



Application of machine learning and deep learning with hybrid model techniques for PM2.5 concentration forecast in Mukdahan Province using aerosol optical depth data and meteorological data

ศิวกร กุลวงศ์¹ ก้องภพ เพิ่มไพศาลสกุล¹
ธีระวุฒิ จันทะพันธ์¹ อรรถจณ์ คงทอง¹

¹โรงเรียนวิทยาศาสตร์จุฬารณราชวิทยาลัย มุกดาหาร, บางทรายใหญ่, เมืองมุกดาหาร, มุกดาหาร, 49000
Email: 05949@pccm.ac.th

บทคัดย่อ

โครงการนี้มีจุดประสงค์เพื่อ(1)ออกแบบโมเดลพยากรณ์ค่าความเข้มข้นของฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน (PM2.5) รายชั่วโมงในจังหวัดมุกดาหาร โดยประยุกต์ใช้ข้อมูลความลึกเชิงแสงของละอองลอย (AOD) ภาพถ่ายดาวเทียม และข้อมูลอุตุนิยมวิทยา ร่วมกับเทคนิคการเรียนรู้ของเครื่อง, การเรียนรู้เชิงลึก และโมเดลแบบผสม (2)เปรียบเทียบประสิทธิภาพของโมเดลการทำนายฝุ่น PM2.5 แบบเดี่ยว (Single Model) และโมเดลแบบผสม (Hybrid Model) โดยใช้ตัวชี้วัดทางสถิติและการประเมินเชิงพื้นที่-เวลา (Spatio-temporal evaluation) (3)ประยุกต์ใช้โมเดลเพื่อพัฒนาเป็นระบบพยากรณ์ความเข้มข้นของฝุ่น PM2.5 โดยมีกลุ่มเป้าหมายคือ เจ้าหน้าที่และผู้เชี่ยวชาญในหน่วยงานภาครัฐเพื่อใช้ในการเตรียมการป้องกันและรับมือกับปัญหาอย่างเร่งด่วนเครื่องมือและเทคโนโลยีที่ใช้คือ(1)แอปพลิเคชัน Visual Studio Code (2) เว็บไซต์คอมไพเลอร์ Lightning AI (3)Pre-Trained Model (4)Model และ Structure Models การศึกษาการออกแบบโมเดลพยากรณ์ฝุ่น PM2.5 พบว่ามีโมเดลที่ดีที่สุดดังนี้ โมเดลการเรียนรู้ของเครื่อง 2 โมเดลคือ Gradient Boosting Machine (GBM) และ Support Vector Regression (SVR) เป็นสองโมเดลที่ให้ผลการพยากรณ์ดีที่สุด โดย GBM มีค่า RMSE ต่ำที่สุด (2.9902) และค่า R^2 สูง (0.7616) แสดงถึงความแม่นยำและความสามารถในการอธิบายความแปรปรวนของข้อมูลได้ดี ขณะที่ SVR ให้ค่า MAE ต่ำที่สุด (2.2663) สะท้อนความคลาดเคลื่อนเฉลี่ยที่ต่ำ และมีค่า Temporal Stability ต่ำ แสดงถึงความเสถียรของการทำนายในช่วงเวลา โมเดลการเรียนรู้เชิงลึก 2 โมเดล คือ Transformer และ FNO เป็นสองโมเดลที่ให้ประสิทธิภาพดีที่สุด โดย Transformer มีค่า MAE(2.6399) และ RMSE(3.4741) ต่ำที่สุด และมีค่า R^2 (0.6780) สูงสุด ขณะที่ค่า CRPS(2.6399) ในขณะเดียวกัน FNO มีค่า Temporal Stability(0.9626) ต่ำที่สุด และเมื่อเปรียบเทียบประสิทธิภาพพบว่าโมเดลผสมมีความสามารถในการพยากรณ์ดีกว่าโมเดลเดี่ยวและเมื่อประเมินประสิทธิภาพพบว่าโมเดลด้วยวิธีการ stacking ดีที่สุดโดยมีค่าความคลาดเคลื่อนเฉลี่ยเพียง $2.9668 \mu\text{g}/\text{m}^3$ (MAE = 2.9668, RMSE = 4.1201, R^2 = 0.9319) จากการทดลองทั้งหมดสรุปได้ว่าสามารถนำไปพัฒนาเป็นระบบพยากรณ์ความเข้มข้นของฝุ่น PM2.5 ได้

คำสำคัญ:ฝุ่นละอองขนาดเล็กน้อยกว่า 2.5 ไมครอน (PM2.5), ความลึกเชิงแสงของละอองลอย (Aerosol Optical Depth: AOD), การเรียนรู้ของเครื่อง, การเรียนรู้เชิงลึก, โมเดลผสม

1. บทนำ

ฝุ่นละออง หมายถึง สารละลายแขวนลอยของอนุภาคของแข็งและหยดละอองของเหลวในอากาศ โดยสามารถถูกสูดดม เข้าไปในปอดและกระแสเลือด ทำให้เกิดอันตรายต่อสุขภาพได้ โดยฝุ่นละอองขนาดเล็กไม่เกิน 2.5 ไมครอน (PM2.5) คืออนุภาคฝุ่นละอองขนาดเล็ก ที่มีขนาดเส้นผ่านศูนย์กลางไม่เกิน 2.5 ไมโครเมตร (US.EPA, 2025)

ข้อมูลจากคลังสุขภาพของกรมสาธารณสุขระบุว่า ในช่วงเดือนตุลาคม พุทธศักราช 2566 ถึงเดือนธันวาคม พุทธศักราช 2567 มีประชาชนได้รับผลกระทบจากฝุ่น PM2.5 มากกว่า 3.32 ล้านคน และมีประชาชนในพื้นที่เสี่ยงมากกว่า 38 ล้านคน ส่งผลให้เกิดการสูญเสียโอกาสด้านสุขภาพมากกว่า 3,000 ล้านบาท และทำให้อายุขัยเฉลี่ยของประชาชนไทยลดลงถึง 1.78 ปี (Hfocus, 2025)

ความเข้มข้นของฝุ่น PM2.5 ได้รับอิทธิพลจากหลายปัจจัย ทั้งปัจจัยทางอุตุนิยมวิทยา เช่น ลม ฝน อุณหภูมิ และความสูงของชั้นบรรยากาศ รวมถึงปัจจัยด้านภูมิประเทศและกิจกรรมของมนุษย์ (Yun et al., 2018) ซึ่งไม่สามารถควบคุมได้โดยตรง ทำให้การพยากรณ์ล่วงหน้าที่มีความแม่นยำมีบทบาทสำคัญต่อการเตรียมความพร้อมของประชาชนและหน่วยงานที่เกี่ยวข้อง

จังหวัดมุกดาหารเป็นหนึ่งในพื้นที่ที่ประสบปัญหาฝุ่น PM2.5 โดยเคยปรากฏค่าฝุ่น PM2.5 สูงถึง $75.4 \mu\text{g}/\text{m}^3$ ซึ่งอยู่ในระดับสูงเกินมาตรฐานกำหนด และส่งผลกระทบต่อสุขภาพ (สำนักงานประชาสัมพันธ์จังหวัดมุกดาหาร, 2567)

แม้ในปัจจุบันจะมีงานวิจัยจำนวนมากที่นำข้อมูลอุตุนิยมวิทยาและเทคนิคการเรียนรู้ของเครื่องมาใช้ในการพยากรณ์ฝุ่น PM2.5 แต่ยังมีบางส่วนที่ยังคงขาดความครอบคลุมของการทดลองในด้านการทดสอบโมเดลที่หลากหลายและไม่มีการใช้ข้อมูลความลึกเชิงแสงละอองลอย (AOD) ซึ่งเป็นหนึ่งในตัวแปรที่มีอิทธิพลต่อค่าความเข้มข้นของฝุ่น PM2.5

2. วิธีการดำเนินงาน

ตอนที่ 1 การออกแบบโมเดลพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5 รายชั่วโมงในจังหวัดมุกดาหาร โดยประยุกต์ใช้ข้อมูลความลึกเชิงแสงของละอองลอย (AOD) ภาพถ่ายดาวเทียม และข้อมูลอุตุนิยมวิทยา ร่วมกับเทคนิคการเรียนรู้ของเครื่อง, การเรียนรู้เชิงลึก และโมเดลแบบผสม

2.1 ขั้นตอนการจัดการข้อมูล

ผู้จัดทำดำเนินการพัฒนาระบบพยากรณ์ความเข้มข้นของฝุ่น PM2.5 เริ่มต้นจากรวบรวมข้อมูลจากหน่วยงาน 1)กรมอุตุนิยมวิทยา 2)กรมควบคุมมลพิษ 3)NASA หลังจากนั้นรวบรวมข้อมูลจากทุกหน่วยงานให้เป็นไฟล์เดียวและจัดการข้อมูลให้เป็นในรูปแบบของข้อมูลอนุกรมเวลาและจัดการค่าผิดปกติ เช่น ค่าว่าง หรือช่วงเวลาที่ไม่มีข้อมูลหลังจากนั้นการปรับมาตราส่วนข้อมูล (Normalization) และบันทึกไฟล์เพื่อนำไปใช้ในขั้นตอนถัดไป

2.2 ขั้นตอนการสร้างโมเดลพยากรณ์การตกของฝน

เริ่มต้นจากการนำเข้าฐานข้อมูลข้อมูลอุตุนิยมวิทยาจากขั้นตอนก่อนหน้า จำนวน 78514 บรรทัด โดยแบ่งข้อมูลด้วยวิธีการแบ่งตามลำดับเวลา(Time-Base Split)การออกเป็นชุดข้อมูลสำหรับฝึกสอนโมเดล (Training set) 62811 บรรทัด และ เพื่อทดสอบประสิทธิภาพ 15703 บรรทัดหลังจากนั้นออกแบบโมเดลสำหรับการพยากรณ์การตกของฝนได้ 5 โมเดลดังนี้

1.Baseline model 2.XGBoost 3.LSTM 4.Transformer และ 5.Fourier Neural Operator
ต่อไปกำหนดตัวชี้วัดประสิทธิภาพของโมเดล (Metrics)
3 ตัวชี้วัดดังนี้ 1.Area Under the Receiver Operating Characteristic Curve (AUC-ROC) 2.Area Under the Precision-Recall Curve(PR-AUC) 3.Brier Score หลังจากนั้นทดสอบรันโปรแกรมและบันทึกผล

2.3 ขั้นตอนเตรียมข้อมูลสำหรับการสร้างโมเดลผสมสำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5

เริ่มจากเข้าฐานข้อมูลข้อมูลอุตุนิยมวิทยา, ความเข้มข้นของฝุ่น PM2.5 และ AOD จำนวน 28176 บรรทัดและแบ่งข้อมูลด้วยวิธีการแบ่งตามลำดับเวลา (Time-based Split) เพื่อใช้ในการฝึกสอนโมเดล จำนวน 25358 บรรทัด และ เพื่อทดสอบประสิทธิภาพ โมเดลจำนวน 2818 บรรทัด

2.4 ขั้นตอนการสร้างโมเดลด้วยเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) สำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM 2.5

การออกแบบโมเดลโมเดลเดี่ยวด้วยเทคนิคการเรียนรู้ของเครื่องสำหรับการพยากรณ์ฝุ่น PM2.5 ทั้งหมด 6 โมเดลดังนี้ 1.Linear Regression 2.Gradient Boosting Machines 3.Random Forest และ ExtraTrees 4.Support Vector Regression (SVR) 5.Linear และ Regularized Linear Models 6.Probabilistic และ Quantile models หลังจากนั้น ปรับ Hyper Parameter

2.5 ขั้นตอนการสร้างโมเดลด้วยการเรียนรู้เชิงลึก (Deep Learning) สำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM 2.5

เริ่มจากออกแบบโมเดลโมเดลเดี่ยวด้วยเทคนิคการเรียนรู้เชิงลึกสำหรับการพยากรณ์ฝุ่น PM2.5 ทั้งหมด 5 โมเดลดังนี้ 1.LSTM 2.Encoder – Decoder (seq2seq) 3.DNN 4.Temporal Transformer 5.Fourier Neural Operator และปรับ Hyper Parameter

2.6 ขั้นตอนการประเมินประสิทธิภาพของโมเดลการเรียนรู้ของเครื่องและการเรียนรู้เชิงลึก

เริ่มต้นจากการกำหนดตัวชี้วัด 7 ตัวชี้วัด ดังนี้ 1.Mean Absolute Error (MAE) 2.Root Mean Squared Error (RMSE) 3.Coefficient of Determination (R^2) 4.Mean Absolute Percentage Error (MAPE) 5.Coverage Percentage (Coverage %) 6.Continuous Ranked Probability Score (CRPS) และ 7.Temporal Stability

(ความเสถียรเชิงเวลา) และรันทุกโมเดลจากขั้นตอนก่อนหน้าและบันทึกผล

2.7 ขั้นตอนการเลือกโมเดลในการทำโมเดลผสม (Hybrid model) โดยเกณฑ์การคัดเลือกคือ รูปแบบการทำงานไม่ซ้ำกับโมเดลที่เลือกแล้ว, MAE, RMSE และ R^2

เริ่มต้นจากขั้นตอนการวางแผนขั้นตอนการทำไฮบริดโมเดลโดยมีทั้งหมด 6 โมเดล ดังนี้ 1.Average 2.Weighted intersection 3.Stacking 4.Mete-learner (Non-linear) 5.Stacking intersection หลังจากนั้นรันและทดสอบประสิทธิภาพตามตัวชี้วัดต่อไป

ตอนที่ 2 การเปรียบเทียบประสิทธิภาพของโมเดลการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5 แบบเดี่ยว (Single Model) และโมเดลแบบผสม (Hybrid Model) โดยใช้ตัวชี้วัดทางสถิติและการประเมินเชิงพื้นที่-เวลา (Spatio-temporal evaluation)

2.8 ขั้นตอนการเปรียบเทียบประสิทธิภาพโมเดลเดี่ยว (Single Model) และโมเดลผสม (Hybrid Model) เปรียบโดยใช้ตัวชี้วัด (Metric) 5 ตัวชี้วัด ดังนี้ 1.MAE 2.RMSE 3. R^2 4.MAPE 5.Temporal Stability และเปรียบเทียบในรูปแบบของตาราง

2.9 ขั้นตอนการคัดเลือกโมเดลที่ดีที่สุดในการนำไปใช้ในระบบพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5

วิธีการคัดเลือกโมเดลที่ดีที่สุดสำหรับการนำไปใช้โดยใช้วิธีการคัดเลือกแบบจำลองด้วยการตัดสินใจหลายเกณฑ์แบบถ่วงน้ำหนัก (Weighted Multi-Criteria Decision Approach) โดยใช้ตัวชี้วัด (Metric) 5 ตัวชี้วัดดังนี้ 1.MAE ให้คะแนน 40 2.RMSE ให้คะแนน 20 3. R^2 ให้คะแนน 15 4.MAPE ให้คะแนน 10 5.Temporal Stability ให้คะแนน 15 เริ่มจากทำขั้นตอนปรับมาตรฐานข้อมูล (Normalize) ให้ค่าอยู่ระหว่าง 0-1 และนำไปคำนวณตามน้ำหนักคะแนนหลังจากนั้นเลือกโมเดลที่ดีที่สุดไปใช้

ตอนที่ 3 การประยุกต์ใช้โมเดลเพื่อพัฒนาเป็นระบบพยากรณ์ความเข้มข้นของฝุ่น PM2.5

2.10 ขั้นตอนการสร้างระบบพยากรณ์ความเข้มข้นของฝุ่น PM 2.5 ภายในเขตพื้นที่จังหวัดมุกดาหาร

เริ่มต้นจากเขียนเว็บไซต์ทั้ง Front-end และ Back-end ด้วย CSS, JavaScript, HTML และ Python ตามลำดับ และวางระบบ Back-end ด้วยโมเดลจากขั้นตอนที่ 2.9 และขั้นตอนการหาที่มาของการเปลี่ยนแปลงของค่าความเข้มข้นของฝุ่น PM2.5 หลังจากนั้นเชื่อมต่อ API จากแหล่งต่าง ๆ เช่น แผนที่จาก Open street map และออกแบบหน้า UX/UI ต่างๆ โดยเชื่อมต่อตำแหน่งกับ Back-end โดยแสดงผลในรูปแบบ Geographic

2.11 ขั้นตอนการทดสอบประสิทธิภาพของระบบพยากรณ์ความเข้มข้นของฝุ่น PM 2.5 ภายในเขตพื้นที่จังหวัดมุกดาหาร

เริ่มต้นด้วยเตรียมชุดข้อมูลที่ระบบยังไม่เคยเห็นสำหรับการทดสอบประสิทธิภาพไว้จำนวน 1 ชุด สำหรับการทดสอบระยะเวลาที่ใช้ในการประมวลผลของระบบและทดสอบการประมวลผลของระบบจำนวน 30 รอบพร้อมบันทึกระยะเวลาในการประมวลผลต่อไป นำผลการประมวลผลทั้งหมดมาสรุปและแสดงผลในรูปแบบแผนภาพกล่อง หลังจากนั้นการทดสอบเปรียบเทียบค่าจากการพยากรณ์ของระบบกับค่าความเข้มข้นของฝุ่น PM2.5 จริง โดยเตรียมชุดข้อมูลสำหรับการทดสอบทดสอบการพยากรณ์ของระบบในชุดข้อมูลทดสอบที่เตรียมมาและเปรียบเทียบค่าจากการพยากรณ์ของระบบกับค่าจริงพร้อมบันทึกผลหลังจากนั้นสรุปผลและนำเสนอข้อมูลในรูปแบบของกราฟ

3.ผลและอภิปรายผล

3.1 การออกแบบโมเดลพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5 รายชั่วโมงในจังหวัดมุกดาหาร

ตารางที่ 1: ตารางเปรียบเทียบค่าประสิทธิภาพของโมเดลสำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5

Models	ROC-AUC	PR-AUC	Brier-Score
Baseline	0.9166	0.6335	0.0510
XGBoost	0.9425	0.7111	0.0449
LSTM	0.9047	0.6167	0.1135
Transformer	0.8808	0.5421	0.1046
FNO	0.8818	0.4903	0.0874
CNN	0.9018	0.6166	0.0546

ตารางที่ 2: ตารางเปรียบเทียบค่าประสิทธิภาพของโมเดลการเรียนรู้ของเครื่อง (Machine Learning) สำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5

Models	MAE	RMSE	R ²	MAPE	Coverage%	CRPS	Temporal-stability
Baseline	3.1103	4.3879	0.5227	27.0200	89.9929	3.1103	1.9172
GBM	2.3005	2.9902	0.7616	20.7387	89.9858	2.3005	0.7474
Extratrees	2.3263	3.0250	0.7560	21.0029	89.0858	2.3263	0.7528
SVR	2.2663	2.9741	0.7641	19.6814	89.9858	2.2663	0.7341
Linear	2.2977	2.9909	0.7615	20.7977	89.9858	2.2977	0.7289
Probabilistic	3.6442	2.9923	0.7612	20.5465	93.33	3.6408	0.8931

ตารางที่ 3: ตารางเปรียบเทียบค่าประสิทธิภาพของโมเดลการเรียนรู้เชิงลึก (Deep Learning) สำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5

Models	MAE	RMSE	R ²	MAPE	CRPS	Temporal-stability
Baseline	7.8899	9.0355	-1.1784	89.58	7.8899	3.0740
Seq2seq	8.0605	9.2097	-1.2629	91.47	8.0605	3.1161
DNN	2.6419	3.5042	0.6724	23.36	2.9219	1.0179
Transformer	2.6399	3.4741	0.6780	24.66	2.6399	0.9470
FNO	2.8099	3.7009	0.6346	23.21	2.8099	0.9626

รูปภาพที่ 1 ผลการทดลองการฝึกโมเดลเดี่ยว

จากตารางที่ 1 พบว่าโมเดลสำหรับการพยากรณ์ฝุ่น XGBoost ให้อำนาจในการพยากรณ์ที่สูงที่สุด จึงเลือกนำไปใช้ในระบบ โดยโมเดลพยากรณ์ฝุ่นซึ่งแบ่งออกเป็น 2 ประเภทคือ Machine learning และ Deep learning ซึ่งได้ผลการทดลองตามตารางที่ 2 และ 3 ตามลำดับ ผู้จัดทำได้เลือกโมเดลที่มีอำนาจในการพยากรณ์สูงสุด 2 อันดับแรกและมีการทำงานที่ต่างกันคือ GBM, SVR, Transformer และ FNO ในการนำไปใช้ขั้นตอนการทำโมเดลผสม

Models	MAE	RMSE	R ²	MAPE	Temporal-stability
Average	28.1020	29.3260	-1.8907	202.88	6.6879
Weighted-intersection	26.2378	27.5352	-1.5484	187.93	6.4281
Stacking	2.9668	4.1201	0.9319	16.99	1.5036
Meta-learner	2.9875	4.1732	0.9302	17.05	1.5662
Stacking-intersection	2.9694	4.1472	0.9310	17.03	1.5360

ตารางที่ 4 ตารางเปรียบเทียบค่าประสิทธิภาพของโมเดลผสม (Hybrid model) สำหรับการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5

จากการออกแบบโมเดลผสมด้วย 4 โมเดลเดี่ยวจากการทดลองก่อนหน้านี้ด้วยเทคนิค stacking พบว่าให้อำนาจในการพยากรณ์ที่สูงที่สุดเมื่อเปรียบเทียบกับโมเดลอื่นๆ โดยมีค่าความคลาดเคลื่อนเฉลี่ยเพียง 2.9668 ซึ่งต่ำกว่า 5 ซึ่งยังไม่ส่งผลต่อสุขภาพและมีแนวทางการผสมที่ถูกคัดออกสองแนวทางคือ Average และ Weighted เนื่องจากที่ผลการทดสอบประสิทธิภาพ (MAE, RMSE, MAPE) สูงผิดปกติ และมีค่า R² ต่ำ

3.2 ผลการเปรียบเทียบประสิทธิภาพของโมเดลการทำนาย PM2.5 แบบเดี่ยว (Single Model) และโมเดลแบบผสม (Hybrid Model) โดยใช้ตัวชี้วัดทางสถิติและการประเมินเชิงพื้นที่-เวลา

Models	MAE	RMSE	R ²	MAPE	Temporal-stability
GBM	2.3005	2.9902	0.7616	20.7387	0.7474
SVR	2.2663	2.9741	0.7641	19.6814	0.7341
Transformer	2.6399	3.4741	0.6780	24.66	0.9470
FNO	2.8099	3.7009	0.6346	23.21	0.9626
Stacking	2.9668	4.1201	0.9319	16.99	1.5036
Meta-learner	2.9875	4.1732	0.9302	17.05	1.5662
Stacking-intersection	2.9694	4.1472	0.9310	17.03	1.5360

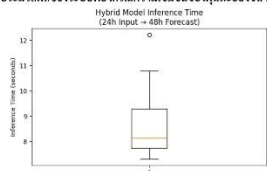
ตารางที่ 5 ตารางเปรียบเทียบค่าประสิทธิภาพโมเดลเดี่ยวและโมเดลผสม

จากการทดลองพบว่าในด้านของตัวชี้วัด MAE และ RMSE โมเดลเดี่ยวมีค่าที่ดีกว่าแต่ละตัวที่ เกิดขึ้นนั้นมีค่าที่ค่อนข้างน้อยและยังไม่สามารถนำไปใช้ได้จริง เนื่องจากค่า R², MAPE และ Temporal Stability นั้นมีค่าที่ไม่เหมาะสมแก่การนำไปใช้จริงในระบบเมื่อเทียบกับโมเดลผสม และมีช่วงค่าตัวชี้วัดที่แตกต่างกันค่อนข้างมาก จึงได้ว่าโมเดลผสมมีอำนาจในการพยากรณ์ที่สูงกว่าโมเดลเดี่ยวและสามารถนำไปใช้ในระบบต่อไปได้

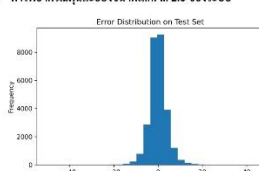
3.3 การประยุกต์ใช้โมเดลเพื่อพัฒนาเป็นระบบ

พยากรณ์ความเข้มข้นของฝุ่น PM2.5

แผนภาพที่ 1 : แผนภาพกล่องความแปรปรวนของเวลาที่ใช้ในการประมวลผลของระบบพยากรณ์ความเข้มข้นของฝุ่นละอองขนาดเล็ก



แผนภาพที่ 2 : กราฟการกระจายของค่าความคลาดเคลื่อนในการพยากรณ์ฝุ่นละอองขนาดเล็ก PM 2.5 ของระบบ



รูปภาพที่ 2 ผลการทดสอบประสิทธิภาพการทดสอบประสิทธิภาพ

จากแผนภาพที่ 1 การทดลองวัดระยะเวลาในการประมวลผลของระบบจำนวน 30 รอบ พบว่าเวลาในการประมวลผลของระบบมีค่าเฉลี่ยประมาณ 8.59 มิลลิวินาที แม้ในกรณีที่ใช้เวลานานที่สุดระบบใช้ระยะเวลาในการประมวลผลเพียง 12.22 มิลลิวินาที ซึ่ง เหมาะสมแก่การใช้งานจริง

จากแผนภาพที่ 2 การทดลองการพยากรณ์ของระบบบนชุดข้อมูลสำหรับทดสอบพบว่า มีค่า MAE เท่ากับ 2.9863 ค่า RMSE เท่ากับ 4.2594 และค่า MAPE เท่ากับ 15.88% โดยมีค่า R² เท่ากับ 0.9390 ซึ่งแสดงว่าระบบสามารถอธิบายความแปรปรวนของข้อมูลได้อย่างมีประสิทธิภาพแม้ค่า MAPE จะอยู่ที่

15.88% แต่ยังคงอยู่ในระดับที่ยอมรับได้สำหรับการพยากรณ์ค่าความเข้มข้นฝุ่น PM 2.5

สรุปผล

จากการออกแบบโมเดลพยากรณ์ค่าความเข้มข้นของฝุ่นละอองขนาดเล็ก (PM2.5) รายชั่วโมงในจังหวัดมุกดาหาร พบว่าโมเดลการเรียนรู้ของเครื่องมีโมเดลที่เหมาะสม 2 โมเดล คือ Gradient Boosting Machine (GBM) และ Support Vector Regression (SVR) เป็นสองโมเดลที่ให้ผลการทดสอบประสิทธิภาพดีที่สุด โดย GBM มีค่า RMSE ต่ำที่สุด (2.9902) และค่า R² สูง (0.7616) ขณะที่ SVR ให้ค่า MAE ต่ำที่สุด (2.2663) ส่วนโมเดลการเรียนรู้เชิงลึกที่เหมาะสมที่สุด 2 โมเดล คือ Transformer และ FNO เป็นสองโมเดลที่ให้ผลการทดสอบประสิทธิภาพดีที่สุด โดย Transformer มีค่า MAE(2.6399) และ RMSE(3.4741) ต่ำที่สุด มีค่า R²(0.6780) สูงที่สุด ส่วน FNO มีค่า Temporal Stability(0.9626) ต่ำที่สุด

จากการเปรียบเทียบประสิทธิภาพของโมเดลการพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5 แบบเดี่ยว (Single Model) และโมเดลแบบผสม (Hybrid Model) ด้วยการศึกษาโดยใช้ตัวชี้วัดทางสถิติและการประเมินเชิงพื้นที่-เวลา (Spatio-temporal evaluation) พบว่า ในด้านของตัวชี้วัด MAE และ RMSE โมเดลเดี่ยวมีผลการทดสอบที่ดีกว่าแต่ยังคงมีผลต่างที่ค่อนข้างน้อยและยังไม่สามารถนำไปใช้ได้จริง เนื่องจากค่า R², MAPE และ Temporal Stability นั้นมีค่าที่ไม่เหมาะสมแก่การนำไปใช้จริงในระบบเมื่อเทียบกับโมเดลผสมและมีช่วงค่าตัวชี้วัดที่แตกต่างกันค่อนข้างมาก ดังนั้นโมเดลผสมจึงมีอำนาจการพยากรณ์ที่สูงกว่าโมเดลเดี่ยว

จากการประยุกต์ใช้โมเดลเพื่อพัฒนาเป็นระบบพยากรณ์ความเข้มข้นของฝุ่น PM2.5 พบว่าระบบใช้ระยะเวลาในการประมวลผลเฉลี่ยเพียง 8.5902 มิลลิวินาที และเมื่อนำค่าจากการพยากรณ์มาเปรียบเทียบกับค่าจริง พบว่ามีค่าความคลาดเคลื่อน

เฉลี่ยเพียง 2.9863 ซึ่งความคลาดเคลื่อนที่ 2.9863 นั้น ไม่ส่งผลต่อร่างกาย โดยระบบพยากรณ์ความเข้มข้นฝุ่น PM2.5 ทำงานเชื่อมต่อระหว่าง Front-end และ Back-end ได้ดีและสามารถพยากรณ์ค่าความเข้มข้นของฝุ่น PM2.5 ได้อย่างรวดเร็วและแม่นยำ

กิตติกรรมประกาศ

ขอขอบคุณคุณครู ชีระวุฒิ จันทะพันธ์ และ คุณครูอรุณ คงทอง ครูที่ปรึกษาโครงการ ที่ช่วยให้คำปรึกษา ข้อชี้แนะ ตลอดจนข้อคิดเห็นต่าง ๆ และขอขอบคุณ ดร. อดี พิลาโหม นักวิจัย ศูนย์วิทยาศาสตร์บรรยากาศและอวกาศ มหาวิทยาลัยราชภัฏสกลนคร ที่ได้ช่วยหาชุดข้อมูล AOD และความรู้เกี่ยวกับการจัดการข้อมูล และขอขอบคุณกรมอุตุนิยมวิทยา กรมควบคุมมลพิษ และมหาวิทยาลัยราชภัฏสกลนคร ที่มอบชุดข้อมูลเพื่อทำโครงการเรื่องนี้

เอกสารอ้างอิง

- [1] จังหวัดมุกดาหาร รายงานสถานการณ์ฝุ่น PM2.5. สำนักงานประชาสัมพันธ์จังหวัดมุกดาหาร. (8 กุมภาพันธ์ พ.ศ.2567). <https://mukdahan.prd.go.th/>
- [2] สธ. เผยคนไทย 38 ล้านคนอยู่ในพื้นที่ฝุ่น PM2.5 เกินค่ามาตรฐาน พบเป็นกลุ่มเปราะบาง 15 ล้านคน | Hfocus.org. Hfocus เจาะลึกระบบสุขภาพ. (9 มกราคม พ.ศ.2568). <https://www.hfocus.org/content/2025/01/32758>
- [3] Karimian, H., Li, Q., Wu, C., Qi, Y., Mo, Y., Chen, G., ... & Sachdeva, S. (2019). Evaluation of different machine learning approaches to forecasting PM2. 5 mass concentrations. Aerosol and Air Quality Research, 19(6), 1400-1410.

- [4] Yun, G., Zuo, S., Dai, S., Song, X., Xu, C., Liao, Y., ... & Ren, Y. (2018). Individual and interactive influences of anthropogenic and ecological factors on forest PM2. 5 concentrations at an urban scale. Remote Sensing, 10(4), 521.
- [5] Particulate matter (PM) basics | US EPA. U.S. Environmental Protection Agency (gov). (2026, May 21). <https://www.epa.gov/pollution/particulate-matter-pm-basics>
- [6] Qiao, W., Tian, W., Tian, Y., Yang, Q., Wang, Y., & Zhang, J. (2019). The forecasting of PM2. 5 using a hybrid model based on wavelet transform and an improved deep learning algorithm. Ieee Access, 7, 142814-142825.