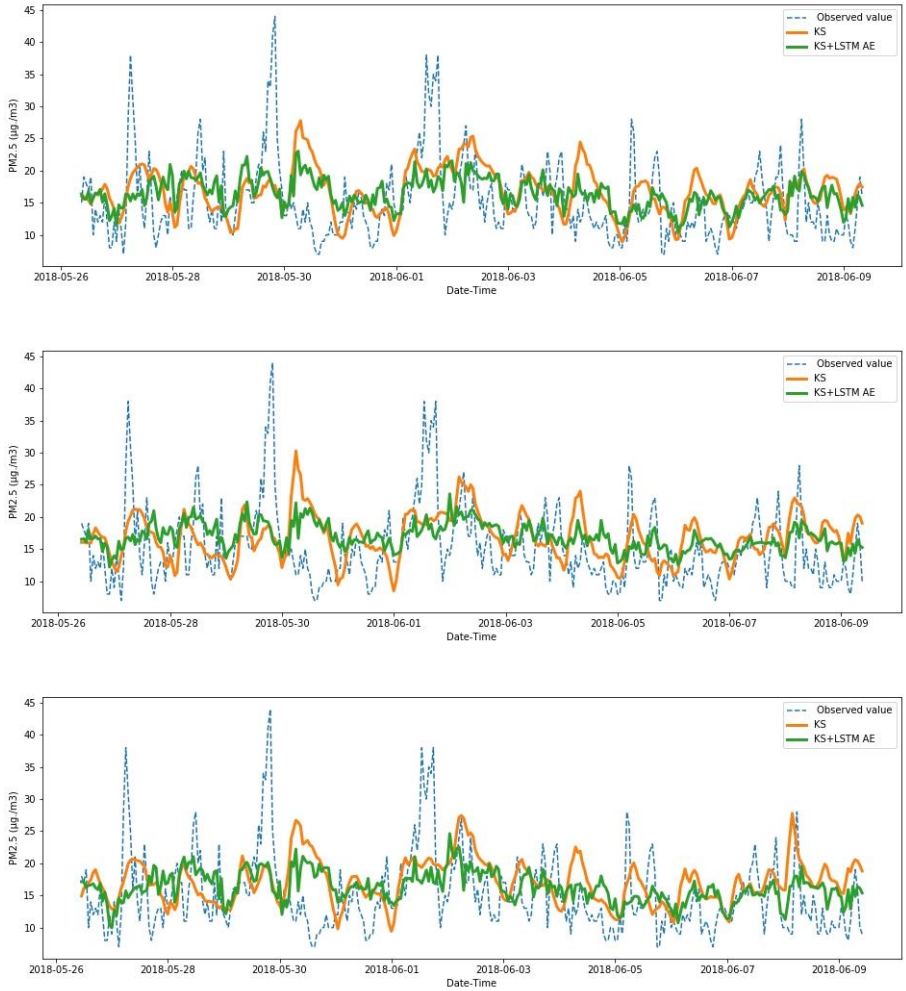


ภาพประกอบ 18 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 4 ถึงชั่วโมงที่ 6

ภาพประกอบ 18 แสดงตัวอย่างผลลัพธ์การพยากรณ์ชั่วโมงที่ 4 ถึงชั่วโมงที่ 6 โดยชั่วโมงที่ 4 และชั่วโมงที่ 5 ผลลัพธ์ของแบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวให้ประสิทธิภาพที่ดีกว่าแบบจำลองที่ทำการสกัดคุณลักษณะ แต่ในชั่วโมงที่ 6 แบบจำลองที่แทนที่แทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะให้ประสิทธิภาพในการพยากรณ์ PM2.5 ดีกว่าแบบจำลองที่ทำการแทนที่ข้อมูลเพียงอย่างเดียวอยู่เล็กน้อยหากวัดค่า RMSE แต่ไม่ได้ดีกว่าหากวัดด้วย MAE

ภาพประกอบ 19 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 7 ถึงชั่วโมงที่ 9

ภาพประกอบ 19 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 7 และชั่วโมงที่ 8 พบว่าแบบจำลองลองทั้งสองแบบให้ประสิทธิภาพเมื่อประเมินด้วย RMSE และ MAE ไม่ต่างกันอย่างมีนัยสำคัญ แต่ในชั่วโมงที่ 9 พบว่าแบบจำลองที่ทำการสกัดคุณลักษณะให้ประสิทธิภาพดีกว่าแบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวเมื่อประเมินด้วย RMSE และ MAE



ภาพประกอบ 20 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 10 ถึงชั่วโมงที่ 12

ภาพประกอบ 20 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 10 ถึงชั่วโมงที่ 12 ในชั่วโมงที่ 10 ผลการประเมิน RMSE และ MAE ของทั้งสองแบบจำลองให้ผลลัพธ์คนละทิศทาง จึงเป็นไปได้ยากที่จะบอกว่าวิธีใดดีกว่ากัน แต่ในชั่วโมงที่ 11 และ 12 แบบจำลองที่ผ่านการสกัดคุณลักษณะมีประสิทธิภาพดีกว่าแบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวอยู่เล็กน้อย

ภาพประกอบ 21 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 13 ถึงชั่วโมงที่ 15

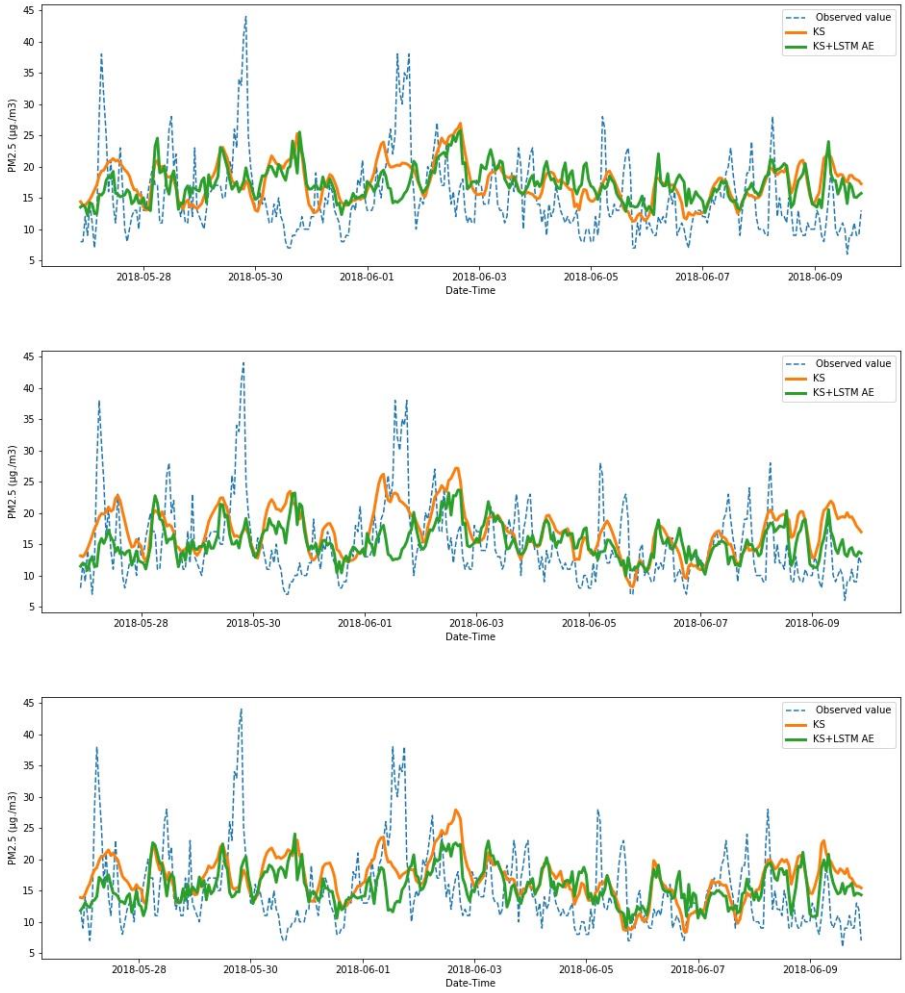
ภาพประกอบ 21 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 13 ถึงชั่วโมงที่ 15 ในชั่วโมงที่ 13 14 และ 15 พบว่าค่า RMSE และค่า MAE ที่ประเมินแบบจำลองทั้งสองแบบมีค่าใกล้เคียงกันมากแสดงถึงประสิทธิภาพในการทำงานที่ใกล้เคียงกัน

ภาพประกอบ 22 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 16 ถึงชั่วโมงที่ 18

ภาพประกอบ 22 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 16 ถึงชั่วโมงที่ 18 ในการพยากรณ์ผลลัพธ์ในช่วงเวลานี้พบว่า แบบจำลองที่ทำการแทนที่ข้อมูลพียงอย่างเดียวมีแนวโน้มผลลัพธ์การพยากรณ์ PM2.5 ดีกว่าแบบจำลองที่ทำการแทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะ

ภาพประกอบ 23 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 19 ถึงชั่วโมงที่ 21

ภาพประกอบ 23 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 19 ถึงชั่วโมงที่ 21 พบว่ายิ่งพยากรณ์ล่วงหน้าไกลมากขึ้นประสิทธิภาพการพยากรณ์ของแบบจำลอง LSTM ยิ่งลดลง ส่วนของผลลัพธ์การพยากรณ์พบว่าแบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวให้ประสิทธิภาพในการพยากรณ์ที่ดีกว่าแบบจำลองที่ทำการแทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะ ทั้งชั่วโมงที่ 19 20 และ 21



ภาพประกอบ 24 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย KS เทียบกับ KS + LSTM AE ชั่วโมงที่ 22 ถึงชั่วโมงที่ 24

ภาพประกอบ 24 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 22 ถึงชั่วโมงที่ 24 ในชั่วโมงที่ 22 23 และ 24 พบว่าแบบจำลองที่ทำการสกัดคุณลักษณะให้ประสิทธิภาพในการพยากรณ์มากกว่าแบบจำลองที่ทำการแทนที่ข้อมูลเพียงอย่างเดียว

4.2 ผลลัพธ์การพยากรณ์โดยการใช้เทคนิคแทนที่ข้อมูลด้วย LWMA และ การพยากรณ์โดยใช้เทคนิคแทนที่ข้อมูลด้วย LWMA ร่วมกับการสกัดคุณลักษณะด้วย LSTM Autoencoder

เมื่อแทนที่ข้อมูลสูญหายด้วยเทคนิค LWMA แล้วนำไปสร้างแบบจำลอง LSTM พบว่าแบบจำลอง LSTM ที่มีประสิทธิที่ภาพที่สุดเมื่อวัดด้วย RMSE ได้ค่าความคลาดเคลื่อนเฉลี่ยอยู่ที่ 5.56 ประสิทธิภาพเพิ่มขึ้น 7.95% เมื่อเทียบจากแบบจำลองที่สร้างจากชุดข้อมูลที่ลบแถวข้อมูลสูญหาย ส่วนเมื่อวัดด้วยเครื่องมือวัด MAE แบบจำลอง LSTM ให้ค่าคลาดเคลื่อนเฉลี่ยน้อยที่สุดอยู่ที่ 4.33 มีประสิทธิภาพดีขึ้นจากแบบจำลองที่ลบข้อมูลสูญหาย 10.91% จากนั้นทำการสกัดคุณลักษณะจากชุดข้อมูลที่แทนที่ข้อมูลสูญหายด้วยเทคนิค LWMA แล้วสร้างแบบจำลอง LSTM พบว่าเมื่อวัดประสิทธิภาพด้วย RMSE พบว่าประสิทธิภาพเพิ่มจากแบบจำลองที่ลบข้อมูลสูญหาย 7.45% แต่ประสิทธิภาพลดลองจากแบบจำลองที่แทนที่ข้อมูลสูญหายเพียงอย่างเดียวอยู่ที่ 0.50% ส่วนเมื่อวัดด้วย MAE แบบจำลอง LSTM ที่สร้างจากชุดข้อมูลที่ทำการสกัดคุณลักษณะนั้นให้ค่าความคลาดเคลื่อนที่ 4.37 ดีกว่าแบบจำลองที่ลบข้อมูลสูญหายอยู่ 10.08% แต่ประสิทธิภาพลดลงไป 0.83% เมื่อเทียบกับแบบจำลองที่ทำการแทนที่ข้อมูลด้วย LWMA เพียงอย่างเดียว โดยตารางค่า RMSE แสดงในตาราง 9 และตาราง MAE แสดงที่ตาราง 10

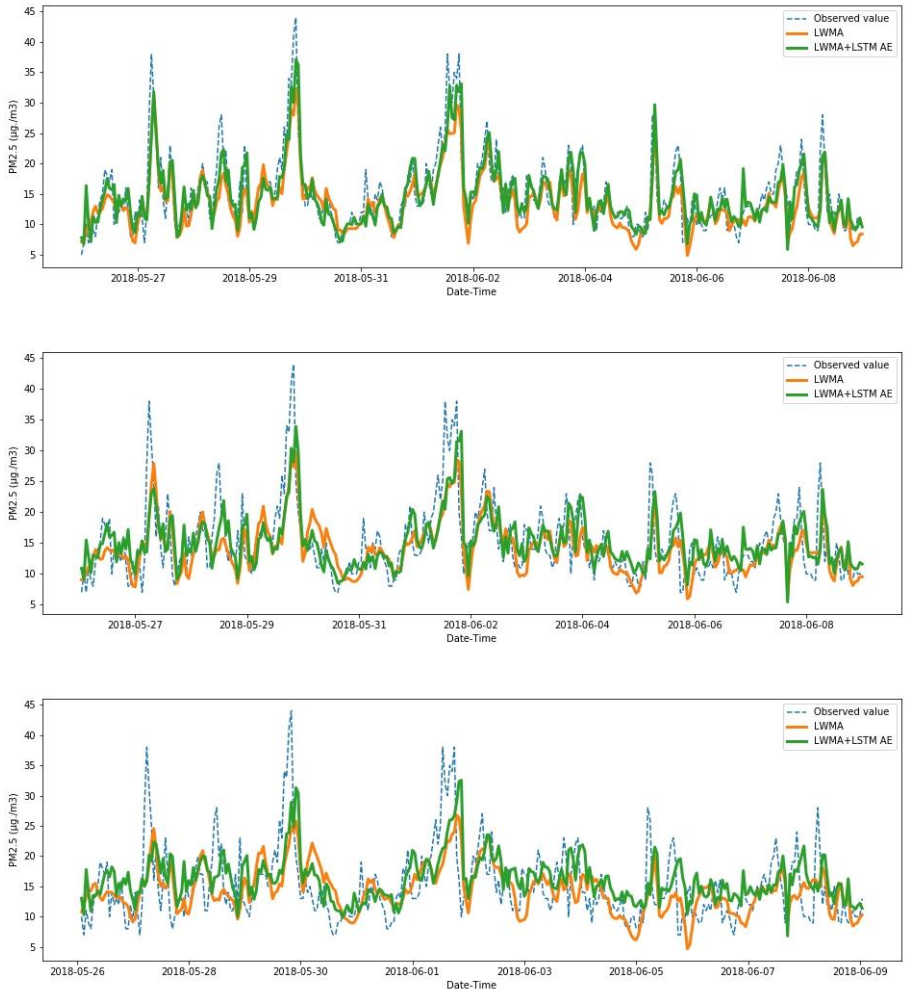
ตาราง 9 ค่า RMSE แสดงประสิทธิภาพแบบจำลอง LSTM ที่แทนที่ข้อมูลด้วย LWMA และสกัดคุณลักษณะด้วย LSTM Autoencoder

Time Step	ประเภทแบบจำลอง LSTM		
	Baseline	LWMA	LWMA + LSTM AE
t+1	3.11	3.13	3.35
t+2	4.14	4.26	4.26
t+3	4.87	4.67	5.00
t+4	5.36	4.91	5.32
t+5	5.60	5.15	5.53
t+6	5.84	5.44	5.51
t+7	5.98	5.57	5.80
t+8	6.09	5.75	5.88
t+9	6.48	5.75	5.74
t+10	6.32	5.84	5.73
t+11	6.47	5.70	5.81
t+12	6.44	5.91	5.69
t+13	6.69	5.92	5.65
t+14	6.52	5.82	5.63
t+15	6.46	6.00	5.75
t+16	6.42	5.65	5.81
t+17	6.52	5.80	5.96
t+18	6.45	5.75	5.89
t+19	6.71	5.99	5.94
t+20	6.60	6.05	5.93
t+21	6.30	5.95	5.96
t+22	6.59	5.99	5.98
t+23	6.54	6.32	5.92
t+24	6.37	6.00	6.03
Average	6.04	5.56	5.59
อัตราประสิทธิภาพเพิ่มขึ้น (%)	-	**+7.95%**	**+7.45%**

ตาราง 10 ค่า MAE แสดงประสิทธิภาพแบบจำลอง LSTM ที่แทนที่ข้อมูลด้วย LWMA และสกัดคุณลักษณะด้วย LSTM Autoencoder

Time Step	ประเภทแบบจำลอง LSTM		
	Baseline	LWMA	LWMA + LSTM AE
t+1	2.36	2.69	2.45
t+2	3.16	3.17	3.27
t+3	3.74	3.71	3.92
t+4	4.11	3.94	4.22
t+5	4.36	4.06	4.27
t+6	4.55	4.21	4.46
t+7	4.70	4.25	4.60
t+8	4.80	4.31	4.54
t+9	5.29	4.30	4.37
t+10	5.07	4.32	4.58
t+11	5.24	4.40	4.41
t+12	5.20	4.48	4.37
t+13	5.56	4.52	4.52
t+14	5.25	4.64	4.55
t+15	5.26	4.69	4.31
t+16	5.27	4.71	4.49
t+17	5.37	4.43	4.75
t+18	5.24	4.83	4.62
t+19	5.54	4.58	4.78
t+20	5.46	4.54	4.60
t+21	5.08	4.86	4.64
t+22	5.50	4.93	4.79
t+23	5.44	4.67	4.66
t+24	5.20	4.73	4.77
Average	4.86	4.33	4.37
อัตราประสิทธิภาพเพิ่มขึ้น (%)	-	**+10.91%**	**+10.08%**

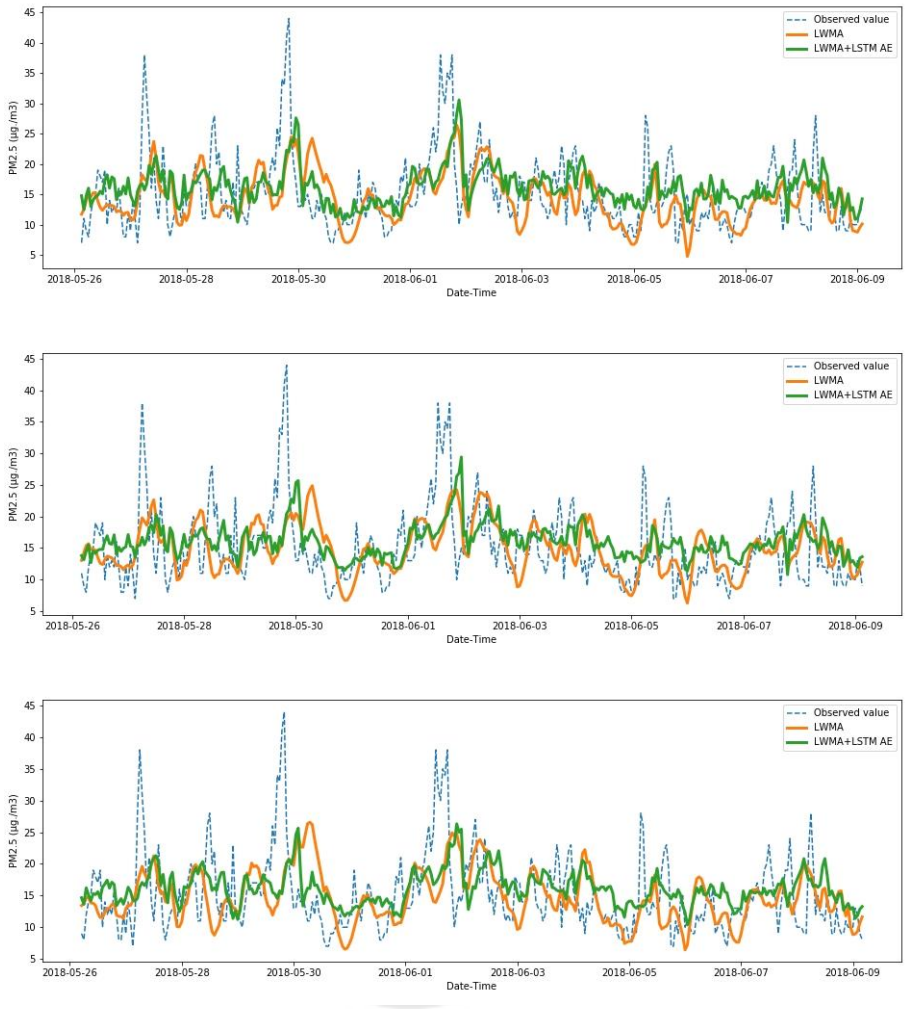
เพื่ออธิบายผลกระทบต่อประสิทธิภาพการพยากรณ์ความหนาแน่นฝุ่นละออง PM2.5 ทำที่ทำการแทนที่ข้อมูลด้วย LWMA และทำการสกัดคุณลักษณะด้วย LSTM Autoencoder ผู้วิจัยได้นำค่าความหนาแน่นที่ได้จากการพยากรณ์รายชั่วโมงตั้งแต่ชั่วโมงที่ 1 จนถึงชั่วโมงที่ 24 ของแบบจำลองที่แทนที่ข้อมูลด้วย LWMA และแบบจำลองที่แทนที่ข้อมูลด้วย LWMA ร่วมกับ LSTM AE ที่ให้ประสิทธิภาพ RMSE ดีที่สุด นำมาพล๊อตกราฟเปรียบเทียบ และเพื่อให้เห็นผลกระทบที่ชัดเจนผู้วิจัยได้เลือกช่วงเวลา 336 ชั่วโมงแรกของ Test Set แบ่งช่วงการอธิบายออกเป็นช่วง ดังนี้



ภาพประกอบ 25 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย LWMA เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 1 ถึงชั่วโมงที่ 3

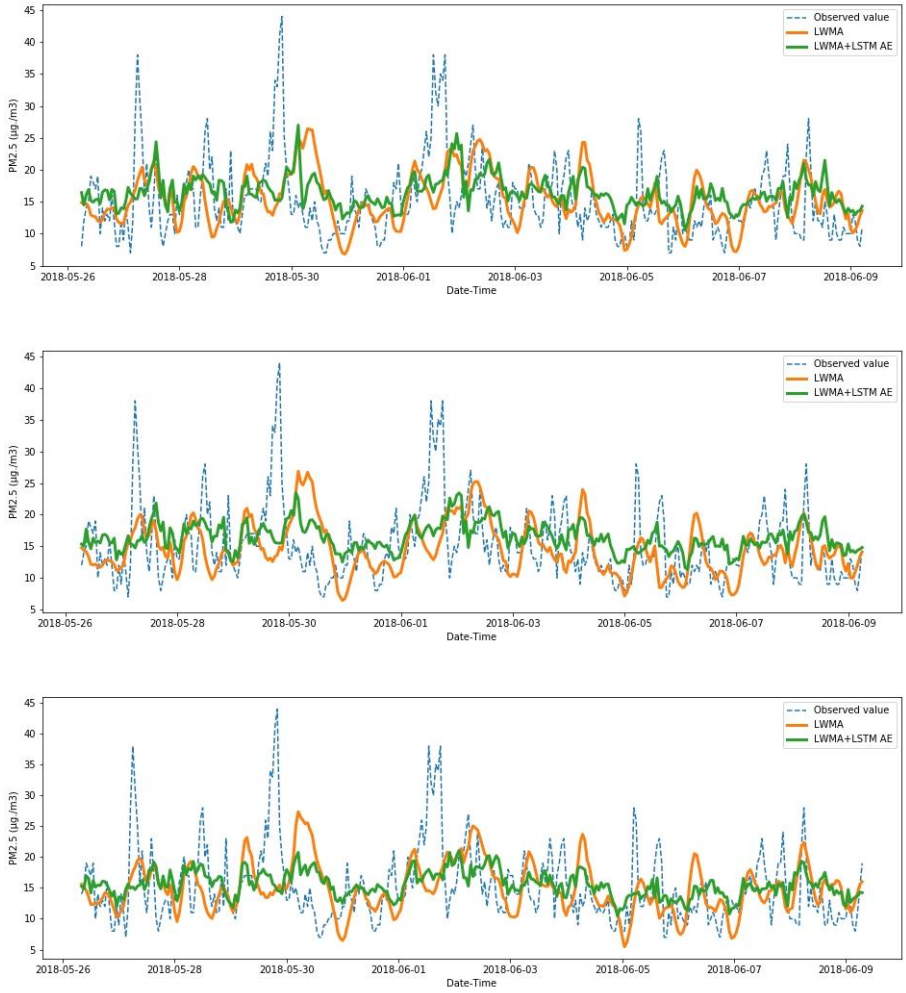
ภาพประกอบ 25 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 1 ถึงชั่วโมงที่ 3 โดยชั่วโมงที่ 1 และ 2 แบบจำลองทั้งสองแบบมีประสิทธิภาพในการพยากรณ์ PM2.5 ใกล้เคียงกัน

แต่ในชั่วโมงที่ 3 แบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวมีประสิทธิภาพในการพยากรณ์ PM2.5 ดีกว่าแบบจำลองที่แทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะ



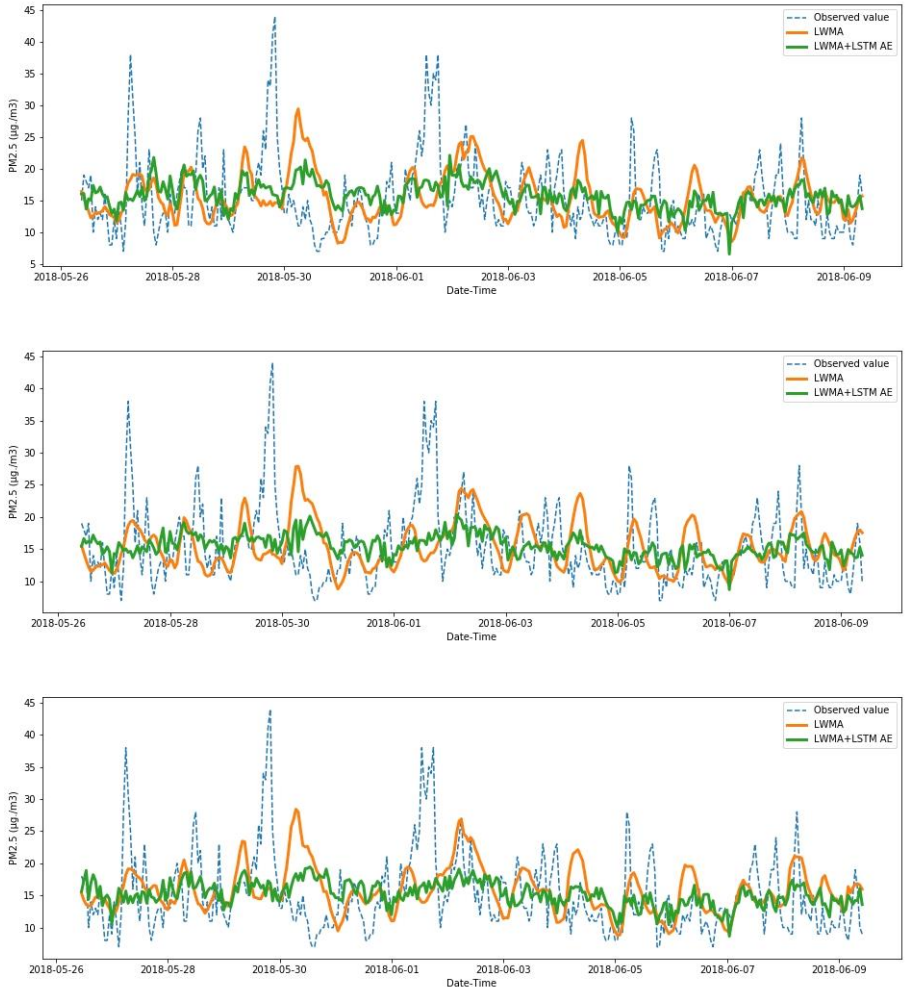
ภาพประกอบ 26 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย LWMA เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 4 ถึงชั่วโมงที่ 6

ภาพประกอบ 26 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 4 ถึงชั่วโมงที่ 6 พบว่าในชั่วโมงที่ 4 5 และ 6 แบบจำลองที่ทำการแทนที่ข้อมูลเพียงอย่างเดียวมีประสิทธิภาพดีกว่าแบบจำลองที่แทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะอย่างชัดเจนเมื่อประเมินค่าด้วย RMSE และ MAE



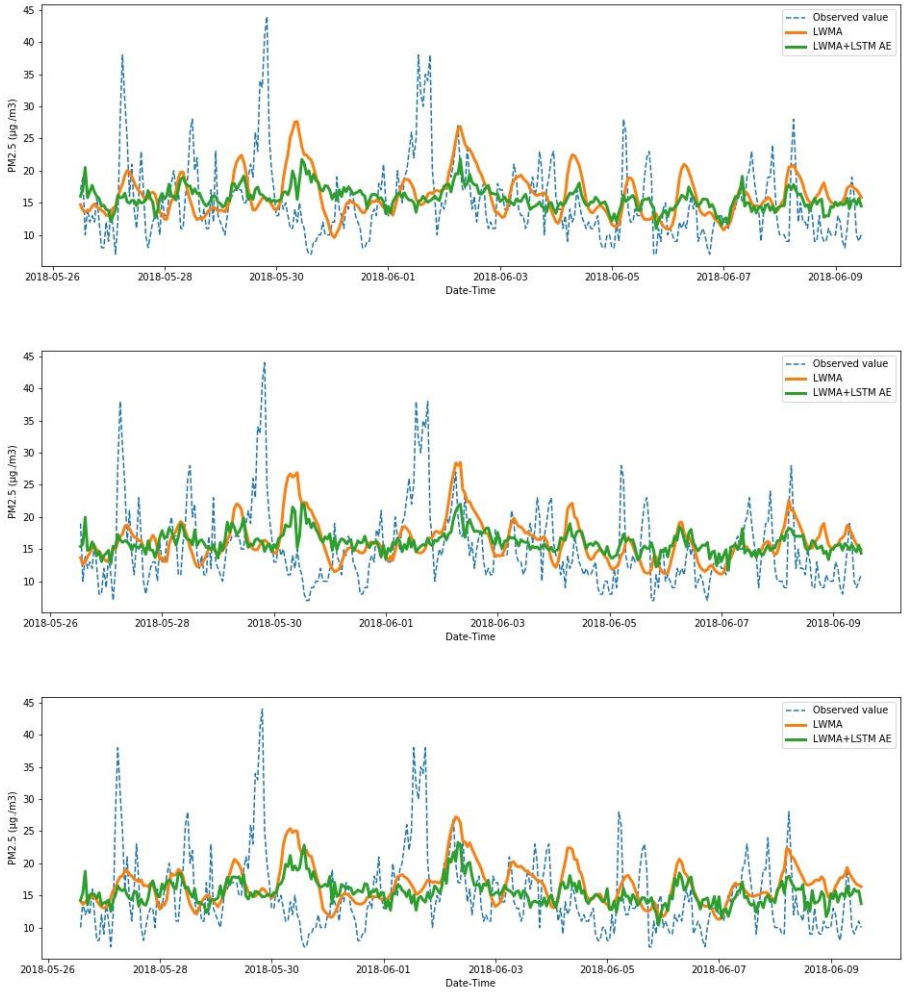
ภาพประกอบ 27 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 7 ถึงชั่วโมงที่ 9

ภาพประกอบ 27 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 7 ถึงชั่วโมงที่ 9 พบว่า การพยากรณ์ผลลัพธ์ในช่วงเวลานี้ แบบจำลองที่แทนที่ข้อมูลเพียงอย่างเดียวมีประสิทธิภาพดีกว่าเล็กน้อยเมื่อเปรียบเทียบกับแบบจำลองที่ทำการสกัดคุณลักษณะ



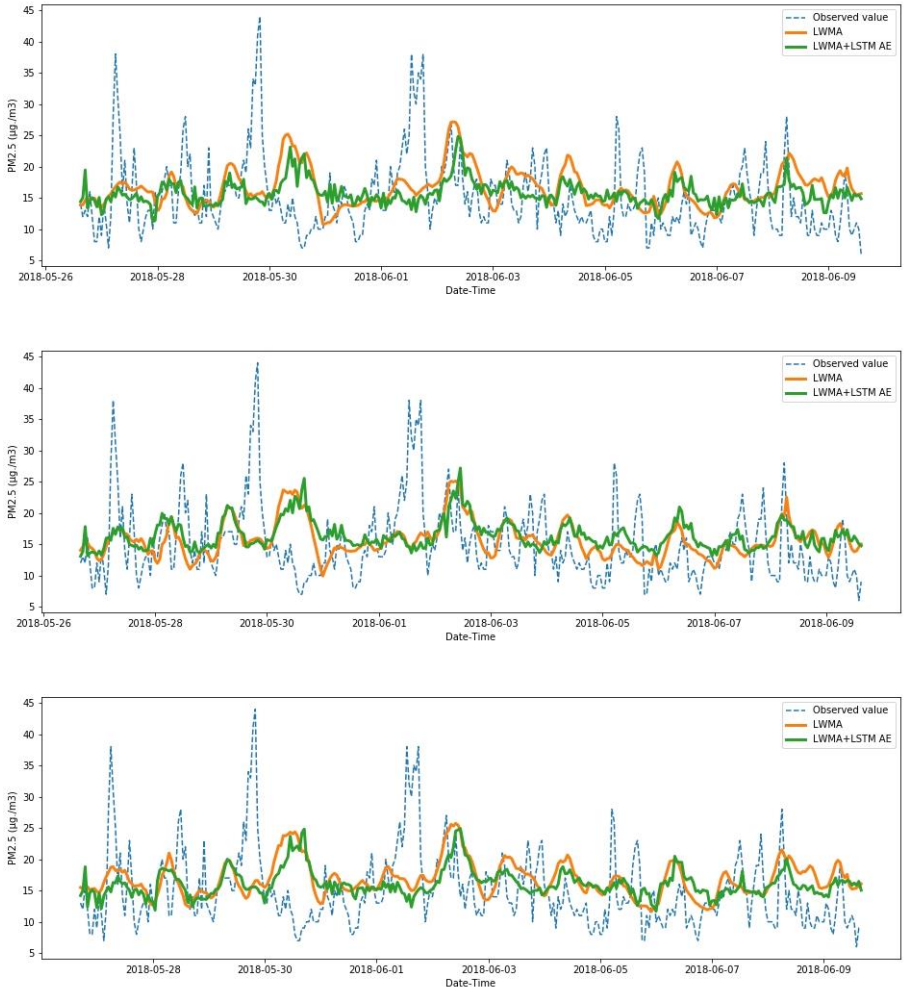
ภาพประกอบ 28 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 10 ถึงชั่วโมงที่ 12

ภาพประกอบ 28 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 10 ถึงชั่วโมงที่ 12 พบว่าในชั่วโมงที่ 10 และ 11 แบบจำลองทั้ง 2 ประเภทมีประสิทธิภาพไม่แตกต่างกันมาก แต่ในชั่วโมงที่ 12 แบบจำลองที่ทำการสกัดคุณลักษณะมีประสิทธิภาพการพยากรณ์ที่ดีกว่าแบบจำลองที่แทนที่ข้อมูลอยู่เล็กน้อยเมื่อประเมินประสิทธิภาพด้วย RMSE และ MAE



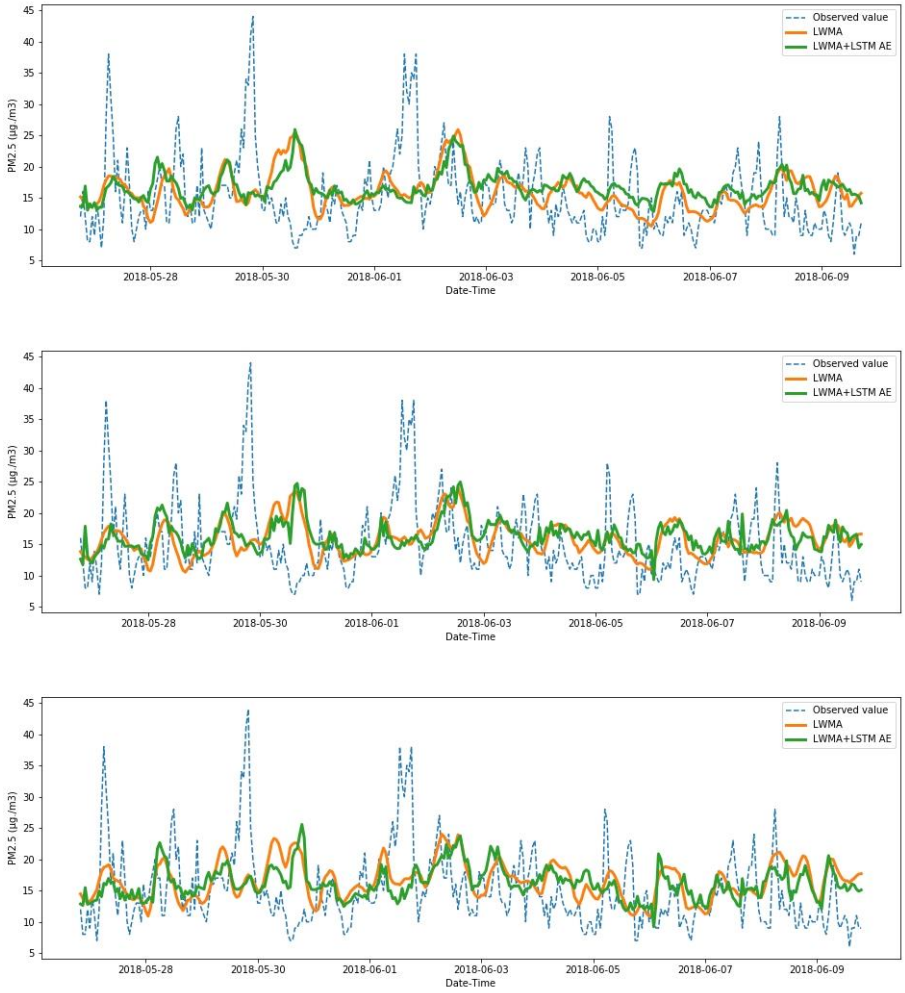
ภาพประกอบ 29 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย LWMA เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 13 ถึงชั่วโมงที่ 15

ภาพประกอบ 29 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 13 ถึงชั่วโมงที่ 15 พบว่าชั่วโมงที่ 14 และ 15 แบบจำลองที่สกัดคุณลักษณะมีประสิทธิภาพในการพยากรณ์ PM2.5 ล่วงหน้าดีกว่าแบบจำลองที่ทำการแทนที่ข้อมูลเพียงอย่างเดียว ส่วนชั่วโมงที่ 13 แบบจำลองที่ทำการสกัดคุณลักษณะให้ผลลัพธ์การพยากรณ์ดีกว่าเมื่อประเมินด้วย RMSE แต่ค่า MAE ของแบบจำลองทั้งสองแบบนั้นไม่ต่างกัน



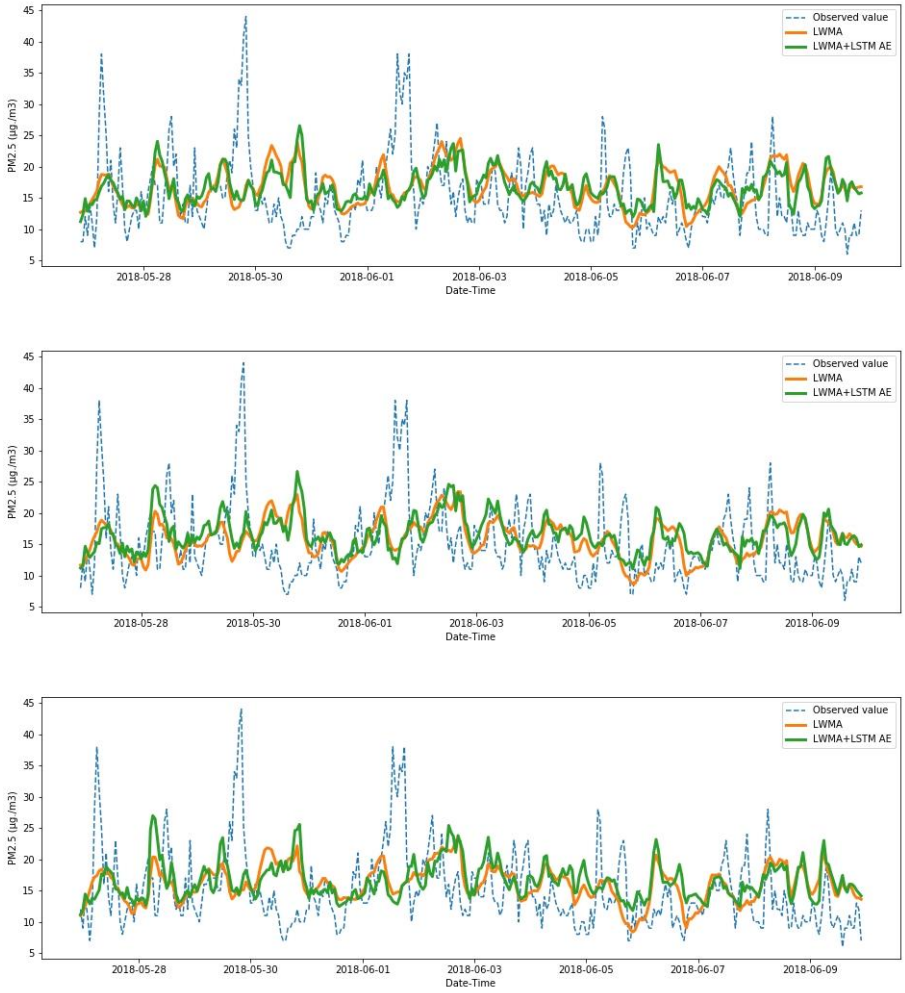
ภาพประกอบ 30 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย LWMA เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 16 ถึงชั่วโมงที่ 18

ภาพประกอบ 30 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 16 ถึงชั่วโมงที่ 18 พบว่าในชั่วโมงที่ 16 และ 18 ผลลัพธ์การพยากรณ์ PM2.5 ล่วงหน้าของทั้งสองแบบจำลองนั้นขัดแย้งกันเมื่อประเมินด้วย RMSE และ MAE ทำให้สรุปไม่ได้ว่าวิธีใดดีกว่ากัน และในชั่วโมงที่ 17 พบว่าแบบจำลองที่ทำการแทนที่ข้อมูลเพียงอย่างเดียวมีประสิทธิภาพในการพยากรณ์ PM2.5 ดีกว่าแบบจำลองที่ทำการแทนที่ข้อมูลร่วมกับการสกัดคุณลักษณะ



ภาพประกอบ 31 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย LWMA เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 19 ถึงชั่วโมงที่ 21

ภาพประกอบ 31 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 19 ถึงชั่วโมงที่ 21 พบว่าแบบจำลองทั้ง 2 มีประสิทธิภาพในการตรวจจับรูปแบบข้อมูลใกล้เคียงกัน แต่เมื่อประเมินประสิทธิภาพด้วย RMSE และ MAE จึงไม่สามารถตัดสินได้ว่าแบบจำลองประเภทใดดีกว่ากัน



ภาพประกอบ 32 ตัวอย่างแสดงผลลัพธ์การพยากรณ์ PM2.5 ที่แทนที่ข้อมูลด้วย LWMA เทียบกับ LSTM + LSTM AE ชั่วโมงที่ 22 ถึงชั่วโมงที่ 24

ภาพประกอบ 32 แสดงตัวอย่างผลลัพธ์การพยากรณ์ล่วงหน้าชั่วโมงที่ 22 ถึงชั่วโมงที่ 24 ถึงแม้ว่าความสามารถของแบบจำลองทั้ง 2 ประเภทจะลดลงเป็นอย่างมากเมื่อเทียบกับการพยากรณ์ฝุ่น PM2.5 ในชั่วโมงแรก แต่เมื่อประเมินประสิทธิภาพด้วย RMSE และ MAE พบว่าการพยากรณ์ PM2.5 ในชั่วโมงที่ 22 และ 23 แบบจำลองที่ทำการสกัดคุณลักษณะมีประสิทธิภาพดีกว่าแบบจำลองที่ทำการแทนที่ข้อมูลเพียงอย่างเดียว

จากการเพิ่มวิธีการสกัดคุณลักษณะ (Feature Extraction) กับการแทนที่ข้อมูลด้วย KS และ LWMA พบว่าโดยรวมประสิทธิภาพของแบบจำลอง LSTM ที่สร้างขึ้นจากการสกัดคุณลักษณะนั้นมีประสิทธิภาพลดลงเมื่อเทียบกับแบบจำลอง LSTM ที่สร้างจากการแทนที่ข้อมูลเพียงอย่างเดียว อาจเป็นเพราะการลด Noise ด้วยการสกัดคุณลักษณะเป็นการลดโอกาสที่แบบจำลอง LSTM จะได้เรียนรู้รูปแบบอนุกรมเวลาของ PM2.5 แต่เมื่อพิจารณาผลลัพธ์การพยากรณ์โดยเพิ่มการสกัดคุณลักษณะในการสร้างแบบจำลอง LSTM พบว่าแบบจำลองที่แทนที่ข้อมูลด้วย KS ร่วมกับการสกัดคุณลักษณะมีประสิทธิภาพดีกว่าแบบจำลองที่แทนที่ด้วย KS เพียงอย่างเดียว ในชั่วโมงที่ 1 2 6 9 11 12 22 23 และ 24 เช่นเดียวกับแบบจำลองที่แทนที่ข้อมูลด้วย LMWA ร่วมกับการสกัดคุณลักษณะมีประสิทธิภาพดีกว่าแบบจำลองที่แทนที่ด้วย LWMA เพียงอย่างเดียวในชั่วโมงที่ 12 13 15 22 และ 23 แสดงให้เห็นว่าการสกัดคุณลักษณะด้วย LSTM AE ช่วยเพิ่มประสิทธิภาพในการพยากรณ์ของแบบจำลอง LSTM ในระยะที่สั้นและในระยะเวลาที่ยาวในบางช่วงเวลา

บทที่ 5
สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในงานวิจัยนี้ได้ศึกษาเทคนิคการแทนที่ข้อมูลสูญหายร่วมกับการสกัดคุณลักษณะเพื่อเพิ่มประสิทธิภาพในการทำนายฝุ่นละออง PM2.5 ในอีก 24 ชั่วโมงข้างหน้า ซึ่งใช้ข้อมูลจากบริเวณพื้นที่สถานีตำรวจนครบาลโชคชัย ในชุดข้อมูลประกอบด้วยข้อมูลฝุ่น PM2.5 และข้อมูลทางอุตุนิยมวิทยา ผู้วิจัยได้วัดประสิทธิภาพแบบจำลองที่ได้ถูกสร้างขึ้นเพื่อนำมาเปรียบเทียบและสรุปผลการวิจัย โดยสามารถแบ่งหัวข้อได้ดังนี้

1. สรุปผลการวิจัย
2. อภิปรายผลการวิจัย
3. ข้อเสนอแนะ

สรุปผลการวิจัย

ในปัจจุบันปัญหาเรื่องฝุ่นละออง PM2.5 กลายเป็นปัญหาระดับประเทศของประเทศไทย เนื่องจากระดับความหนาแน่นของฝุ่น PM2.5 ในบริเวณเมืองใหญ่ที่สูงเกินมาตรฐานของทางรัฐบาล ส่งผลให้เกิดผลกระทบต่อการใช้ชีวิตและสุขภาพของผู้คนเป็นจำนวนมาก ซึ่งการสร้างเครื่องมือที่มีความสามารถในการพยากรณ์ค่าความหนาแน่นของฝุ่น PM2.5 ล่วงหน้าที่มีความแม่นยำจึงเป็นตัวช่วยให้ผู้คนสามารถวางแผน ป้องกัน และเตรียมการรับมือปริมาณฝุ่น PM2.5 ที่ความหนาแน่นอาจเพิ่มขึ้นจนมีผลกระทบกับคุณภาพชีวิตของผู้คนได้อย่างทันที

ในการวิจัยนี้ ผู้วิจัยได้เลือกเทคนิคการแทนที่ข้อมูลสูญหาย ได้แก่ Kalman Smoothing และ Linearly Weighted Moving Average ในส่วนของการสกัดคุณลักษณะได้ใช้ LSTM Autoencoder เป็นเครื่องมือและใช้ Long Short-Term Memory ในการสร้างแบบจำลอง โดยการทดลองจะทำการแทนที่ข้อมูลสูญหายแล้วสร้างแบบจำลอง และการแทนที่ข้อมูลสูญหายร่วมกับการสกัดคุณลักษณะของข้อมูล PM2.5 ย้อนหลัง 168 ชั่วโมง และประเมินประสิทธิภาพแบบจำลองที่สร้างขึ้นด้วย Root Mean Square Error และ Mean Absolute Error

ผลกระทบของการแทนที่ข้อมูลที่สูญหาย (Data Imputation) นั้นควรทำเป็นอย่างยิ่งในขั้นตอนการเตรียมข้อมูล (Data Pre-Processing) เพราะเมื่อแทนที่ข้อมูลด้วย KS และ LWMA พบว่าแบบจำลองให้ประสิทธิภาพเพิ่มขึ้นจากแบบจำลองที่สร้างด้วยข้อมูลที่ทำการลบแถวข้อมูลสูญหายทิ้งอยู่ประมาณ 7%-8% เมื่อเปรียบเทียบด้วยค่าความคลาดเคลื่อนด้วย RMSE และเมื่อ

ทำการสกัดคุณลักษณะ (Feature Extraction) PM2.5 ด้วย LSTM AE พบว่า LSTM ช่วยเพิ่มความสามารถการพยากรณ์ผลลัพธ์ในชั่วโมงแรกๆและตั้งแต่ชั่วโมงที่ 22 เป็นต้นไป

อภิปรายผลการวิจัย

จากผลการทดลองในบทที่ 4 พบว่าการแทนที่ข้อมูลสูญหายสามารถลดข้อมูลรบกวนและช่วยให้ LSTM ที่เป็นแบบจำลองโครงข่ายประสาทเทียมแบบเชิงลึก (Deep Neural Network) สามารถเรียนรู้ และพยากรณ์ค่าความเข้มข้น PM2.5 ได้แม่นยำขึ้นมากขึ้น แต่การสกัดคุณลักษณะ PM2.5 ด้วย LSTM AE ช่วยเพิ่มประสิทธิภาพในการพยากรณ์ในบางช่วงเวลาเท่านั้น โดยช่วงเวลาที่แบบจำลองสามารถพยากรณ์ผลลัพธ์ออกมาได้ดีคือช่วงชั่วโมงท้ายของการพยากรณ์ ถึงการแทนที่ข้อมูลด้วย KS และ LWMA และการสกัดคุณลักษณะด้วย LSTM AE นั้นมีประสิทธิภาพได้ดีกว่าแบบจำลอง LSTM ที่ทำการลบข้อมูลที่สูญหาย [1] แต่ประสิทธิภาพของแบบจำลอง LSTM นั้นยังให้ผลลัพธ์ที่ออกมายังดีไม่สุดโดยเฉพาะการพยากรณ์ค่า PM2.5 ในช่วงเวลาที่ไกลออกไป (Long-Term) ถึงแม้ว่าการแทนที่ข้อมูลด้วย KS และ LWMA สามารถเพิ่มประสิทธิภาพให้แบบจำลอง LSTM ได้เป็นอย่างดีแต่เมื่อเจอเหตุการณ์ข้อมูลนั้นสูญหายต่อเนื่องเป็นจำนวนมากค่าที่แทนที่ของทั้งสองเทคนิคได้ทำการแทนที่ค่าสูญหายนั้นเป็นค่าคงที่ซึ่งทำให้เกิด Noise ในชุดข้อมูลได้ การศึกษาเทคนิคการแทนที่ข้อมูลที่สามารถจัดการกับข้อมูลที่สูญหายเป็นจำนวนมากได้ดีอาจจะช่วยเพิ่มประสิทธิภาพในการพยากรณ์ให้เพิ่มขึ้น เช่นเดียวกันกับการสกัดคุณลักษณะ LSTM Autoencoder ให้ประสิทธิภาพที่ดีในการพยากรณ์ผลลัพธ์ในบางช่วงเวลา ถ้าสามารถหาวิธีการสกัดคุณลักษณะอื่นที่ส่งผลดีกับผลลัพธ์การพยากรณ์ทุกช่วงเวลาอาจส่งผลให้แบบจำลอง LSTM มีประสิทธิภาพเพิ่มขึ้นได้เนื่องจาก LSTM มีลักษณะการทำงานเป็นกล่องดำส่งผลให้ Interpretability ของคุณลักษณะต่ำ ทำให้ไม่สามารถอธิบายความสำคัญของคุณลักษณะ (Feature Importance) ไปมากกว่านี้

ข้อเสนอแนะ

1. ในงานวิจัยนี้ใช้ข้อมูลทำการวิจัยเพียงแห่งเดียว หากใช้ข้อมูลจากสถานีอื่นมาประกอบอาจช่วยให้แบบ LSTM ที่เป็นแบบจำลองแบบโครงข่ายประสาทเทียมเชิงลึกสามารถหารูปแบบของค่าความหนาแน่นของฝุ่น PM2.5 ได้แม่นขึ้น

2. ในงานวิจัยนี้ได้ออกแบบโครงข่ายประสาทเทียมของ LSTM แบบง่าย ถ้าออกแบบโครงข่ายประสาทเทียมให้มีความซับซ้อนมากขึ้นอาจส่งผลให้ผลการพยากรณ์แม่นมากขึ้น

3. ในงานวิจัยนี้เทคนิคในการแทนที่ข้อมูลสูญหายทั้ง KS และ LSTM มีปัญหากับการแทนที่ข้อมูลที่สูญหายต่อเนื่อง การใช้ Gated Recurrent Unit (GRU) [19] อาจช่วยแทนที่ข้อมูลใกล้เคียงกับข้อมูลที่สูญหาย นำไปสู่ผลการพยากรณ์ที่แม่นยำต่อไป

4 ปัจจุบันมีการประยุกต์ใช้ Convolutional Neural Network (CNN) [20, 21] ทำการสกัดคุณลักษณะกับการพยากรณ์ความหนาแน่นของฝุ่น PM2.5 ได้ผลลัพธ์ที่ดีขึ้น CNN จึงเป็นเทคนิคที่ อาจเพิ่มประสิทธิภาพให้แบบจำลอง LSTM ในการพยากรณ์ค่าหนาแน่น PM2.5 บริเวณสถานีตำรวจนครบาลโชคชัย

บรรณานุกรม

1. Thaweephol, K. and N. Wiwatwattana. *Long Short-Term Memory Deep Neural Network Model for PM2. 5 Forecasting in the Bangkok Urban Area*. in 2019 17*th International Conference on ICT and Knowledge Engineering (ICT&KE)*. 2019. IEEE.
2. Tsai, Y.-T., Y.-R. Zeng, and Y.-S. Chang. *Air pollution forecasting using RNN with LSTM*. in 2018 *IEEE* 16*th Intl Conf on Dependable, Autonomic and Secure Computing,* 16*th Intl Conf on Pervasive Intelligence and Computing,* 4*th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. 2018. IEEE.
3. Park, J.-H., et al. *PM*10 *density forecast model using long short term memory*. in 2017 *Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*. 2017. IEEE.
4. Soldaat, L., et al., *Smoothing and trend detection in waterbird monitoring data using structural time-series analysis and the Kalman filter.* Journal of Ornithology, 2007. 148(2): p. 351-357.
5. Kalman, R.E., *A new approach to linear filtering and prediction problems.* 1960.
6. Zhang, J., Y. Chi, and L. Xiao, *Solar Power Generation Forecast Based on LSTM*. 2018. 869-872.
7. LIU, D.-h. and J.-j. WANG, *A PCA-LSTM Model for Stock Index Prediction.* DEStech Transactions on Engineering and Technology Research, 2018(ecar).
8. Huang, K., et al. *Mood detection from daily conversational speech using denoising autoencoder and LSTM*. in 2017 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2017.
9. Marchi, E., et al. *A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional LSTM neural networks*. in 2015 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2015.

10. Zhu, L. and N. Laptev. *Deep and confident prediction for time series at uber*. in 2017 *IEEE International Conference on Data Mining Workshops (ICDMW)*. 2017. IEEE.
11. Laptev, N., et al. *Time-series extreme event forecasting with neural networks at uber*. in *International Conference on Machine Learning*. 2017.
12. Hochreiter, S. and J. Schmidhuber, *Long short-term memory*. Neural computation, 1997. 9(8): p. 1735-1780.
13. Bhardwaj, R. and D. Pruthi. *Time series and predictability analysis of air pollutants in Delhi*. in 2016 2*nd International Conference on Next Generation Computing Technologies (NGCT)*. 2016. IEEE.
14. Athira, V., et al., *Deepairnet: Applying recurrent networks for air quality prediction.* Procedia computer science, 2018. 132: p. 1394-1403.
15. Ausati, S. and J. Amanollahi, *Assessing the accuracy of ANFIS, EEMD-GRNN, PCR, and MLR models in predicting PM*2. 5. Atmospheric environment, 2016. 142: p. 465-474.
16. Coto-Jimenez, M., et al. *Hybrid Speech Enhancement with Wiener filters and Deep LSTM Denoising Autoencoders*. in 2018 *IEEE International Work Conference on Bioinspired Intelligence (IWOBI)*. 2018. IEEE.
17. Huang, K.-Y., et al., *Speech emotion recognition using autoencoder bottleneck features and LSTM*. 2016. 1-4.
18. Moritz, S. and T. Bartz-Beielstein, *imputeTS: time series missing value imputation in R.* The R Journal, 2017. 9(1): p. 207-218.
19. Che, Z., et al., *Recurrent neural networks for multivariate time series with missing values.* Scientific reports, 2018. 8(1): p. 1-12.
20. Huang, C.-J. and P.-H. Kuo, *A deep cnn-lstm model for particulate matter (PM*2. 5) *forecasting in smart cities.* Sensors, 2018. 18(7): p. 2220.
21. Li, T., M. Hua, and X. Wu, *A hybrid CNN-LSTM model for forecasting particulate matter (PM*2. 5). IEEE Access, 2020. 8: p. 26933-26940.

ประวัติผู้เขียน

ชื่อ-สกุล	ภานุพงษ์ ร่องอ้อ
วัน เดือน ปี เกิด	30 มิถุนายน 2535
สถานที่เกิด	นครสวรรค์
วุฒิการศึกษา	พ.ศ. 2554 วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ จาก มหาวิทยาลัยศรีนครินทรวิโรฒ