

แบบจำลองการพยากรณ์ค่า PM2.5 โดยใช้โครงข่ายประสาทเทียมแบบ LSTM ในพื้นที่กรุงเทพมหานคร

ศรียักษ์ ศรีทองชัย*

*ภาควิชาวิศวกรรมอุตสาหการและการจัดการ คณะวิศวกรรมศาสตร์และเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยศิลปากร,
นครปฐม, ประเทศไทย

*ผู้ประพันธ์บรรณกิจ อีเมล : srithongchai_s@su.ac.th

รับต้นฉบับ : 22 ธันวาคม 2564; รับประทานฉบับแก้ไข : 18 กุมภาพันธ์ 2565; ตอรับบทความ : 1 มีนาคม 2565
เผยแพร่ออนไลน์ : 27 มิถุนายน 2565

บทคัดย่อ

PM2.5 ถือเป็นปัญหาสำคัญของประเทศไทยโดยเฉพาะในกรุงเทพมหานครและปริมณฑล ส่งผลกระทบต่อสุขภาพ เศรษฐกิจและสังคม การพยากรณ์โดยใช้แบบจำลองทางคณิตศาสตร์เพื่อตรวจสอบระดับความเข้มข้นของ PM2.5 เป็นทางเลือกที่ทำให้ทราบค่าประมาณของ PM2.5 งานวิจัยนี้นำเสนอวิธีโครงข่ายประสาทเทียมแบบหน่วยความจำระยะสั้นแบบยาวในการสร้างแบบจำลองเพื่อใช้ในการทำนายปริมาณฝุ่นอนุภาคขนาดเล็กดังกล่าวบริเวณพื้นที่ทั่วไปและริมถนนในกรุงเทพมหานคร ใช้การวิเคราะห์สหสัมพันธ์คัดเลือกตัวแปรหลักเข้าสู่แบบจำลองและเกณฑ์ RMSE และ MAPE เป็นเครื่องมือวัดประสิทธิภาพการพยากรณ์ ซึ่งพัฒนาแบบจำลองที่แตกต่างกันได้ 8 แบบโดยอาศัยข้อมูลคุณภาพอากาศและข้อมูลอุตุนิยมวิทยา ผลการวิจัยพบว่าสำหรับพื้นที่ทั่วไปแบบจำลองที่ 2 ที่ประกอบด้วยตัวแปร PM10, NO2, CO, O3, SO2 และ LAGPM2.5 เป็นแบบจำลองที่ให้ความแม่นยำสูงสุด ให้ค่าเฉลี่ย RMSE = 8.05, ค่าเฉลี่ย MAPE = 27.22 ส่วนพื้นที่ริมถนนแบบจำลองที่ 8 มีตัวแปร PM10, NO2, CO, Temp, Hum, Press, WindSp, WindDir และ LAGPM2.5 เป็นแบบจำลองที่เหมาะสมที่สุดให้ค่าเฉลี่ย RMSE= 4.83, ค่าเฉลี่ย MAPE = 22.57 แบบจำลองบริเวณริมถนน (แบบจำลองที่ 5-8) มีค่าผิดพลาดน้อยกว่าแบบจำลองในพื้นที่ทั่วไป (แบบจำลองที่ 1-4) และการพยากรณ์ที่ใช้ข้อมูลย้อนหลังเป็นระยะสั้น 1 วันมีแนวโน้มให้ผลลัพธ์ที่ถูกต้องมากขึ้น

คำสำคัญ : ฝุ่นละอองขนาดเล็ก การพยากรณ์ โครงข่ายประสาทเทียม หน่วยความจำระยะสั้นแบบยาว



A PM2.5 Prediction Model Using LSTM Neural Network in Bangkok Area

Sriruk Srithongchai*

**Department of Industrial Engineering and Management, Faculty of Engineering and Industrial Technology,
Silpakorn University, Nakhon Pathom, Thailand*

*Corresponding Author. E-mail address: srithongchai_s@su.ac.th

Received: 22 December 2021; Revised: 18 February 2022; Accepted: 1 March 2022

Published online: 27 June 2022

Abstract

PM2.5 has become a serious concern in Thailand, particularly in Bangkok and its vicinity. It has significant impacts on human health, economy and society. A prediction method based on mathematical models is an alternative approach to obtain estimates of PM2.5 concentrations. This research proposes long short-term memory (LSTM) neural networks to develop models to forecast the levels of fine particulate matter in Bangkok, ambient and roadside areas. Correlation analysis was used to select the key variables, and the RMSE and MAPE criteria were employed as forecasting performance measures. Eight different models were constructed using air quality and meteorological data. The results demonstrated that for ambient area, model 2 (which includes the variables PM10, NO2, CO, O3, SO2 and LAGPM2.5) was the best model with an average RMSE of 8.05 and an average MAPE of 27.22. In roadside area, model 8 (which contains the variables PM10, NO2, CO, Temp, Hum, Press, WindSp, WindDir and LAGPM2.5) showed the best performance with an average RMSE of 4.83 and an average MAPE of 22.57. Additionally, the prediction models at roadside site (model 5-8) were more accurate than the others (model 1-4). Estimates based upon short-term past data, 1 day, tended to have smaller forecast errors.

Keywords: PM2.5, Forecasting, Neural network, LSTM

1) บทนำ

สถานการณ์ปัญหามลพิษทางอากาศ (Air Pollution) เกิดขึ้นในหลายภูมิภาคทั่วโลกและนับวันจะทวีความรุนแรงเพิ่มมากขึ้น ช่วงหลายปีที่ผ่านมาเว็บไซต์รายงานคุณภาพอากาศ IQAir.com ได้จัดให้เชียงใหม่และกรุงเทพมหานครติดอันดับเมืองที่พบมลพิษทางอากาศมากที่สุดในโลก [1], [2] ในฐานะที่กรุงเทพมหานครเป็นเมืองหลวงและเมืองศูนย์กลางทางเศรษฐกิจการค้า จึงได้เกิดการอพยพหลังไหลของผู้คนเข้ามาอยู่อาศัยเพื่อแสวงหาโอกาสในการประกอบอาชีพ การศึกษา และความมั่นคงทางด้านรายได้ จำนวนประชากรที่เพิ่มขึ้นก่อให้เกิดมลพิษทางอากาศเพิ่มมากขึ้นตามไปด้วย ซึ่งส่วนใหญ่มีสาเหตุมาจากการใช้ยานพาหนะและการประกอบกิจการอุตสาหกรรม การจราจรที่ติดขัดนานทำให้น้ำมันถูกเผาผลาญมากขึ้น การเผาไหม้ของเชื้อเพลิงที่ไม่สมบูรณ์ของเครื่องยนต์ทำให้เกิดเขม่าควันและสารพิษปล่อยสู่บรรยากาศ เช่น ก๊าซออกไซด์ของไนโตรเจน ก๊าซคาร์บอนมอนอกไซด์ สารประกอบไฮโดรคาร์บอน ก๊าซซัลเฟอร์ไดออกไซด์และฝุ่นละอองที่มีขนาดเล็ก โรงงานอุตสาหกรรมถือเป็นอีกแหล่งกำเนิดหลักของมลพิษทางอากาศที่มาจากการเผาไหม้เชื้อเพลิงและกระบวนการผลิต มลพิษทางอากาศที่อยู่ในรูปของฝุ่นละอองขนาดเล็กมากจนมองไม่เห็นด้วยตาเปล่านั้น ยิ่งฝุ่นเล็กลงเท่าใดก็ยิ่งเป็นอันตรายต่อสุขภาพ เนื่องจากอนุภาคขนาดเล็กสามารถถูกสูดหายใจเข้าไปได้ลึกถึงปอดและทางเดินหายใจ เสี่ยงต่อการเกิดโรคร้ายแรงถึงขั้นเสียชีวิต เช่น มะเร็งปอด หัวใจขาดเลือด และปอดอักเสบ ปัญหาจากฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน หรือ PM2.5 กำลังเป็นวิกฤตการณ์ด้านสิ่งแวดล้อมในกรุงเทพมหานคร และอีกหลายจังหวัดในประเทศไทย มีแนวโน้มจะเกิดขึ้นต่อเนื่องช่วงต้นปีและปลายปีของทุกปีและกินเวลานานขึ้น การตรวจวัด PM2.5 ในอากาศมีความสำคัญมากเพื่อเฝ้าระวัง เตรียมป้องกันและแก้ไขปัญหา รวมถึงแจ้งเตือนให้ประชาชนดูแลสุขภาพอนามัยอย่างถูกต้อง กรมควบคุมมลพิษติดตามตรวจสอบคุณภาพอากาศโดยสถานีตรวจวัดคุณภาพอากาศที่กระจายทั่วพื้นที่ประเทศ ซึ่งยังมีข้อจำกัดด้านงบประมาณและเครื่องมือวัดที่ไม่เพียงพอ การประเมินความเข้มข้น PM2.5 โดยใช้แบบจำลองทางคณิตศาสตร์ อย่างเช่น โครงข่ายประสาทเทียม (Artificial Neural Network) วิธีการวิเคราะห์การถดถอย (Regression Analysis) ฟัชซีลอจิก (Fuzzy Logic) และวิธีเชิงพันธุกรรม (Genetic Algorithm) จึงเป็นอีกแนวทางหนึ่งที่น่าสนใจ เนื่องจากเสียเวลาและค่าใช้จ่ายน้อยกว่าการตรวจวัดจริง

ปัจจุบันโครงข่ายประสาทเทียมถูกนำมาประยุกต์ใช้กันอย่างแพร่หลายสำหรับการจำแนกข้อมูล (Classification) การจัดกลุ่มข้อมูล (Clustering) การรู้จดจำแบบ (Pattern Recognition) และการพยากรณ์ (Prediction) งานวิจัยหลาย ๆ ชิ้นใช้โครงข่ายประสาทเทียมศึกษาทำนายปริมาณของสารมลพิษที่เจือปนอยู่ในอากาศ (เช่น PM2.5, PM10, CO และ SO₂) Kurt [3] Perez [4] Voukantsis [5] และคณะ นำเสนอโครงข่ายแบบป้อนไปข้างหน้า (Feed-Forward Neural Network: FNN) ที่มีการเรียนรู้แบบแพร่ย้อนกลับ (Backpropagation) มาสร้างแบบจำลองพยากรณ์ Oprea [6] Chang-Hoi [7] และผู้ร่วมวิจัย ทำการศึกษาโดยใช้โครงข่ายแบบวนซ้ำ (Recurrent Neural Network: RNN) Tsai [8] Bai [9] และคณะ ใช้โครงข่ายที่เรียกว่าหน่วยความจำระยะสั้นแบบยาว (Long-Short Term Memory: LSTM) ในส่วนของงานวิจัยในประเทศที่เกี่ยวข้อง เช่น Thaweephol และ Wiwatwattana [10] เปรียบเทียบประสิทธิภาพของแบบจำลองโครงข่าย LSTM และแบบจำลองทางสถิติ SARIMAX สำหรับพยากรณ์ฝุ่น PM2.5 รายชั่วโมง ระหว่างปี พ.ศ. 2560-2561 ของพื้นที่กรุงเทพมหานครชั้นใน (สถานีตำรวจนครบาลโชคชัย) พบว่า LSTM แม่นยำมากกว่า SARIMAX โดยใช้ตัวแปรสภาพอากาศและอุตุนิยมวิทยา 16 ตัวแปรและลบทิ้งแถวข้อมูลสูญหาย (Missing Data) Rong-O และ Wiwatwattana [11] ศึกษาผลกระทบการแทนค่าสูญหายและคัดเลือกตัวแปรเข้าสู่แบบจำลอง (Feature Selection) ซึ่งการทำนาย PM2.5 ด้วยวิธี LSTM ในพื้นที่เดียวกับที่ได้กล่าวมาข้างต้นมีความผิดพลาดน้อยลง ทั้งนี้แต่ละอัลกอริทึมโครงข่ายประสาทเทียมมีคุณลักษณะและสมรรถนะแตกต่างกัน โดย FNN ทำงานด้วยการส่งผ่านข้อมูลไปข้างหน้าทิศทางเดียว ขณะที่ RNN ป้อนข้อมูลย้อนกลับจึงเรียนรู้ได้มากขึ้น และถูกพัฒนาต่อยอดเป็น LSTM ช่วยให้เก็บข้อมูลไว้ได้นานและมักใช้กับข้อมูลอนุกรมเวลา (Time Series Data) LSTM จึงเหมาะสมกับข้อมูลที่ใช้ในการศึกษาครั้งนี้ซึ่งก็คือ ข้อมูลคุณภาพอากาศและอุตุนิยมวิทยา (Air Quality and Metrological Data) งานวิจัยนี้เสนอแบบจำลองโครงข่าย LSTM เพื่อพยากรณ์ค่า PM2.5 ในกรุงเทพมหานคร โดยเพิ่มเติมจากงานวิจัยที่ผ่านมาด้วยการพิจารณาเปรียบเทียบแบบจำลองหลายตัวแปรที่แตกต่างกันรวมทั้งศึกษาในพื้นที่ทั่วไป (Ambient Area) ที่ใช้ข้อมูลจากสถานีตรวจวัดคุณภาพอากาศบางนาและสถานีอุตุนิยมวิทยาบางนาและบริเวณริมถนน (Roadside Area) ใช้ข้อมูลของสถานีโรงพยาบาลจุฬาลงกรณ์และสถานีศูนย์ประชุมแห่งชาติสิริกิติ์

2) ข้อมูลการวิจัย

กรุงเทพมหานคร [12] ตั้งอยู่ภาคกลางตอนล่างของประเทศ บริเวณละติจูดที่ 13.44 องศาเหนือ ลองจิจูด 100.34 องศาตะวันออก ลักษณะภูมิประเทศเป็นพื้นที่ราบลุ่ม แม่น้ำเจ้าพระยาเป็นแม่น้ำสายหลักไหลผ่าน มีพื้นที่ทั้งหมดประมาณ 1,568.74 ตารางกิโลเมตร อากาศเป็นแบบร้อนชื้นมีอุณหภูมิเฉลี่ยตลอดปี 28-30 องศาเซลเซียส มีสถานีอุตุนิยมวิทยาจำนวน 4 สถานีและสถานีตรวจวัดคุณภาพอากาศจำนวน 62 สถานี (ภายใต้การกำกับดูแลของกรมควบคุมมลพิษ 12 สถานีและกรุงเทพมหานคร 50 สถานี) ผู้วิจัยสนใจศึกษาปริมาณฝุ่นละออง PM2.5 ในกรุงเทพมหานครทั้งบริเวณพื้นที่ทั่วไปและพื้นที่ริมถนน โดยใช้ข้อมูลสารก่อกมลพิษทางอากาศจากสถานีตรวจวัดคุณภาพอากาศบางนาเป็นตัวแทนข้อมูลพื้นที่ทั่วไป และสถานีโรงพยาบาลจุฬาลงกรณ์เป็นตัวแทนของพื้นที่ริมถนน ซึ่งพิจารณาจากเกณฑ์ความสมบูรณ์ของข้อมูลเพื่อให้ได้ข้อมูลที่ครบถ้วนมากที่สุดในการดำเนินการวิจัย รวมไปถึงข้อมูลสภาพภูมิอากาศที่มีผลต่อความหนาแน่นของฝุ่นพิษ จึงได้คัดเลือกสถานีอุตุนิยมวิทยาที่อยู่ในสถานที่ที่ใกล้เคียง ห่างจากสถานีตรวจวัดคุณภาพอากาศไม่เกิน 3 กิโลเมตร ดังนั้นข้อมูลพื้นที่ทั่วไปได้มาจากสถานีตรวจวัดคุณภาพอากาศที่บางนาและสถานีอุตุนิยมวิทยาบางนา ข้อมูลพื้นที่ริมถนนมาจากสถานีโรงพยาบาลจุฬาลงกรณ์และสถานีอุตุนิยมวิทยาศูนย์ประชุมแห่งชาติสิริกิติ์ โดยข้อมูลดิบที่ได้จากการเก็บรวบรวมจะได้ข้อมูลทั้งในแบบรายชั่วโมง (ข้อมูลคุณภาพอากาศ) ราย 3 ชั่วโมง (ความกดอากาศ ความเร็วลม ทิศทางลม) และรายวัน (อุณหภูมิ ความชื้นสัมพัทธ์ ปริมาณน้ำฝน)

งานวิจัยใช้ข้อมูลรายวัน ข้อมูลที่ไม่ใช่รายวันให้ใช้ค่าเฉลี่ยต่อวันแทนข้อมูลรายวัน ตั้งแต่เดือนมกราคม พ.ศ. 2560 ถึงเดือนธันวาคม พ.ศ. 2563 แบ่งได้ 2 ส่วน ประกอบด้วย ข้อมูลคุณภาพอากาศที่เป็นฝุ่นละอองและสารมลพิษในอากาศได้แก่ ฝุ่น PM2.5 ($\mu\text{g}/\text{m}^3$) ฝุ่น PM10 ($\mu\text{g}/\text{m}^3$) ก๊าซคาร์บอนมอนอกไซด์ (ppm) ก๊าซซัลเฟอร์ไดออกไซด์ (ppb) ก๊าซไนโตรเจนไดออกไซด์ (ppb) และก๊าซโอโซน (ppb) และข้อมูลอุตุนิยมวิทยา ได้แก่ อุณหภูมิ ($^{\circ}\text{C}$) ความชื้นสัมพัทธ์ (%) ปริมาณน้ำฝน (mm) ความกดอากาศ (hPa) ความเร็วลม (knot) และทิศทางลม (degree) ในตารางที่ 1 และตารางที่ 2 สรุปผลข้อมูลคุณภาพอากาศและอุตุนิยมวิทยาพบว่าพื้นที่ริมถนนมีค่าเฉลี่ยมลพิษทางอากาศสูงกว่าพื้นที่ทั่วไป ค่า PM2.5 สูงสุดในแต่ละพื้นที่อยู่ในระดับที่เกินกว่ามาตรฐาน ($50 \mu\text{g}/\text{m}^3$) แต่อย่างไรก็ตามค่าเฉลี่ย 24 ชั่วโมง (รายวัน) ของฝุ่น

PM2.5 ส่วนใหญ่จะต่ำกว่าค่ามาตรฐานของประเทศไทย ดังรูปที่ 1 และ 2

ตารางที่ 1 : ค่าสถิติเบื้องต้นข้อมูลรายวันของพื้นที่ทั่วไป ปี 2560-2563

ข้อมูล (ตัวแปร)	ค่าเฉลี่ย	ค่า	ค่าต่ำสุด/สูงสุด
ฝุ่น PM2.5 (PM2.5)	24.04	14.55	3.50-100.00
ฝุ่น PM10 (PM10)	40.63	19.88	7.42-153.46
คาร์บอนมอนอกไซด์ (CO)	0.47	0.21	0.04-1.40
ซัลเฟอร์ไดออกไซด์ (SO2)	1.55	1.06	0.00-6.83
ไนโตรเจนไดออกไซด์	16.27	9.70	0.57-58.22
โอโซน (O3)	19.14	9.38	1.68-52.52
อุณหภูมิ (Temp)	29.10	1.56	20.70-33.40
ความชื้นสัมพัทธ์ (Hum)	74.39	7.94	46.00-96.00
ปริมาณน้ำฝน (Rainfall)	6.89	14.67	0.00-119.00
ความกดอากาศ (Press)	1009.4	2.66	1018.99-1018.65
ความเร็วลม (WindSp)	2.83	1.25	0.00-7.00
ทิศทางลม (WindDir)	114.93	58.21	0.00-290.00

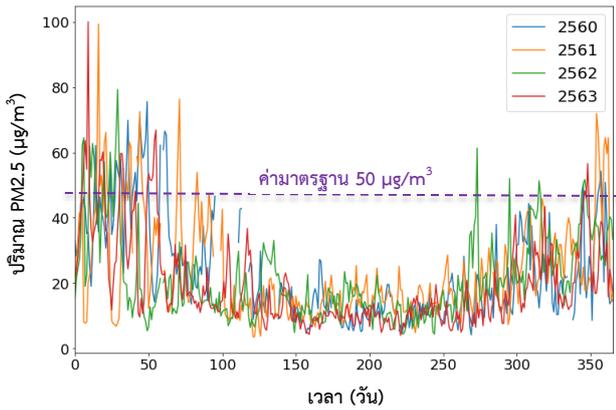
ตารางที่ 2 : ค่าสถิติเบื้องต้นข้อมูลรายวันของพื้นที่ริมถนน ปี 2560-2563

ข้อมูล (ตัวแปร)	ค่าเฉลี่ย	ค่า	ค่าต่ำสุด/สูงสุด
ฝุ่น PM2.5 (PM2.5)	26.22	12.92	7.83-92.25
ฝุ่น PM10 (PM10)	51.13	19.81	19.71-141.71
คาร์บอนมอนอกไซด์ (CO)	1.15	0.42	0.11-2.51
ซัลเฟอร์ไดออกไซด์ (SO2)	NA	NA	NA
ไนโตรเจนไดออกไซด์	34.77	20.49	8.17-123.30
โอโซน (O3)	NA	NA	NA
อุณหภูมิ (Temp)	29.19	1.62	20.60-33.60
ความชื้นสัมพัทธ์ (Hum)	74.68	7.75	47.00-97.00
ปริมาณน้ำฝน (Rainfall)	7.10	15.79	0.00-188.30
ความกดอากาศ (Press)	1008.9	2.65	1001.21-
ความเร็วลม (WindSp)	1.22	0.72	0.00-4.00
ทิศทางลม (WindDir)	71.64	46.27	0.00-290.00

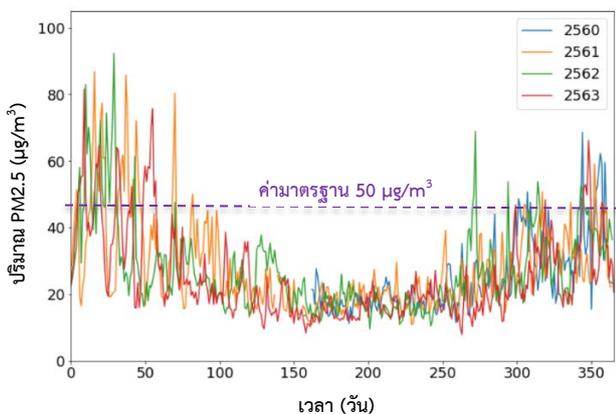
หมายเหตุ: NA หมายถึงไม่มีข้อมูล

3) การดำเนินงานวิจัย

โครงข่ายประสาทเทียมเลียนแบบมาจากการทำงานของสมองมนุษย์ จะเรียนรู้ผ่านข้อมูลตัวอย่างและประมวลผลข้อมูลอื่นโดยอาศัยประสบการณ์ความรู้ที่ได้รับมา ในที่นี่จะใช้โครงข่ายประสาทเทียม LSTM สร้างแบบจำลองทำนาย PM2.5 โดยคัดเลือกตัวแปรสำคัญ (Feature Selection) ที่มีผลต่อตัวแปร ถูกทำนายด้วยสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson Correlation Coefficient) และวัดประสิทธิภาพของแบบจำลองด้วยค่ารากที่สองความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) และเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE)



รูปที่ 1 : ปริมาณฝุ่น PM2.5 เฉลี่ย 24 ชั่วโมงของพื้นที่ทั่วไป ปี 2560-2563



รูปที่ 2 : ปริมาณฝุ่น PM2.5 เฉลี่ย 24 ชั่วโมงของพื้นที่ริมถนน ปี 2560-2563

3.1) การเตรียมข้อมูล

ข้อมูลที่น่าสนใจมาเป็นข้อมูลอนุกรมเวลาลักษณะรายวัน ซึ่งข้อมูลคุณภาพอากาศได้รับจากผลการตรวจวัดโดยสถานีตรวจวัดคุณภาพอากาศบางนาและสถานีโรงพยาบาลจุฬาลงกรณ์ และข้อมูลอุตุนิมิตวิทยาจากสถานีอุตุนิมิตวิทยาบางนาและสถานีศูนย์ประชุมแห่งชาติสิริกิติ์ โดยใช้ข้อมูลย้อนหลัง 4 ปีล่าสุด (พ.ศ. 2560-2563) ข้อมูลรายวันของแต่ละพื้นที่มีทั้งหมด 1,461 แถว แต่เนื่องจากข้อมูลไม่ครบถ้วนสมบูรณ์ ข้อมูลที่ขาดหายไปอย่างต่อเนื่องติดต่อกันนานจะถูกตัดทิ้งและทดแทนค่าสูญหายกรณีอื่นด้วยค่าเฉลี่ย (Mean) เช่น ในช่วง 9 เดือนหลังของปี 2563 พื้นที่ทั่วไปไม่มีข้อมูล “คาร์บอนมอนนอกไซด์” “ซัลเฟอร์ไดออกไซด์” และ “ไนโตรเจนไดออกไซด์” ค่า “โอโซน” ขาดหายมีมากถึง 92% จึงลบทิ้งแถวข้อมูลของช่วงเวลาดังกล่าว และ “ปริมาณน้ำฝน” ทั้ง 4 ปี ค่าหายไปกว่า 25% จึงตัดปริมาณน้ำฝนออกไปจากตัวแปรที่ทำการศึกษา [13] ดังนั้นพื้นที่ทั่วไปมีขนาดข้อมูลลดลงเหลือเท่ากับ 1,370 แถวและบริเวณพื้นที่ริมถนนมีจำนวนคงเหลือ 1,247 แถว

ข้อมูลข้างต้นถูกจัดแบ่งเป็น 3 ชุดคือข้อมูลชุดสอน (Training Set) นำไปสอนให้โครงข่ายเกิดการเรียนรู้ ข้อมูลชุดตรวจสอบ (Validation Set) เพื่อประเมินการทำงานของแบบจำลองที่สร้างขึ้นและข้อมูลชุดทดสอบ (Test Set) ใช้ในการทดสอบประสิทธิภาพการทำงานของแบบจำลอง ในลำดับแรกทำการแบ่งข้อมูลออกเป็น 2 ส่วน 80% สำหรับสอนและ 20% สำหรับทดสอบโดยให้ข้อมูลล่าสุดเป็นชุดทดสอบ จากนั้นตัดเอาส่วนท้าย 20% ของข้อมูลชุดสอนที่ได้มาไว้เป็นชุดตรวจสอบ ส่วนที่เหลืออยู่จะเป็นข้อมูลชุดสอนที่แท้จริง

3.2) การคัดเลือกตัวแปร

แบบจำลองโครงข่ายประสาทเทียมสำหรับทำนายปริมาณฝุ่น PM2.5 มีตัวแปรผลลัพธ์ (Output Variable) 1 ตัวแปรคือความเข้มข้นของ PM2.5 และตัวแปรป้อนเข้า (Input Variable) ที่ส่งผลต่อตัวแปรผลลัพธ์มีหลายตัว การพิจารณาคัดเลือกตัวแปรป้อนเข้าดูจากค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันที่บ่งชี้ความสัมพันธ์ระหว่างตัวแปรสองตัว ค่ายิ่งสูงแสดงถึงตัวแปรป้อนเข้า นั้นจะยังมีบทบาทสำคัญต่อการพยากรณ์ รูปที่ 3 และ 4 แสดงตารางเมทริกซ์สหสัมพันธ์แบบแผนภูมิความร้อน (Heat Map) ใช้ความเข้มของสีในการนำเสนอข้อมูล ยิ่งสีเข้มขนาดความสัมพันธ์ก็จะยิ่งสูงขึ้นและแสดงผลเป็นตัวเลขในแต่ละช่องของตารางด้วย

หลังการจัดการกับข้อมูลสูญหาย ตัวแปร Rainfall ถูกตัดทิ้งไม่อยู่ในขอบข่ายการศึกษา เนื่องจากจำนวนข้อมูลสูญหายมีมากเกินไป ดังนั้นพื้นที่ทั่วไปจึงมีตัวแปร 11 ตัว ได้แก่ PM2.5 PM10 CO SO2 NO2 O3 Temp Hum Press WindSp และ WindDir (ดูตารางที่ 1) จากรูปที่ 3 จะเห็นว่าตัวแปรป้อนเข้าที่มีอิทธิพลต่อตัวแปรผลลัพธ์ (PM2.5) สูงสุด 5 อันดับแรกคือ PM10 NO2 CO O3 และ SO2 ส่วนริมถนนมีตัวแปรเกี่ยวข้อง 9 ตัว ได้แก่ PM2.5 PM10 CO NO2 Temp Hum Press WindSp WindDir และตัวแปรป้อนเข้าสำคัญสุด 5 อันดับคือ PM10 Press Hum CO และ NO2 ดังรูปที่ 4 ตัวแปรป้อนเข้าที่นำมาใช้สร้างแบบจำลองถือเป็นปัจจัยหนึ่งที่มีผลต่อความถูกต้องของการพยากรณ์ งานวิจัยนี้คัดเลือกตัวแปรป้อนเข้าที่แตกต่างกันลงในแบบจำลอง ทำให้ได้แบบจำลอง 8 แบบดังตารางที่ 3 โดยใช้หลักเกณฑ์ดังนี้

- เลือกตัวแปรป้อนเข้าที่มีสัมประสิทธิ์สหสัมพันธ์สูงสุด 5 อันดับ
- ใช้ตัวแปรเช่นเดียวกับเกณฑ์ที่ 1 และเพิ่มตัวแปรค่า PM2.5 ของวันก่อนหน้านั้น (LAGPM2.5)



รูปที่ 3 : แผนภูมิความสัมพันธ์สหสัมพันธ์สำหรับพื้นที่ทั่วไป



รูปที่ 4 : แผนภูมิความสัมพันธ์สหสัมพันธ์สำหรับพื้นที่ริมถนน

ตารางที่ 3 : ตัวแปรในแบบจำลอง

แบบจำลอง	ตัวแปรป้อนเข้า
1	PM10 NO2 CO O3 SO2
2	PM10 NO2 CO O3 SO2 LAGPM2.5
3	PM10 NO2 CO O3 SO2 Temp Hum Press WindSp WindDir
4	PM10 NO2 CO O3 SO2 Temp Hum Press WindSp WindDir LAGPM2.5
5	PM10 Press Hum NO2 CO
6	PM10 Press Hum NO2 CO LAGPM2.5
7	PM10 NO2 CO Temp Hum Press WindSp WindDir
8	PM10 NO2 CO Temp Hum Press WindSp WindDir LAGPM2.5

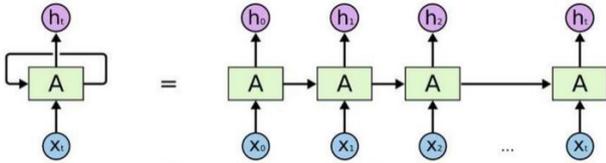
- รวมเอาตัวแปรทั้งหมดเข้าไว้ในแบบจำลอง
- เลือกใช้ตัวแปรตามเกณฑ์ที่ 3 และเพิ่มตัวแปร LAGPM2.5

3.3) การสร้างแบบจำลองและพยากรณ์

3.3.1) หน่วยความจำระยะสั้นแบบยาว แนวคิดเรื่องโครงข่ายประสาทเทียมถูกกำเนิดขึ้นมาตั้งแต่ต้นทศวรรษ 40 (พ.ศ. 2486) McCulloch และ Pitts [14] นำเสนอโครงข่ายประสาทเทียมอย่างง่ายมีรากฐานมาจากศาสตร์ทางด้านสรีรวิทยาและการทำงานของเซลล์ประสาทในสมองมนุษย์ ซึ่งได้รับความสนใจจากนักวิจัยและเกิดการพัฒนามาอย่างต่อเนื่องเพื่อให้ครอบคลุมปัญหาที่มีความยุ่งยากและซับซ้อน เมื่อราว ๆ 25 ปีก่อน LSTM ซึ่งจัดว่าเป็นโครงข่ายประเภท RNN ได้ถูกคิดค้นขึ้นโดย Hochreiter และ Schmidhuber [15] เพื่อใช้แก้ปัญหาบางอย่างของ RNN โดยที่ RNN เป็นโครงข่ายประสาทเชิงลึก (Deep Neural Network) มีลักษณะการทำงานแบบวนซ้ำ นอกจากจะรับข้อมูลจากข้อมูลขาเข้าแล้วยังนำเอาผลลัพธ์ของเซลล์ประสาทก่อนหน้าเข้ามาาร่วมด้วย ทำให้ RNN สามารถใช้งานได้ดีในข้อมูลแบบเรียงลำดับ (Sequential Data) เช่น คลิปเสียง คลิปวิดีโอ ข้อความตัวอักษรและข้อมูลอนุกรมเวลา รูปที่ 5 ทางซ้ายแสดงโครงสร้าง RNN ที่มีลูปวนกลับนำผลลัพธ์กลับมาใช้ใหม่อีกครั้ง เป็นข้อมูลเข้า ส่วนในรูปด้านขวาเป็นรูปที่คลี่รูปด้านซ้ายออกเพื่อแสดงถึงขั้นตอนการทำงานทีละขั้นตอน ตัวแปร x_t และ h_t แทนข้อมูลนำเข้าและข้อมูลส่งออกที่เวลา t ตามลำดับ และ A หมายถึงชั้นซ่อน (Hidden Layer) แต่บ่อยครั้งพบปัญหาเรื่องค่าเกรเดียนที่น้อยมากในข้อมูลที่มีขนาดยาว หรือที่เรียกว่าปัญหาการสูญหายของเกรเดียน (Vanishing Gradient Problem) จึงยากที่จะเก็บรักษาความจำระยะยาวได้

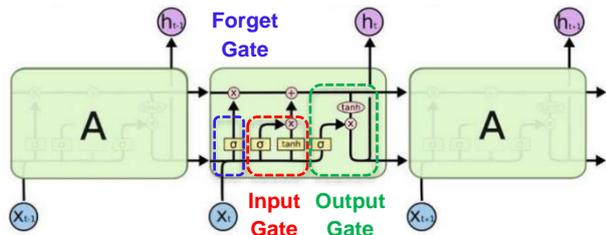
LSTM ต่อยอดมาจาก RNN เพื่อช่วยแก้ไขข้อจำกัดของ RNN ดังกล่าวข้างต้น ด้วยการออกแบบให้โครงข่ายมีความสามารถในการเรียนรู้และจดจำได้นานขึ้น ใช้ฟังก์ชันที่เปรียบเสมือนประตู (Gate) เป็นตัวคอยควบคุมการไหลของข้อมูลประกอบด้วย ประตูลืม (Forget Gate) ประตูทางเข้า (Input Gate) และประตูทางออก (Output Gate) รูปที่ 6 ตัวอย่างการทำงานภายในเซลล์ของ LSTM ประตูลืม มีหน้าที่กำหนดว่าข้อมูลที่เข้ามาควร จะเก็บเอาไว้หรือทิ้งไป โดยรับข้อมูลปัจจุบันและข้อมูลจากสถานะซ่อนก่อนหน้าซึ่งใช้เป็นแหล่งข้อมูลเข้าสำหรับประตูอื่น ๆ ด้วยเช่นกันแล้วคำนวณผ่านฟังก์ชันซิกมอยด์ ประตูทางเข้า เพื่อเปิดรับข้อมูลที่เข้ามาใหม่แล้วจึงบันทึกข้อมูลลงในเซลล์ของตัว

มันเอง คำนวณผลด้วยฟังก์ชันซิกมอยด์ ในขณะที่เดียวกันจะปรับค่าของสถานะเซลล์ด้วยฟังก์ชัน tanh *ประตูทางออก* ใช้กำหนดค่าสถานะที่ซ่อนอยู่ถัดไปรวมถึงส่งผลลัพธ์ไปเป็นข้อมูลออก ข้อมูลเข้าถูกส่งผ่านฟังก์ชันซิกมอยด์และสถานะเซลล์ผ่านฟังก์ชัน tanh



รูปที่ 5 : โครงข่ายประสาทเทียมแบบวนซ้ำ (RNN)

แหล่งที่มา: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>



รูปที่ 6 : หน่วยความจำระยะสั้นแบบยาว (LSTM)

แหล่งที่มา: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

3.3.2) *แบบจำลองการพยากรณ์* แบบจำลองโครงข่ายประสาทเทียม LSTM ที่ใช้คาดการณ์ค่าความเข้มข้น PM2.5 มีโครงสร้างเป็นแบบหลายชั้นคือ ชั้นอินพุท (Input layer) มีเพียงชั้นเดียวและจำนวนโหนด (Node) ในชั้นเท่ากับจำนวนตัวแปรป้อนเข้า ชั้นเอาต์พุท (Output Layer) มีชั้นเดียวและจำนวนโหนดมีค่าเท่ากับจำนวนตัวแปรผลลัพธ์ และชั้นซ่อนเป็นชั้นที่อยู่ระหว่างกลางชั้นอินพุทและเอาต์พุท ในการเรียนรู้ของแบบจำลองต้องอาศัยค่าไฮเปอร์พารามิเตอร์ (Hyperparameter) ซึ่งถูกกำหนดไว้ดังนี้

- จำนวนชั้นซ่อน 2 ชั้น เหมาะกับปัญหาที่ไม่ซับซ้อน [16]
- จำนวนโหนดในแต่ละชั้นซ่อน 32 โหนด ซึ่งตามหลักการทั่วไป (Rule of Thumb) จำนวนโหนดในชั้นซ่อนไม่เกินสองเท่าของ จำนวนโหนดชั้นอินพุท การศึกษาวิจัยนี้มีจำนวนโหนดชั้นอินพุทสูงสุดเท่ากับ 11 ดังนั้นจำนวนโหนดชั้นซ่อนควรน้อยกว่า $2 \times 11 = 22$ แต่ใช้จำนวนโหนดที่มากกว่าโดยแก้ปัญหการจดจำ

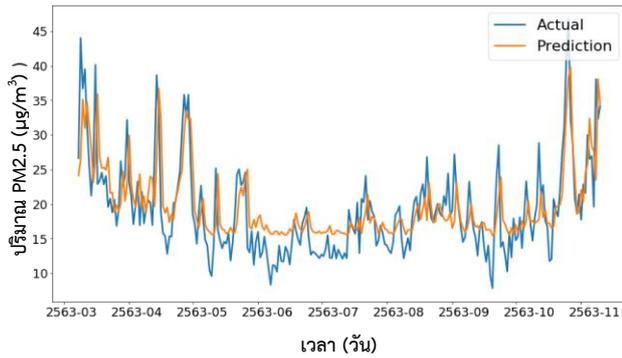
รูปแบบของข้อมูลฝึกสอนมากเกินไป (Overfitting) ด้วยการสุ่มลดจำนวน โหนดบางส่วน (Dropout)

- ขนาดข้อมูลที่ใช้เพื่อการปรับค่าน้ำหนัก (Weight) ในแต่ละครั้ง (Batch Size) เท่ากับ 32 เป็นค่าทั่วไปที่ใช้สำหรับกรณีที่มีจำนวนข้อมูลไม่มากนัก (~1,000 รายการ) [17]
- จำนวนรอบที่สอน (Epoch) 100 รอบ ใช้วิธีการลองผิดลองถูก (Trial and Error) ดูจากฟังก์ชันการสูญเสีย (Loss Function) ที่ค่าลดลงไปเรื่อย ๆ จนไม่เปลี่ยนแปลงหรือเปลี่ยนแปลงน้อย
- สุ่มลดจำนวนโหนด 20% เป็นค่าเบื้องต้นที่นำมาพิจารณาการลดด้วยสัดส่วนที่มากหรือน้อยเกินไป จะทำให้ความสามารถในการทำนายผลไม่ดีเท่าที่ควร ตัวอย่างงานวิจัยที่วิเคราะห์การลดจำนวนโหนดอย่างเหมาะสมสำหรับ LSTM เสนอโดย Cheng และคณะ [18]

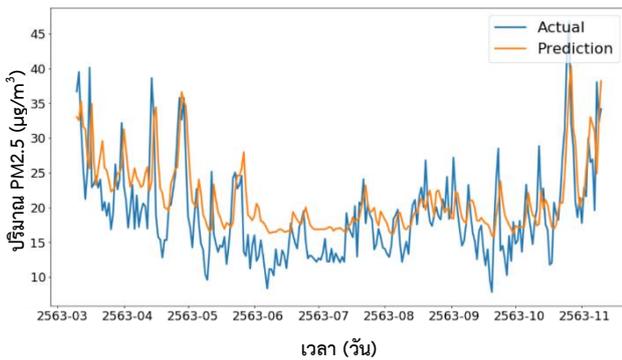
แบบจำลองการพยากรณ์นี้สร้างขึ้นโดยการใช้ Keras และได้ นำเอาวิธีการ MinMaxScaler มาปรับขอบเขตของข้อมูลให้อยู่ในช่วงเดียวกัน (Normalization) คือระหว่าง 0 กับ 1 คำนวณจากความแตกต่างของค่าปัจจุบันกับค่าต่ำสุดหารด้วยพิสัย (Range) เหมาะกับข้อมูลไม่เป็นรูประฆังคว่ำ (Normal Distribution) และสอดคล้องกับผลลัพธ์ไม่ติดค่าลบ ซึ่งได้จากฟังก์ชันการกระตุ้นรีลู (ReLU Function) ใช้อัลกอริทึม Adam (Adaptive Moment Estimation) และให้ความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) เป็นฟังก์ชันการสูญเสียที่วัดความผิดพลาดจากการทำนาย รูปที่ 7, 8 และ 9 เป็นผลการพยากรณ์ปริมาณ PM2.5 ล่วงหน้า 1 วัน ที่ได้จากแบบจำลองที่ 6 ช่วงเดือนมีนาคมถึงพฤศจิกายนของปี 2563 โดยใช้ข้อมูลย้อนหลัง 1 วัน, 3 วัน และ 5 วัน ตามลำดับ

4) ผลการวิจัย

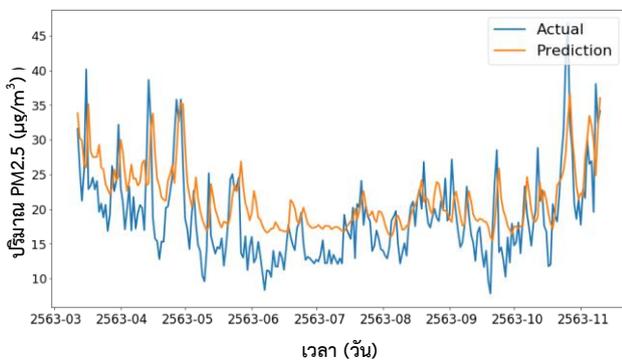
แบบจำลองโครงข่าย LSTM ทั้ง 8 แบบ (ดูตารางที่ 3) ถูกนำมาใช้พยากรณ์ความหนาแน่นฝุ่นละออง PM2.5 ของวันรุ่งขึ้น บริเวณพื้นที่ที่ทั่วไปและริมถนน โดยอาศัยข้อมูลในอดีตที่ผ่านมา 1 วัน, 3 วัน และ 5 วัน รูปที่ 7-9 กราฟแสดงผลเปรียบเทียบค่าพยากรณ์ PM2.5 กับข้อมูลจริง (Actual Data) จะเห็นได้ว่าค่าพยากรณ์เพิ่มขึ้นหรือลดลงผันแปรตามฤดูกาลที่เปลี่ยนแปลงไปในลักษณะเช่นเดียวกับค่าจริงในวันนั้น ๆ ค่าพยากรณ์ในรูปที่ 7 มีความถูกต้องใกล้เคียงความจริงมากกว่าในรูปที่ 8 และ 9 ที่ซึ่งความแม่นยำของการคาดการณ์หาจากตัวชี้วัดสองตัว คือ RMSE และ MAPE ดังตาราง 4 พบว่าค่า RMSE อยู่ระหว่าง 4.55-9.76



รูปที่ 7 : เปรียบเทียบค่าพยากรณ์และค่าจริง PM2.5 (ข้อมูลย้อนหลัง 1 วัน)



รูปที่ 8 : เปรียบเทียบค่าพยากรณ์และค่าจริง PM2.5 (ข้อมูลย้อนหลัง 3 วัน)



รูปที่ 9 : เปรียบเทียบค่าพยากรณ์และค่าจริง PM2.5 (ข้อมูลย้อนหลัง 5 วัน)

และ MAPE มีค่าในช่วง 20.47-37.28 แบบจำลองที่ให้ค่า RMSE หรือ MAPE ต่ำสุดจัดว่าเป็นแบบจำลองพยากรณ์ที่ดีที่สุดในพื้นที่ทั่วไปพบว่าแบบจำลองที่ 2 ซึ่งใช้ตัวแปรในการศึกษาจากการคัด เลือกตัวแปรสำคัญร่วมกับตัวแปร LAGPM2.5 (ตัวแปรค่า PM2.5 ของวันที่ผ่านมา) เป็นแบบจำลองที่มีความแม่นยำสูงสุด สำหรับพื้นที่ริมถนนพบว่าแบบจำลองที่ 8 ประกอบด้วยตัวแปรทั้งหมดและ LAGPM2.5 ทำนายล่วงหน้าอย่างถูกต้องแม่นยำที่สุด และยังเห็นได้ชัดว่าแบบจำลองบริเวณริมถนน (แบบจำลองที่ 5-8) ให้ความผิดพลาดที่ได้จากการพยากรณ์น้อย

กว่าแบบ จำลองในพื้นที่ทั่วไป (แบบจำลองที่ 1-4) และการทำนายด้วยข้อมูลย้อนหลังในระยะเวลาสั้นกว่าจะให้ความแม่นยำที่เพิ่มขึ้น ซึ่งข้อมูลย้อนหลัง 1 วันให้ความคลาดเคลื่อนต่ำสุด

ตารางที่ 4 : ผลเปรียบเทียบความแม่นยำของแบบจำลองพยากรณ์

แบบจำลอง	RMSE			MAPE		
	1 วัน	3 วัน	5 วัน	1 วัน	3 วัน	5 วัน
1	7.98	8.29	9.76	32.68	34.22	33.58
2	7.73	7.89	8.52	27.81	26.28	27.57
3	8.31	8.49	9.36	32.02	33.75	37.28
4	7.82	7.66	8.69	31.43	29.67	29.78
5	5.09	5.71	5.90	24.45	28.42	28.88
6	4.82	5.05	5.36	22.88	24.27	25.57
7	5.02	4.84	6.58	22.96	21.96	30.08
8	4.55	4.75	5.18	20.47	23.11	24.13

5) สรุปและอภิปรายผล

โครงข่ายประสาทเทียมหน่วยความจำระยะสั้นแบบยาวได้ถูกนำมาประยุกต์ใช้ในการสร้างแบบจำลองการทำนายระดับความเข้มข้นของสารมลพิษอากาศ PM2.5 ในกรุงเทพมหานคร บริเวณพื้นที่ทั่วไปและพื้นที่ริมถนน รวมทั้งสิ้นจำนวน 8 แบบจำลอง โดยเปรียบเทียบประสิทธิภาพการพยากรณ์ด้วยค่าความคลาดเคลื่อน RMSE และ MAPE พบว่าพื้นที่ทั่วไปแบบจำลองที่ 2 สามารถทำนาย PM2.5 ได้แม่นยำที่สุดเพราะเลือกใช้ตัวแปร 5 ตัวที่มีอิทธิพลสูงสุดต่อตัวแปรผลลัพธ์รวมถึง LAGPM2.5 ที่เป็นค่า PM2.5 ของวันก่อนหน้า ในพื้นที่ริมถนนแบบจำลองที่ 8 เป็นตัวแบบพยากรณ์ที่ดีที่สุด ใช้ตัวแปรทุกตัวร่วมกับ LAGPM2.5 อีกทั้งแบบจำลองอื่น (ตารางที่ 4) ที่มี LAGPM2.5 เป็นตัวแปรป้อนเข้าให้ผลการคาดการณ์ได้ดีกว่ากรณีละเลย LAGPM2.5 จึงสรุปได้ว่า LAGPM2.5 เป็นตัวแปรสำคัญที่ควรมีอยู่ในแบบจำลองการพยากรณ์ เมื่อพิจารณาตัวแปรในแบบจำลองที่ 8 จะเห็นว่าตัวแปรบางตัวมีความสัมพันธ์เกี่ยวข้องกับน้อยมากกับตัวแปรผลลัพธ์ ดังนั้นการคัดเลือกตัวแปรเข้าด้วยค่าสัมประสิทธิ์สหสัมพันธ์อาจเป็นวิธีการที่มิได้เท่าที่ควรกับปัญหาวิจัยนี้ ผลการศึกษาปรากฏว่าการใช้ข้อมูล 1 วันย้อนหลังให้ความถูกต้องแม่นยำของค่าพยากรณ์มากขึ้นด้วยเหตุผลที่ข้อมูลล่าสุดเป็นข้อมูลที่ใกล้เคียงกับปัจจุบันมากที่สุด ในการสร้างแบบจำลองการพยากรณ์นั้นควรคำนึงถึงปัจจัยที่อาจส่งผลกระทบต่อความแม่นยำของการคาดประมาณ ตัวอย่างเช่น การกำหนดค่าไฮเปอร์พารามิเตอร์ และวิธีการประมาณค่าสูญหาย เป็นต้น

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนอุดหนุนการวิจัยจากกองทุนสนับสนุนการวิจัย นวัตกรรมและการสร้างสรรค์ มหาวิทยาลัยศิลปากร ประจำปี 2562 และได้รับความอนุเคราะห์อย่างยิ่งด้านข้อมูลเพื่อการวิจัยจากกองจัดการคุณภาพอากาศและเสียง กรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม และกรมอุตุนิยมวิทยา กระทรวงดิจิทัลเพื่อเศรษฐกิจและสังคม ผู้วิจัยจึงขอขอบพระคุณอย่างสูงมา ณ โอกาสนี้

REFERENCES

- [1] Coconuts Bangkok. "2 Thai cities cracked world's top 10 worst air pollution list this morning." COCONUTS.co. <https://coconuts.co/bangkok/news/2-thai-cities-cracked-worlds-top-10-worst-air-pollution-list-morning/> (accessed Feb. 28, 2022).
- [2] Prachachat. "PM 2.5 dust pushes Chiang Mai, the world's number 1 toxic air." PRACHACHAT.net. <https://www.prachachat.net/general/news-625935> (accessed Nov. 16, 2021).
- [3] A. Kurt, B. Gulbagci, F. Karaca, and O. Alagha, "An online air pollution forecasting system using neural networks," *Environ. Int.*, vol. 34, no. 5, pp. 592–598, Jul. 2008.
- [4] P. Perez, A. Trier, and J. Reyes, "Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile," *Atmos. Environ.*, vol. 34, no. 8, pp. 1189–1196, Jun. 2000.
- [5] D. Voukantsis, K. Karatzas, J. Kukkonen, T. Räsänen, A. Karppinen, and M. Kolehmainen, "Intercomparison of air quality data using principal component analysis, and forecasting of PM10 and PM2.5 concentrations using artificial neural networks, in Thessaloniki and Helsinki," *Sci. Total Environ.*, vol. 409, no. 7, pp. 1266–1276, Mar. 2011.
- [6] M. Opera, M. Popescu, and S. Mihalache, "A neural network based model for PM2.5 air pollutant forecasting," in *20th Int. Conf. Syst. Theory, Control and Comput. (ICSTCC)*, Sinaia, Romania, Oct. 2016, pp. 776–781, doi: 10.1109/ICS TCC.2016.7790762.
- [7] H. Chang-Hoi et al., "Development of a PM2.5 prediction model using a recurrent neural network algorithm for the Seoul metropolitan area, Republic of Korea," *Atmos. Environ.*, vol. 245, Jan. 2021, Art. no. 118021.
- [8] Y.-T. Tsai, Y.-R. Zeng, and Y.-S. Chang, "Air Pollution Forecasting Using RNN with LSTM," in *IEEE 16th Intl. Conf. Dependable, Autonomic and Secure Comput., 16th Intl. Conf. Pervasive Intell. and Comput., 4th Intl. Conf. Big Data Intell. Comput. Cyber Sci. and Technol. Congr. (DASC/PiCom /DataCom/CyberSciTech)*, Athens, Greece, Aug. 2018, pp. 1074–1079.
- [9] Y. Bai, B. Zeng, C. Li, and J. Zhang, "An ensemble long short-term memory neural network for hourly PM2.5 concentration forecasting," *Chemosphere*, vol. 222, pp. 286–294, May 2019.
- [10] K. Thaweephol and N. Wiwatwattana, "Long short-term memory deep neural network model for PM2.5 forecasting in the Bangkok urban area," in *17th Int. Conf. ICT and Knowl. Eng. (ICT&KE)*, Bangkok, Thailand, Nov. 2019, pp. 1–6.
- [11] P. Rong-O and N. Wiwatwattana, "The impact of data imputation and feature extraction on PM2.5 forecasting performance in Bangkok using long short-term memory neural networks," in *11th Int. Conf. Advances Inf. Technol. (IAIT202)*, Bangkok, Thailand, Jul. 2020, pp. 1–10.
- [12] "Bangkok climate," Thai Meteorological Department, Bangkok, Thailand. Accessed: Nov. 18, 2021. [Online]. Available: <https://shorturl.asia/eYHt5>
- [13] D. Berdikulov. "Dealing with missing data." MEDIUM.com. <https://medium.com/@danberdov/dealing-with-missing-data-8b71cd819501> (accessed Jan. 23, 2022).
- [14] D. Andina and D. T. Pham, *Computational intelligence: for engineering and manufacturing*. Dordrecht, Netherlands: Springer, 2007.
- [15] C. Olah. "Understanding LSTM networks." COLAH.GITHUB.io. <http://colah.github.io/posts/2015-08-Understanding-LSTMs> (accessed Feb. 10, 2022).
- [16] J. Heaton. "The number of hidden layers." HEATON RESEARCH.com. <https://www.heatonresearch.com/2017/06/01/hidden-layers.html> (accessed Feb. 10, 2022).
- [17] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of The Trade*, G. Montavon, G. Orr, K. R. Müller, Eds., Heidelberg, Germany: Springer, 2012, pp. 437–478.
- [18] G. Cheng, V. Peddinti, D. Povey, V. Manohar, S. Khudanpur, and Y. Yan, "An exploration of dropout with LSTMs," in *Interspeech 2017*, Aug. 2017, pp. 1586–1590, doi: 10.21437/Interspeech.2017-129.