

การพยากรณ์ฝนระยะปัจจุบันจากข้อมูลน้ำฝนตรวจวัดด้วยเครื่องจักรการเรียนรู้ Rainfall Nowcasting Using Machine Learning from Observed Rainfall Data

ระพี ปันวิชัย¹ ดนัยภพ มะณี² และ นภาพร เปี่ยมสง่า^{3,*}

^{1,2,3} ภาควิชาวิศวกรรมทรัพยากรน้ำ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพมหานคร

*Corresponding author; E-mail address: fengnpr@ku.ac.th

บทคัดย่อ

กรุงเทพมหานครเป็นศูนย์กลางเศรษฐกิจที่มีขนาดพื้นที่ประมาณ 1,568.7 ตารางกิโลเมตร และมีสถานีตรวจวัดปริมาณน้ำฝนจำนวน 130 สถานี ที่ดำเนินการตรวจวัดทุก 5 นาที หรือแต่ละสถานีสามารถครอบคลุมพื้นที่ 12.17 ตารางกิโลเมตร ซึ่งมีความหนาแน่นที่มากและความถี่การเก็บข้อมูลสูงเนื่องจากลักษณะของฝนในกรุงเทพมหานครจะตกหนักในช่วงเริ่มต้นเมื่อฝนตก ทำให้จึงต้องมีการลดระดับน้ำในคลองเพื่อรองรับปริมาณฝนดังกล่าว ซึ่งต้องใช้เวลาในการสูบน้ำไม่น้อยกว่า 1 ชั่วโมงสำหรับความลึกน้ำประมาณ 1 เมตร การพยากรณ์ล่วงหน้าจึงมีความจำเป็น จึงมีแนวคิดในการในการใช้ข้อมูลฝนสถานีมา พยากรณ์ล่วงหน้า 120 นาทีในทุกช่วงเวลา 15 นาที ด้วยเทคนิค Random Forest ที่มีการพิสูจน์ว่าเหมาะสมในการพยากรณ์สถานะฝนตก การศึกษาในครั้งนี้ สถานีพยากรณ์ที่อยู่บริเวณกลางพื้นที่ พบว่าระยะเวลาย้อนหลังที่เหมาะสมที่สุดสำหรับการพยากรณ์ฝนล่วงหน้า 120 นาที 540 นาที ให้ความสามารถในการตรวจจับ 0.95 โดยพิจารณาเปรียบเทียบการจัดกลุ่ม 3 กลุ่ม ได้แก่ การใช้สถานีเดียว การใช้สถานีใกล้เคียง และการสุ่มเลือกสถานี ข้อมูลนำเข้าเงื่อนไขต่างกันทั้งหมด 12 กรณี โดยการพยากรณ์โดยใช้ข้อมูลซึ่งให้ค่าการตรวจจับฝน 0.95 - 0.99 ในช่วงฤดูฝนระหว่างเวลา 12:00–24:00 น.ของทั้งการเรียนรู้ในการสุ่มสถานีข้างเคียงและสถานีที่ใกล้ที่สุด และการเรียนรู้ด้วยข้อมูลจากทุกช่วงเวลา ซึ่งให้ค่าความแม่นยำ 0.99 - 0.98 ตามการพยากรณ์ล่วงหน้าในทุกช่วงเวลา 15 นาที การเรียนรู้จากสถานีใกล้เคียง 30 สถานี ให้ประสิทธิภาพในการตรวจจับฝนดีที่สุด 0.95 และการเรียนรู้แบบสุ่มสถานี ให้ความแม่นยำในการพยากรณ์สูงที่สุด ที่ 0.99

คำสำคัญ: เครื่องจักรเรียนรู้, การพยากรณ์ฝนระยะปัจจุบัน, ฝนตรวจวัด, กรุงเทพมหานคร

Abstract

Bangkok, covering an area of 1,568.7 km², operates a dense network of 130 rainfall monitoring stations, each recording data every 5 minutes and covering approximately 12.17 km². Given the city's frequent occurrence of intense rainfall at the onset of

storms, and the need for at least one hour to drain floodwater one meter deep, short-term rainfall forecasting is crucial for urban water management. This study applies the Random Forest algorithm to forecast rainfall 120 minutes ahead, at 15-minute intervals, using data from rainfall stations. The Phaya Thai station, located centrally, achieved the highest detection performance (0.95) when using 540 minutes of historical data. Twelve input scenarios were tested based on three station selection strategies: single station, nearby stations, and randomly selected stations. During the rainy season (12:00–24:00), the detection rates ranged from 0.95 to 0.99. Models trained on data across all time periods achieved forecast accuracies between 0.98 and 0.99. Using 30 nearby stations yielded the best detection capability (0.95), while random station selection produced the highest overall accuracy (0.99). These findings highlight the effectiveness of using spatially distributed rainfall data with Random Forest for high-resolution short-term rainfall forecasting in urban areas.

Keywords: Machine Learning , Rainfall Nowcasting , Observed rainfall , Bangkok

1. บทนำ

กรุงเทพมหานครมีการจัดการระบบระบายน้ำในรูปแบบคันปิดล้อม (Polder System) ที่จะมีคันกั้นน้ำจากภายนอกเข้าพื้นที่ ส่วนปริมาณน้ำภายในอาศัยเครื่องสูบน้ำระบายน้ำออกจากคลองหลัก ทั้งนี้คลองระบายน้ำบางคลองนอกจากทำหน้าที่ขนส่งน้ำยังเป็นแหล่งเก็บกักน้ำชั่วคราวโดยการลดระดับน้ำเพื่อรองรับปริมาณฝนที่ตก ซึ่งการลดระดับน้ำยังจำเป็นต้องใช้ระยะเวลาหนึ่ง เช่น คลองสามเสนนั้นต้องการลดน้ำลง 1 เมตร ต้องใช้เวลา 1 ชั่วโมง 40 นาที เป็นต้น ดังนั้นถ้าทราบว่าฝนจะตกในช่วงเวลาใดได้อย่างแน่ชัด เพื่อที่จะช่วยให้การบริหารจัดการน้ำมีประสิทธิภาพยิ่งขึ้น

การพยากรณ์ฝนในปัจจุบันสามารถดำเนินการได้หลายแนวทาง ได้แก่ วิธีการทางฟิสิกส์ ที่ต้องใช้ข้อมูลปริมาณมาก มีความซับซ้อนในการคำนวณ

สูง และเวลานาน [1] วิธีเชิงสถิติ ซึ่งมีข้อจำกัดในรูปแบบข้อมูลมีความซับซ้อนหรือไม่เป็นเชิงเส้น อาจทำให้ความแม่นยำลดลง และการคาดการณ์อาจไม่สามารถปรับตัวได้กับสภาวะอากาศที่เปลี่ยนแปลง [2] และ วิธีเชิงปัญญาประดิษฐ์

การเกิดฝนของกรุงเทพมหานครมีช่วงเวลาการตกสั้น [3] คืออยู่ในช่วง 1 – 3 ชั่วโมงเป็นส่วนใหญ่ โดยฝนจะตกสูงสุดในช่วง 15 – 60 นาทีแรกของการเกิดฝน ทั้งนี้เมื่อเป็นฝนหนักเวลาที่เกิดฝนสูงสุดจะเพิ่มขึ้น และฝนโดยส่วนใหญ่ของกรุงเทพมหานครนั้นจะเกิดอยู่ในช่วงเวลาย่ำค่ำ

กรุงเทพมหานครนั้นได้มีการติดตั้งสถานีวัดน้ำฝนอัตโนมัติมากขึ้นจนในปัจจุบันนั้นมีถึง 130 สถานี [4] ซึ่งมีความหนาแน่นค่อนข้างมาก การนำข้อมูลปริมาณฝนตรวจวัดมาเพื่อพยากรณ์สถานะฝนตกล่วงหน้า 1 – 2 ชั่วโมง จะช่วยให้การจัดการน้ำในคลองมีประสิทธิภาพดีขึ้น ในการศึกษาได้เลือกใช้ข้อมูลจากสถานีฝนฟ้า ซึ่งตั้งอยู่ใจกลางพื้นที่กรุงเทพมหานคร เพื่อเป็นตัวแทนของพื้นที่ศึกษาในการพัฒนาแบบจำลองการพยากรณ์ฝน โดยประยุกต์ใช้เทคนิคการเรียนรู้ของเครื่องจักรเรียนรู้ (Machine Learning) เพื่อให้ได้ผลการพยากรณ์ที่รวดเร็วและแม่นยำยิ่งขึ้น ซึ่งมีการพิสูจน์แล้วว่า Random forest ดีที่สุดในพื้นที่กรุงเทพมหานคร ในการพยากรณ์ล่วงหน้าได้มากกว่า 90 นาที [5]

2. พื้นที่ศึกษา

2.1 ลักษณะทั่วไปของกรุงเทพมหานคร

กรุงเทพมหานครนั้นมีพื้นที่ 1568.7 ตารางกิโลเมตร นั้นอยู่บริเวณที่ราบลุ่มแม่น้ำเจ้าพระยาตอนล่างเหนืออ่าวไทยประมาณ 25 กิโลเมตร ซึ่งได้รับอิทธิพลจากการน้ำขึ้นน้ำลงของน้ำทะเล และประกอบกับกรุงเทพมหานครนั้นมีลักษณะเป็นพื้นที่ลุ่มต่ำ ที่มีระดับพื้นดินต่ำกว่าระดับน้ำด้านนอก กรุงเทพมหานครจึงใช้ลักษณะระบบปิดล้อม (Polder System) ทำให้ต้องใช้ท่อ คลอง และสถานีสูบน้ำ เพื่อระบายปริมาณน้ำภายในพื้นที่ออกสู่มแม่น้ำเจ้าพระยา

2.2 ในการบริหารจัดการน้ำของกรุงเทพมหานคร จะมีการใช้คลองเป็นระบบขนส่งทางน้ำ และ แหล่งเก็บกักน้ำชั่วคราว โดยการลดระดับน้ำในคลองก่อนฝนตกเพื่อรองรับปริมาณฝนในช่วงแรก ซึ่งการลดระดับน้ำ ดังกล่าวนั้นต้องใช้ระยะเวลา 1 – 2 ชั่วโมง เช่น คลองสามเสนมีความยาว 11.636 กิโลเมตร ความกว้าง 5-10 เมตร เมื่อทำการสูบน้ำเพื่อลดระดับน้ำในคลองลง 2 เมตรต้องใช้เวลาประมาณ 3 ชั่วโมง

2.3 ลักษณะภูมิอากาศของกรุงเทพมหานคร

กรุงเทพมหานครอยู่ภายใต้อิทธิพลของมรสุม 2 ชนิด คือมรสุมตะวันออกเฉียงเหนือ ซึ่งพัดจากทิศตะวันออกเฉียงเหนือปกคลุมในช่วงฤดูหนาว ทำให้มีอากาศหนาวเย็นและแห้งแล้ง กับมรสุมตะวันตกเฉียงใต้ ซึ่งพัดจากทิศตะวันตกเฉียงใต้ปกคลุมในช่วงฤดูฝน ทำให้อากาศชุ่มชื้นและมีฝนตกทั่วไป โดยแบ่งฤดูกาลของกรุงเทพมหานคร ออกเป็น 3 ฤดูดังนี้ [6]

ฤดูหนาวเป็นช่วงที่ลมมรสุมตะวันออกเฉียงเหนือจากประเทศจีนมีกำลังแรงขึ้นจนแผ่ลงมาปกคลุมประเทศไทยในช่วงตั้งแต่มกราคมถึงกลางเดือนกุมภาพันธ์ส่งผลให้อากาศเย็นและแห้ง

ฤดูร้อนอยู่ในช่วงประมาณกลางเดือนกุมภาพันธ์ถึงกลางเดือนพฤษภาคม เมื่อมรสุมตะวันออกเฉียงเหนืออ่อนกำลังลงจนไม่สามารถเข้าปกคลุมพื้นที่ได้ โดยในช่วงระยะเวลานี้เป็นช่วงว่างของฤดูมรสุม จะมีหย่อมความกดอากาศต่ำเนื่องจากความร้อนเข้าปกคลุม และลมที่พัดปกคลุมบริเวณดังกล่าวเป็นลมใต้และตะวันออกเฉียงใต้ ทำให้อากาศร้อนอบอ้าวทั่วไป โดยมีอากาศร้อนที่สุดในเดือนเมษายน

ส่วนฤดูฝนเป็นช่วงที่ฤดูมรสุมตะวันตกเฉียงใต้ ซึ่งเป็นลมที่พัดจากมหาสมุทรอินเดียนำน้ำฝนและความชุ่มชื้นเข้าสู่ประเทศไทย ประกอบกับร่องความกดอากาศต่ำที่พัดผ่านบริเวณภาคใต้ของประเทศไทยจะเลื่อนขึ้นมาพัดผ่านบริเวณภาคกลางและภาคเหนือเป็นลำดับทำให้มีฝนตกชุกขึ้นตั้งแต่กลางเดือนพฤษภาคมเป็นต้นไปถึงกลางเดือนตุลาคม โดยเดือนกันยายนเป็นช่วงที่มีความชื้นสูงสุดส่งผลให้มีฝนตกชุกมากที่สุดในรอบปี

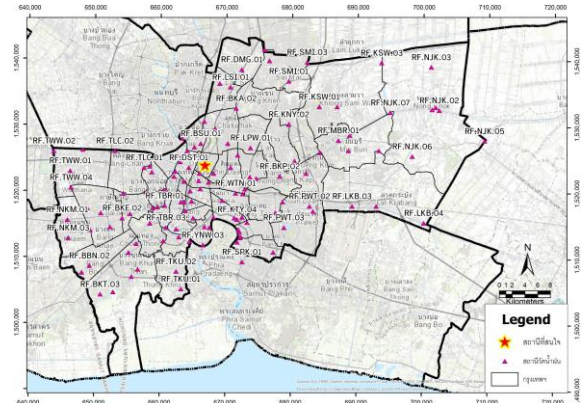
2.4 ลักษณะของฝนในพื้นที่กรุงเทพมหานคร

ส่วนใหญ่กรุงเทพมหานครมีช่วงเวลาการเกิดฝนประมาณ 1-3 ชั่วโมง ซึ่งจะเกิดฝนสูงสุดในช่วง 30 นาทีแรกถึง 55.7 เปอร์เซ็นต์ [3] ส่วนช่วงเวลาของการเกิดฝนตกบ่อยครั้ง คือ เวลาประมาณ 17:00 – 20:00 นาฬิกา

ปริมาณฝนเฉลี่ยของพื้นที่กรุงเทพในรอบ 30 ปี นั้น อยู่ที่ 1689.7 มม. และปริมาณฝนเฉลี่ยในแต่ละเดือนนั้นอยู่ที่ 25.2 , 26.5 , 51.1 , 97.3 , 215.1 , 197.2 , 190.4 , 229.9 , 338.4 , 285.5 , 46.5 และ 13.4 มิลลิเมตร ตามลำดับ [6]

2.5 สถานีวัดน้ำฝนในเครือข่ายของกรุงเทพมหานคร

กรุงเทพมหานครนั้นมีสถานีตรวจวัดฝนในเครือข่ายจำนวน 130 สถานี [7] ที่มีการตรวจวัดทุก 5 นาที โดยแบ่งเป็นฝั่งกรุงเทพตะวันออกจำนวน 88 สถานี และกรุงเทพตะวันตก 42 สถานี ดังแสดงในรูปที่ 1 ซึ่งแสดงตำแหน่งของสถานีวัดน้ำฝนของกรุงเทพมหานคร



รูปที่ 1 แผนที่แสดงตำแหน่งสถานีวัดน้ำฝนของกรุงเทพมหานคร

3. ทฤษฎีที่เกี่ยวข้อง

3.1 การพยากรณ์อากาศ

การคาดหมายสภาพลมฟ้าอากาศในอนาคต เช่น ฝน อุณหภูมิ เมฆ หมอก คลื่นลม รวมทั้งภัยธรรมชาติที่รุนแรงและไม่รุนแรง จะพยากรณ์นั้น ต้อง ความเข้าใจในปรากฏการณ์ต่างๆสภาวะอากาศในปัจจุบันและความสามารถในการผนวกความ เข้าใจและสภาวะอากาศปัจจุบัน ซึ่งแบ่ง ชนิดการพยากรณ์ได้ตามช่วงเวลา ดังแสดงตารางที่ 1

ตารางที่ 1 ชนิดของการพยากรณ์ตามช่วงเวลา

ชนิดของการพยากรณ์	ช่วงเวลา
การพยากรณ์ปัจจุบัน	ไม่เกิน 3 ชั่วโมง
การพยากรณ์ระยะสั้นมาก	ไม่เกิน 12 ชั่วโมง
การพยากรณ์ระยะสั้น	ไม่เกิน 3 วัน
การพยากรณ์ระยะปานกลาง	3 – 10 วัน
การพยากรณ์ระยะยาว	10 – 30 วัน
การพยากรณ์ระยะนาน	30 วัน เป็นต้นไป

3.2 การพยากรณ์ฝนด้วยข้อมูลต่างๆ

3.2.1 การพยากรณ์โดยใช้ข้อมูลภาพดาวเทียมตรวจอากาศ

งานวิจัยในกลุ่มของการพยากรณ์ฝนโดยใช้ภาพดาวเทียมนั้นจะมี ลักษณะที่ศึกษาการเคลื่อนตัวของกลุ่มเมฆที่ต่อเนื่องไป โดยวิธีการต่างๆ เช่น Picture Fuzzy Clustering , Recurrent Neural Network, PFC-STAR [8,9,10] แต่ในด้านการแปรผลภาพเป็นข้อมูลผิวนั้นยังมีน้อย เนื่องจากข้อมูลเมฆทุกก้อนไม่ได้ทำให้เกิดฝน และแม้แต่ในเมฆฝนฟ้าคะนองก็ไม่ได้เกิดเป็นฝนตกทุกก้อนหรือตกทั่วพื้นที่ที่เมฆปกคลุม

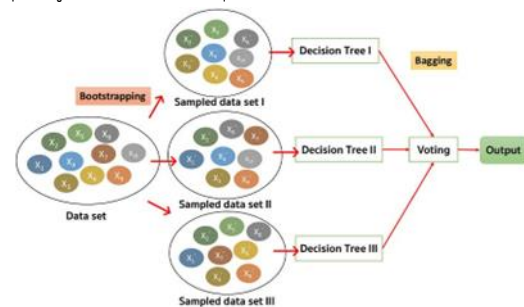
3.2.2 การพยากรณ์โดยใช้ข้อมูลเรดาร์ตรวจอากาศ

เรดาร์ตรวจอากาศใช้การสะท้อนกลับของเรดาร์มาแปลงเป็นปริมาณฝนซึ่งต้องหาความสัมพันธ์ระหว่างปริมาณฝนกับค่าการสะท้อนกลับ ทั้งนี้มีความผิดพลาดที่สามารถเกิดขึ้นได้ แม้ว่าเรดาร์ตรวจอากาศภาคพื้นดินจะให้ความละเอียดเชิงพื้นที่ประมาณ 1 กิโลเมตร และเชิงเวลาประมาณ 5 นาที แต่การสอบเทียบเรดาร์ทำได้ยาก [11] ส่วนการนำไปพยากรณ์มีหลายเทคนิค ซึ่งมีพื้นฐานเดียวกันที่มาจาก Optical Flow [12,13,14] ซึ่งวิธีการส่วนใหญ่นั้นเป็นการมองการเคลื่อนตัวของกลุ่มเมฆ

3.3 Random Forest

เป็นอัลกอริทึมการเรียนรู้สำหรับจัดงานจัดกลุ่มในกลุ่มของ Ensemble Learning ที่พัฒนาโดย Adele Cutler [15] ซึ่งเป็นการสร้างแบบจำลอง Decision Tree ตั้งแต่ 10 เป็นต้นไป แต่ละแบบจำลองจะได้ข้อมูลสำหรับการฝึกฝนเป็นบางส่วนของชุดข้อมูลทั้งหมดที่แตกต่างกัน ในขั้นตอนของการทำการพยากรณ์แต่ละ Decision Tree ก็จะทำพยากรณ์ด้วยตัวมันเอง แล้วนำเอาการตัดสินใจของแบบจำลองเหล่านั้นมาโหวตกันว่าผลการพยากรณ์ฝน (Class) ไດถูกเลือกมากที่สุดในการณ์ (Classification) ในของ แต่ละ Decision Tree

Bootstrap Sampling ในการสร้าง Random Forest นั้น จะใช้การสุ่มตัวอย่างข้อมูลที่มีทดแทน (Bootstrap) เพื่อสร้างชุดข้อมูลย่อย ๆ หลายชุด (Bootstrap Samples) ชุดข้อมูลนี้จะถูกนำไปใช้ในการสร้าง Decision Tree แต่ละต้น โดยลำดับการทำงาน เริ่มจาก การเลือก Feature แบบสุ่ม ขณะที่สร้างแต่ละ Node ใน Decision Tree จะมีการสุ่มเลือก Feature บางส่วนขึ้นมาแทนที่จะใช้ Feature ทั้งหมดในการแยกข้อมูล ทำให้ Decision Trees แต่ละต้นมีลักษณะเฉพาะตัวและลดการพึ่งพา Feature ใด Feature หนึ่งเกินไป ต่อมาเป็นการสร้าง Decision Tree แต่ละชุดข้อมูลย่อยที่ได้จากขั้นตอน Bootstrap Sampling จะถูกนำไปสร้าง Decision Tree โดยใช้การตัดสินใจตาม Feature ที่สุ่มเลือกในแต่ละ Node จนกว่าจะได้ Decision Tree ครอบคลุมจำนวนที่กำหนด (เช่น 100, 500 ต้น) สุดท้ายจะเป็นการรวมผลลัพธ์จากหลายต้นนี้ คือ การนำผลลัพธ์จากแต่ละ Decision Tree มาทำการโหวต (Voting) โดยผลลัพธ์ที่มีจำนวนการโหวต สูงที่สุดจะถูกเลือกเป็นผลลัพธ์สุดท้าย



รูปที่ 2 แผนผังการทำงานของ Random Forest

3.4 การทดสอบประสิทธิภาพ

การทดสอบประสิทธิภาพสำหรับการพยากรณ์อากาศ เป็นการทดสอบประสิทธิภาพของการพยากรณ์ฝนเชิงคุณภาพ คือ ประสิทธิภาพที่บอกว่าแบบจำลองพยากรณ์ว่าฝนจะตกหรือไม่ มีความถูกต้องแม่นยำมากน้อยขนาดไหน โดย True Positive (TP) คือการพยากรณ์ออกมาเป็นฝนแล้วเกิดขึ้นจริง True Negative (TN) คือ การพยากรณ์ออกมาเป็นฝนไม่ตรงกับความเป็นจริง False Positive (FP) คือ การพยากรณ์ออกมาเป็นฝนตกแต่ความเป็นจริงนั้นไม่เกิดขึ้นจริง False Negative (FN) คือ การพยากรณ์ออกมาเป็นฝนไม่ตกแต่ความเป็นจริงนั้นฝนตก ซึ่งแสดงในตารางที่ 2

ตารางที่ 2 confusion matrix ของการพยากรณ์ฝน

		ผลการพยากรณ์	
		ฝนตก (Positive)	ไม่ตก(Negative)
ผลจริง	ฝนตก (Positive)	TP เกิดขึ้นจริง	FN ฝนไม่ตกแต่บอกตก
	ไม่ตก (Negative)	FP ฝนตกแต่บอกไม่ตก	TN ฝนไม่ตกจริง

Precision หรือ False Alarm Ratio (FAR) หมายถึง ความแม่นยำของการพยากรณ์ออกมาเป็นฝนตกว่ามีความถูกต้องมากแค่ไหน โดยสามารถคำนวณได้จากสมการที่ (1) และ (2)

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

โดยค่าที่ใกล้เคียง 1 หมายถึง แบบจำลองมีการแจ้งเตือนฝนตกอย่างแม่นยำสูง มีอัตราการแจ้งเตือนผิดพลาด (False Alarm) ต่ำ

Recall หมายถึง ความสามารถในการตรวจจับฝนตกออกมาได้เพียงใดนั้น สามารถคำนวณได้จากสูตร

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

โดยค่าที่ใกล้เคียง 1 หมายถึง แบบจำลองสามารถตรวจจับเหตุการณ์ฝนตกได้เกือบครบทุกเหตุการณ์จริง ลดโอกาสการพลาดการแจ้งเตือน

4. วิธีการศึกษา

4.1 การตรวจสอบความสมบูรณ์ของข้อมูลฝน

ในการศึกษานี้ใช้ข้อมูลฝนราย 5 นาที จำนวน 130 สถานี ทั้งนี้ต้องมีการตรวจสอบข้อมูลฝนตรวจวัดการนำมาใช้ โดยจะตรวจสอบความสมบูรณ์ของข้อมูลตั้งแต่ปี 2565 -2567 โดยข้อมูล 1 สถานีใน 1 ปี นั้นจะมีข้อมูลฝนราย 15 นาที 35,040 ค่า และการตรวจสอบข้อมูลฝนด้วยวิธีการทางสถิติของฝนรายปี การตรวจสอบฝนรายเดือนเพื่อหาค่าที่ผิดปกติของปริมาณฝนในเดือนต่างๆ และการตรวจสอบปริมาณฝนรายวันว่ามีมากเกินปกติ และการตรวจสอบปริมาณฝนราย 15 นาทีสำหรับการใช้งานที่เกิดฝนมากกว่า ฝนออกแบบในรอบ 100 ปี ซึ่งมีอยู่ที่ 64.73 มิลลิเมตร [5] จึงจะตัดเหตุการณ์ฝนมีมากกว่า 65 มิลลิเมตรใน 15 นาทีออกจากข้อมูล

4.2 สร้างการเรียนรู้

4.2.1 การจัดการกับสถานีนำเข้า

การจัดการกับสถานีนำเข้านั้นเพื่อหาผลกระทบของลักษณะการนำเข้าข้อมูลที่มีลักษณะที่แตกต่างกันโดยแบ่งออกเป็น 3 ลักษณะในการนำเข้า ซึ่งประกอบไปด้วย การนำเข้าข้อมูลเพียงสถานีที่สนใจ การนำเข้าสถานีที่สนใจกับสถานีข้างเคียงที่ใกล้ที่สุดจำนวน 30 สถานี และ การสุ่มหยิบสถานีข้างเคียงในระยะ 10 20 และ 30 กิโลเมตร อย่างละ 10 สถานี จึงทำให้แยกออกเป็น 3 ลักษณะในการนำเข้า ได้แก่ สถานีที่สนใจ 30 สถานีที่ใกล้ที่สุด และสุ่มหยิบสถานีนำเข้า ดังแสดงในตารางที่ 3

4.2.2 การจัดการช่วงเวลานำเข้าแบบจำลอง

เนื่องจากจำนวนเหตุการณ์ที่เกิดฝนในแต่ละช่วงการเรียนรู้นั้นมีค่าแตกต่างกันไปในแต่ละเดือนและช่วงเวลาในแต่ละวัน เพื่อเป็นการศึกษาผลกระทบของช่วงเวลาในการนำเข้าข้อมูลจึงได้แยกช่วงเวลาออกเป็น การ

เรียนรู้ข้อมูลทั้งหมด ข้อมูลเฉพาะช่วงฤดูฝน ข้อมูลเที่ยงวันถึงเที่ยงคืน และข้อมูลฤดูฝนในช่วงเที่ยงวันถึงเที่ยงคืน ดังแสดงในตารางที่ 3

ตารางที่ 3 สรุปรูปแบบการเรียนรู้

กรณี	สถานีนำเข้า	ช่วงเวลานำเข้า
A	สถานีที่สนใจ	ข้อมูลทั้งหมด
B		ข้อมูลเฉพาะช่วงฤดูฝน
C		ข้อมูลเที่ยงวันถึงเที่ยงคืน
D		ข้อมูลฤดูฝนในช่วงเที่ยงวันถึงเที่ยงคืน
E	สถานีที่สนใจ และ 30 สถานีที่ใกล้ที่สุด	ข้อมูลทั้งหมด
F		ข้อมูลเฉพาะช่วงฤดูฝน
G		ข้อมูลเที่ยงวันถึงเที่ยงคืน
H		ข้อมูลฤดูฝนในช่วงเที่ยงวันถึงเที่ยงคืน
I	สถานีที่สนใจ และ สุ่มหยิบสถานีนำเข้า	ข้อมูลทั้งหมด
J		ข้อมูลเฉพาะช่วงฤดูฝน
K		ข้อมูลเที่ยงวันถึงเที่ยงคืน
L		ข้อมูลฤดูฝนในช่วงเที่ยงวันถึงเที่ยงคืน

4.3 การจัดการกับความไม่สมดุลของผลเฉลย

เนื่องจากความไม่สมดุลระหว่างอัตราส่วนระหว่างเหตุการณ์ ฝนตก และไม่ตก ส่งผลให้การเรียนรู้ได้มีแนวโน้มลำเอียงไปทางคลาสที่มีมากกว่า (คลาสไม่ตกฝน) และอาจทำให้โมเดลจำแนกคลาสที่มีจำนวนน้อยกว่า (คลาสฝนตก) ผิดพลาดได้ ดังนั้นจึงได้ใช้ Synthetic Minority Oversampling Technique (SMOTE) ในการทำ Over sampling โดยการทำให้ข้อมูลที่มีขนาดเล็กให้ใหญ่ขึ้น [16,17,18]

4.4 การทดสอบประสิทธิภาพการพยากรณ์ล่วงหน้า

เพื่อประเมินภาพรวมของประสิทธิภาพการพยากรณ์ โดยในงานวิจัยนี้ได้กำหนด เกณฑ์ขั้นต่ำ สำหรับการยอมรับแบบจำลอง โดยให้ค่า Precision ≥ 0.7 และ Recall ≥ 0.7 หากแบบจำลองใดไม่สามารถทำได้ตามเกณฑ์ดังกล่าว จะถือว่าไม่ผ่านเกณฑ์คุณภาพที่กำหนด และไม่นำมาพิจารณาต่อในขั้นตอนการเลือกแบบจำลองที่ดีที่สุด

4.4.1 ผลกระทบระยะเวลาย้อนหลังที่มีผลต่อการพยากรณ์

การทดสอบนี้มีจุดประสงค์เพื่อหาความสัมพันธ์ระหว่างช่วงเวลาย้อนหลังที่ดีที่สุดโดยเริ่มตั้งแต่ระยะเวลาย้อนหลัง 15 นาทีจนถึง 540 นาที เพื่อพยากรณ์ล่วงหน้าทีระยะเวลาระยะเวลา 120 นาที ซึ่งจะวัดผลโดยใช้ค่า ความแม่นยำและความสามารถในการตรวจจับฝนตก โดยเริ่มจากสถานี สถานีพญาไท ซึ่งอยู่บริเวณใจกลางของกรุงเทพมหานครและมีสถานีล้อมรอบครอบคลุมทุกเงื่อนไขดังแสดงในรูปที่ 1

4.4.2 ประสิทธิภาพในการพยากรณ์ล่วงหน้า

การทดสอบประสิทธิภาพในการพยากรณ์ล่วงหน้า 15 30 45 60 75 90 105 และ 120 นาที เพื่อหาแบบจำลองที่ดีที่สุดสำหรับการพยากรณ์ล่วงหน้าให้ได้ 120 นาที จากระยะเวลาย้อนหลังที่ดีที่สุด และผลกระทบของการ

พยากรณ์ที่ระยะเวลานั้นเพิ่มขึ้นมีผลต่อการพยากรณ์อย่างไร ในการตรวจจับฝนตกความแม่นยำของการพยากรณ์

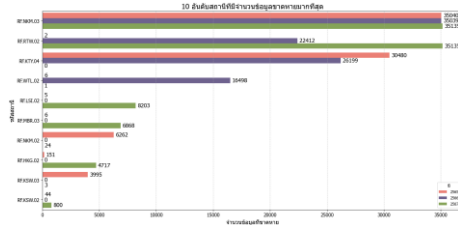
4.4.3 ผลกระทบเนื่องจากตำแหน่งที่ตั้งของสถานี

เพื่อเปรียบเทียบตำแหน่งของสถานีในการนำเข้ากับข้อมูลในการนำเข้านั้นมีลักษณะต่างๆ ในสถานีนำเข้าที่แตกต่างกัน เพื่อหาลักษณะในการนำเข้าเนื่องจากการเลือกสถานีข้างเคียงว่าลักษณะเป็นอย่างไรสำหรับการพยากรณ์ล่วงหน้า 120 นาที ที่ให้ความแม่นยำสูงสุดและความสามารถในการตรวจจับที่ดีที่สุด

5. ผลการศึกษา

5.1 การตรวจสอบความสมบูรณ์ของข้อมูล

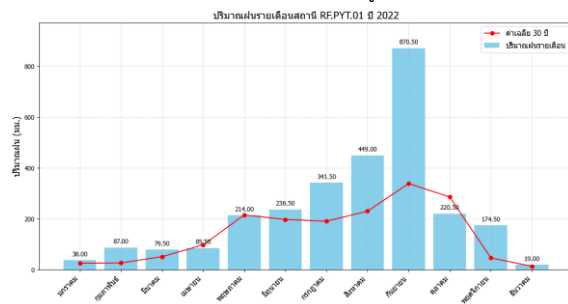
จากปริมาณฝนที่ 130 สถานี ได้ตรวจสอบความถูกต้องแล้วพบ ในปี 2566 2567 และ 2568 มีจำนวนข้อมูลราย 15 นาที ขาดหายมากที่สุดจำนวน 10 สถานีที่แสดงในรูปที่ 3 จะเห็นได้ว่า ที่สถานีฝน NKM.03 RTW.02 และ KTY.04 นั้นมีปริมาณข้อมูลที่ขาดหายในแต่ละปีเกินกว่ากึ่งหนึ่งของข้อมูลทั้งหมด จึงตัดสถานีดังกล่าวทิ้งไม่นำมาใช้



รูปที่ 3 จำนวนข้อมูลขาดหายในแต่ละปีโดยแสดง 20 สถานีที่มากที่สุด

จากข้อมูลฝนสะสมรายปีจะเห็นได้ว่าในปี 2565 นั้นจะมีการกระจายตัวที่กว้างกว่าปีอื่นๆ ดังแสดงในตารางที่ 4 มีปริมาณฝนสูงสุดถึง 2,815 มิลลิเมตร ต่ำสุด 1170 มิลลิเมตร

ปริมาณฝนรายเดือนของสถานีสำนักงานเขตพญาไทซึ่งมีฝนที่สูงสุดในปี 2565 นั้นมีปริมาณฝนที่สูงกว่าค่าเฉลี่ย 30 ปีในเดือนสิงหาคมและกันยายนซึ่งมีปริมาณเกิน 2 เทา ดังแสดงในรูปที่ 4



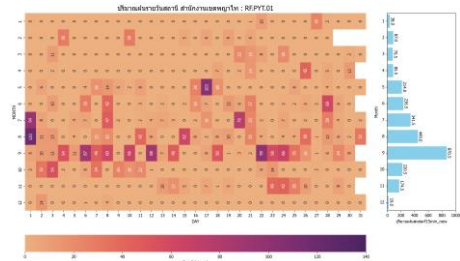
รูปที่ 4 ปริมาณฝนสะสมรายเดือนของสถานีพญาไท

ตารางที่ 4 ข้อมูลทางสถิติปริมาณฝนสะสมรายปี

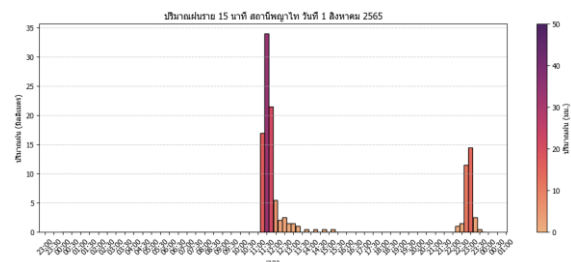
ปี	ค่าเฉลี่ย	Std	ต่ำสุด	สูงสุด
2565	1889.45	379.18	1170.0	2815.5
2566	1183.83	263.13	758.5	1964.5
2567	1380.98	259.36	754.5	1961.5

ปริมาณฝนรายวันของสถานีสำนักงานเขตพญาไทจะพบได้ว่าปริมาณฝนในแต่ดั่งกล่าวนั้นมีปริมาณฝนสะสมสูงในเดือนกันยายนนั้นเกิดจากปริมาณฝนในเดือนนั้นมีฝนตกในหลายวัน ซึ่งไม่มีความผิดปกติ ดังแสดงในรูปที่

แต่จะเห็นว่าวันที่ 1 สิงหาคมนั้นมีความผิดปกติที่มีปริมาณสูงจึงได้ดูปริมาณฝนในวันดังกล่าวนั้นมีปริมาณสูงถึง 120 มิลลิเมตร เมื่อตรวจสอบดูในวันดังกล่าวแล้วไม่พบความผิดปกติแต่อย่างใด ซึ่งแสดงในรูปที่ 7 ซึ่งจะมีการตรวจสอบแบบนี้ในทุกสถานี



รูปที่ 5 ปริมาณฝนสะสมรายวันสถานี สำนักงานเขตพญาไท



รูปที่ 6 ปริมาณฝนสะสมราย 15 นาทีของสถานีพญาไท วันที่ 1 สิงหาคม 2565

5.2 การทดสอบผลกระทบระยะเวลาย้อนหลังที่มีผลต่อการพยากรณ์

ในการพยากรณ์ครั้งนี้ เป็นการพยากรณ์ฝนของสถานีสำนักงานเขตพญาไท ซึ่งตำแหน่งของสถานีอยู่บริเวณใจกลางของพื้นที่กรุงเทพมหานคร และมีปริมาณของสถานีข้างเคียงที่ใกล้เคียงที่ใกล้เคียงและครอบคลุมเงื่อนไขต่างๆ ครบถ้วนที่สุดในการทดสอบ การจัดกลุ่มทดสอบเป็น 3 กลุ่มตามตารางที่ 3 ซึ่งได้ผลดังนี้

กลุ่มที่ 1 สถานีนำเข้า คือ สถานีที่สนใจ (กรณี A B C D)

ประสิทธิภาพในการหาฝนของการเรียนรู้ในกลุ่มที่ 1 ที่เป็นการพยากรณ์โดยใช้การนำเข้าสถานีที่สนใจ จะเห็นได้ว่าระยะเวลาย้อนหลังที่เพิ่มขึ้นนั้นส่งผลต่อการตรวจจับฝนที่เพิ่มขึ้นเมื่อระยะเวลาย้อนหลัง เมื่อระยะเวลาย้อนหลังเริ่มต้นที่ 15 นาที จะให้ความถูกต้อง ในกรณี A B C D อยู่ที่ 0.27 0.28 0.33 และ 0.31 ตามลำดับ และระยะเวลาย้อนหลังที่มาก

ถึง 540 นาที นั้น ให้ประสิทธิภาพในการตรวจจับอยู่ที่ 0.50 0.46 0.54 และ 0.49 ตามลำดับ โดย การเรียนรู้ที่ดีที่สุดในการตรวจจับฝนในการ นำเข้าเพียงสถานีเดียวนั้น จะเป็นของ กรณี C ดังแสดงในรูปที่ 7

ความแม่นยำในการพยากรณ์ของการเรียนรู้ในกลุ่มนี้นั้นมีสูงมากกว่า 0.9 หรือ 90% โดยมีความแม่นยำสูงในการพยากรณ์ในกรณี B ซึ่ง ประสิทธิภาพสูงที่สุดในช่วง ระยะเวลาย้อนหลัง 480 นาที ดังแสดงในรูปที่ 10

กลุ่มที่ 2 สถานีที่สนใจกับ 30 สถานีที่ใกล้ที่สุด(กรณี E F G H)

ประสิทธิภาพในการตรวจจับฝนของการเรียนรู้ในกลุ่มนี้นั้นมีลักษณะที่ เหมือนกับการเรียนรู้ในกลุ่มแรกว่าการเพิ่มระยะเวลาย้อนหลังนั้นส่งผลต่อ การเพิ่มขึ้นขึ้นของประสิทธิภาพในการตรวจจับ ดังนี้ โดยเมื่อระยะเวลาค่า เริ่มต้นอยู่ที่ 0.64 0.67 0.68 และ 0.65 ตามลำดับ E F G H และเมื่อ ระยะเวลาย้อนหลังเพิ่มขึ้นถึงที่ระยะเวลาย้อนหลัง 540 นาที นั้นจะอยู่ที่ 0.91 0.94 0.92 และ 0.95 ซึ่งความสามารถในการตรวจจับที่ดีที่สุดอยู่ที่ กรณี H ดังแสดงในรูปที่ 9

ความแม่นยำของการพยากรณ์ฝนโดยข้อมูลนำเข้ากลุ่มที่ 2 นั้นจะให้ ประสิทธิภาพสูงสุดคือการพยากรณ์ล่วงหน้าโดยใช้ข้อมูลย้อนหลัง 180 นาทีเพื่อให้ได้การพยากรณ์ที่สามารถบอกฝนตกได้ถูกต้องถึง 99% หรือ

0.99 และกรณีที่แม่นยำต่ำสุดในกลุ่มนั้นให้ความแม่นยำสูงถึง 0.97 หรือ 97% ดังแสดงในรูปที่ 10

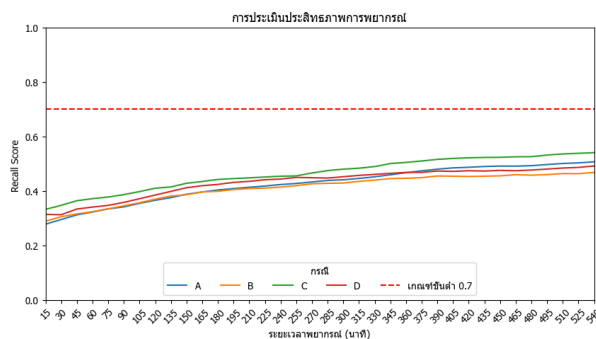
กลุ่มที่ 3 สถานีนำเข้า คือ สถานีที่สนใจกับสี่หมอบสถานีนำเข้า (กรณี I J K L)

ประสิทธิภาพในการตรวจจับฝนของการเรียนรู้ในกลุ่มนี้นั้นจะเห็นได้ว่า ความสามารถในการตรวจจับมีแนวโน้มเพิ่มขึ้นตามระยะเวลาย้อนหลัง เพิ่มขึ้นนั้นจะสูงที่สุดในกรณี L ซึ่งเป็นการเรียนรู้ในช่วงระยะเวลาย้อนหลัง 240 นาทีเป็นต้นไป

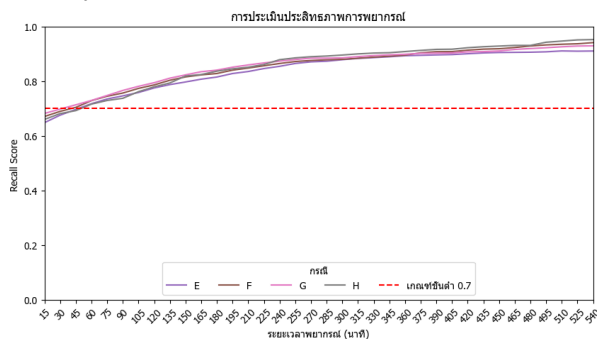
ประสิทธิภาพในการตรวจจับเริ่มต้นนั้นจะอยู่ที่ 0.65 0.67 0.68 และ 0.65 ตามลำดับ แต่เมื่อระยะเวลาย้อนหลังเพิ่มขึ้นถึง 540 นาทีแล้วจะเห็น ถึงประสิทธิภาพในการตรวจจับที่เพิ่มขึ้นจนให้ประสิทธิภาพอยู่ที่ 0.91 0.94 0.92 และ 0.95 ตามลำดับ ดังแสดงในรูปที่ 14

ความแม่นยำสูงสุดของการพยากรณ์ในกลุ่มการเรียนรู้ที่ กรณี I ที่ให้ ความแม่นยำสูงถึง 0.99 หรือ 99% และ กรณี L ที่มีความแม่นยำต่ำสุด ในกลุ่มนี้นั้นให้ความแม่นยำสูงถึง 0.97 หรือ 97% ดังแสดงในรูปที่ 15

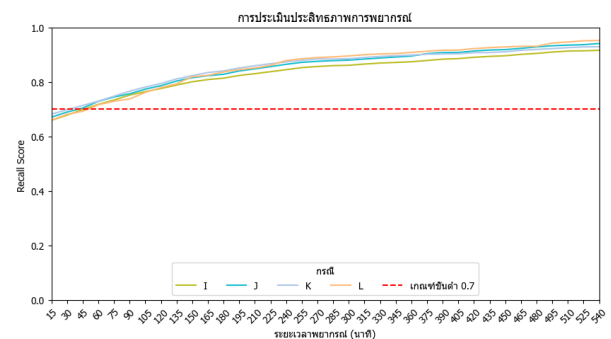
ดังนั้นระยะเวลาย้อนหลังในการพยากรณ์ที่ 540 นาที จะดีที่สุดในการ พยากรณ์ล่วงหน้า 120 นาทีในทั้ง 3 กลุ่มการเรียนรู้



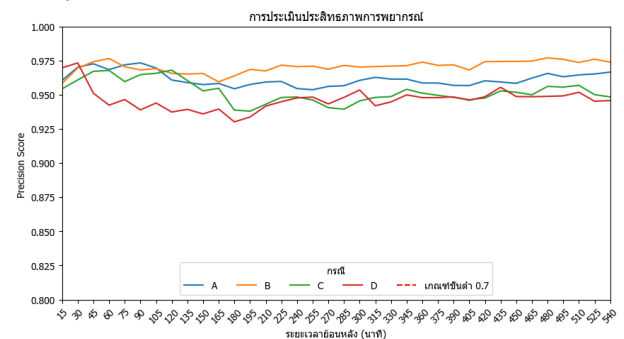
รูปที่ 7 ความสามารถในการตรวจจับ Recall กรณี A B C D



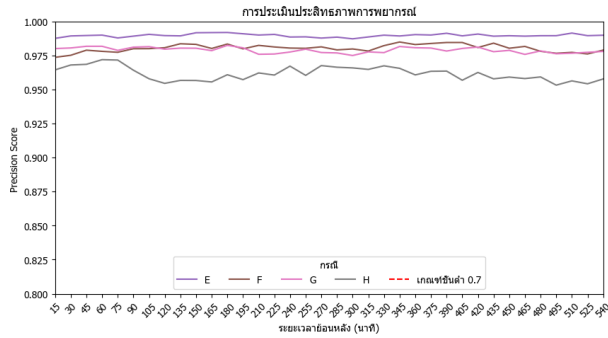
รูปที่ 8 ความสามารถในการตรวจจับ Recall กรณี E F G H



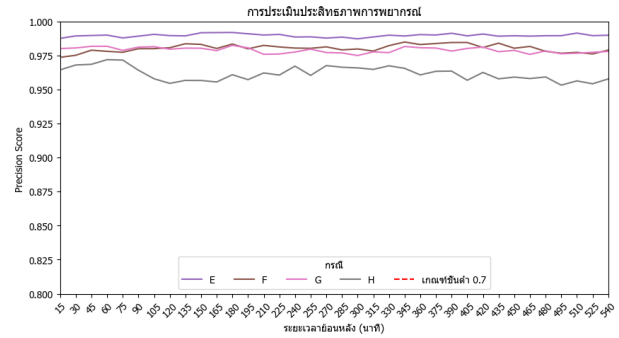
รูปที่ 9 ความสามารถในการตรวจจับ Recall กรณี I J K L



รูปที่ 10 ความสามารถในการตรวจจับ Precision กรณี A B C D



รูปที่ 11 ความสามารถในการตรวจจับ Precision E F G H



รูปที่ 12 ความสามารถในการตรวจจับ Precision I J K L

5.3 การทดสอบประสิทธิภาพในการพยากรณ์ล่วงหน้า

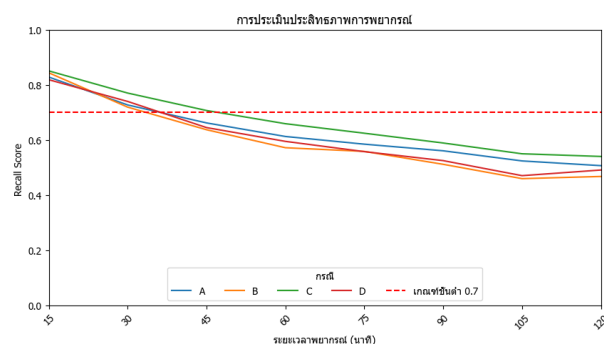
ซึ่งผลของการพยากรณ์ล่วงหน้ามีผลต่อความแม่นยำในการพยากรณ์ล่วงหน้า 120 นาทีเมื่อระยะเวลาเพิ่มขึ้นประสิทธิภาพในการพยากรณ์นั้นลดลง

กลุ่มที่ 1 สถานีนำเข้า คือ สถานีที่สนใจ (กรณี A B C D)

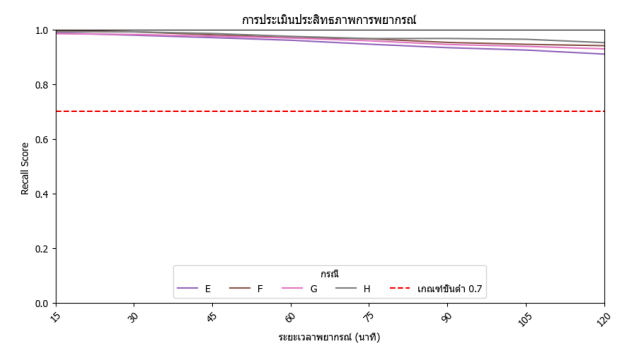
ผลของการพยากรณ์ล่วงหน้าทุก 15 จะเห็นได้ว่าการในลักษณะนี้นั้นสามารถพยากรณ์ล่วงหน้าได้เพียง 30 นาที จะเห็นได้ว่าการเรียนรู้ที่ดีที่สุดคือ การพยากรณ์โดยใช้ข้อมูล กรณี C ให้ recall ในการพยากรณ์ล่วงหน้า 0.85 , 0.77 , 0.70 , 0.65 , 0.62 , 0.58 , 0.54 , 0.54 ตามลำดับการพยากรณ์ล่วงหน้า ดังแสดงในรูปที่ 13

กลุ่มที่ 2 สถานีนำเข้า สถานีที่สนใจกับ 30 สถานีที่ใกล้เคียงที่สุด (กรณี E F G H)

ผลของช่วงเวลาในการพยากรณ์ล่วงหน้า 120 นาทีของการเรียนรู้ในรูปแบบการนำเข้ากลุ่ม 2 นั้นจะให้ประสิทธิภาพสูงคือการเรียนรู้ในรูปแบบ H ที่ให้ความสามารถในการตรวจจับ ณอยู่ที่ 0.99 0.99 0.98 0.97 0.96 0.96 0.96 0.95 ตามลำดับ ดังแสดงในรูปที่ 14



รูปที่ 13 ความสามารถตรวจจับสถานะฝน ล่วงหน้า 120 นาที กรณี A B C D



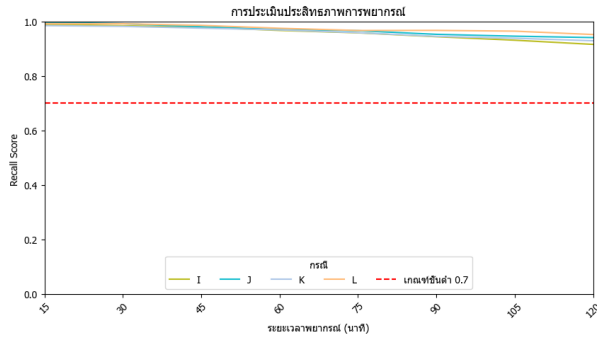
รูปที่ 14 ความสามารถตรวจจับสถานะฝน ล่วงหน้า 120 นาที กรณี E F G H

ซึ่งในทางกลับกันจะเห็นได้ว่าความแม่นยำของการพยากรณ์นั้นจะกลับกันเนื่องจากการพยากรณ์ที่ออกมาจะมีการพยากรณ์ที่ผิดพลาดเพิ่มขึ้นจึงทำให้ประสิทธิภาพแม่นยำสูงสุดอยู่ที่ กรณี E ดังแสดงในรูปที่ 17 กลุ่มที่ 3 สถานีนำเข้า คือ สถานีที่สนใจกับกลุ่มหีบสถานีนำเข้า (กรณี I J K L)

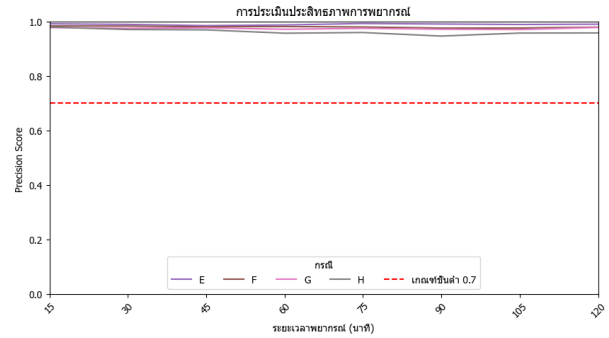
การพยากรณ์ล่วงหน้า 120 นาทีนั้น ความสามารถในการตรวจจับฝนที่ดีที่สุดนั้นอยู่ที่กรณี L 0.99 0.99 0.95 0.97 0.96 0.96 0.96 และ 0.95 ตามลำดับระยะเวลาการพยากรณ์ ดังแสดงในรูปที่ 15 ซึ่งมีลักษณะคล้ายคลึงกับการเรียนรู้ในกลุ่มที่ 2 ที่เป็นการเรียนรู้ข้อมูลในช่วงเวลาฤดูฝนในช่วงเที่ยงวันถึงเที่ยงคืน

ซึ่งความแม่นยำในการพยากรณ์ในกลุ่มที่ 3 นั้นให้ค่าความแม่นยำสูงสุดอยู่ที่การเรียนรู้ กรณี I ที่ให้ค่าความแม่นยำในการพยากรณ์ 0.99 0.98 0.98 0.98 0.98 0.98 และ 0.98 ตามลำดับการพยากรณ์ล่วงหน้า ดังแสดงในรูปที่ 18

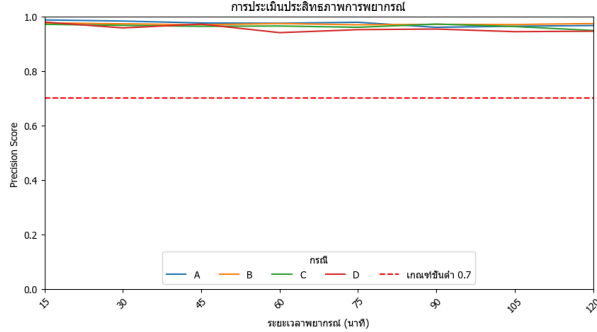
แต่ความสามารถในการตรวจจับที่ดีที่สุดนั้นกลับมีความแม่นยำที่น้อยกว่ากรณีที่มีการตรวจจับได้น้อยกว่า และมีลักษณะที่คล้ายคลึงกับการพยากรณ์ในกลุ่มที่ 2 ที่ความแม่นยำสูงในกรณีที่เรียนรู้ข้อมูลทั้งหมด



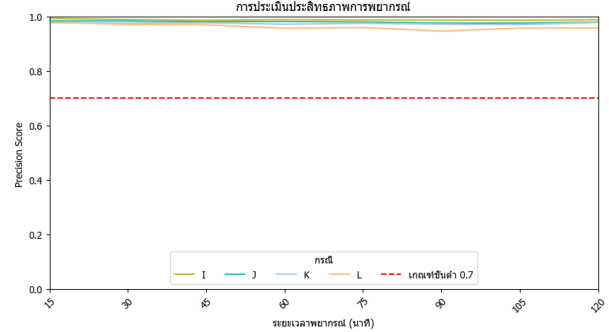
รูปที่ 15 ความสามารถตรวจจับสถานะฝน ล่วงหน้า 120 นาที กรณี I J K L



รูปที่ 17 ความแม่นยำของการพยากรณ์ล่วงหน้า 120 นาที กรณี E F G H



รูปที่ 16 ความแม่นยำของการพยากรณ์ล่วงหน้า 120 นาที กรณี A B C D



รูปที่ 18 ความแม่นยำของการพยากรณ์ล่วงหน้า 120 นาที กรณี I J K L

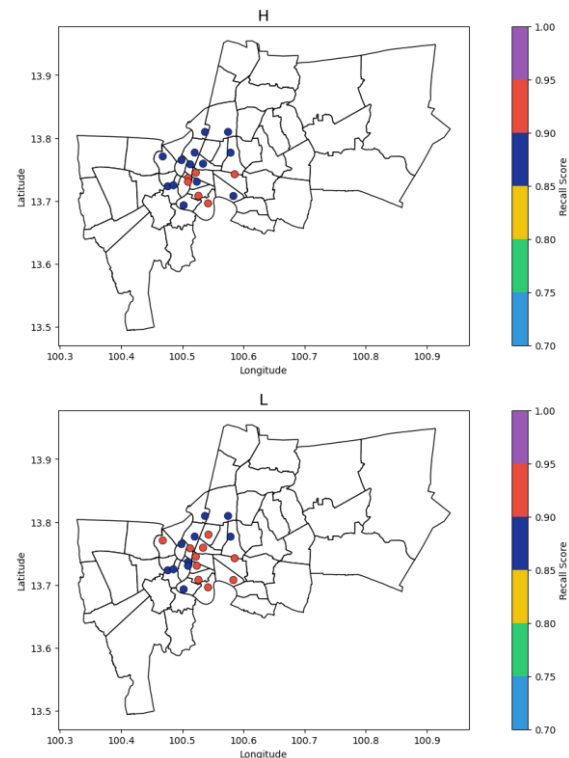
5.4 การทดสอบผลกระทบเนื่องจากตำแหน่งที่ตั้งของสถานี

เนื่องจากการพยากรณ์ล่วงหน้าของการพยากรณ์โดยใช้ข้อมูลในกลุ่มที่ 1 การเรียนรู้เฉพาะ สถานีที่สนใจนั้นไม่สามารถพยากรณ์ล่วงหน้าได้จึงไม่นำมาพิจารณาในขั้นตอนนี้

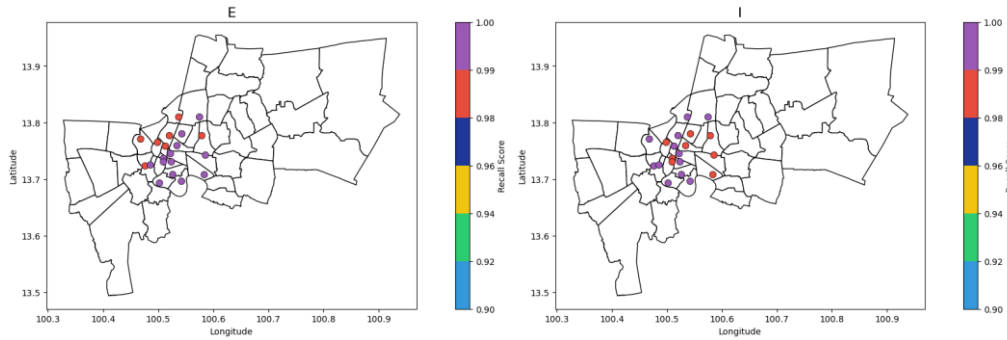
ผลกระทบเนื่องจากตำแหน่งของสถานีที่ระยะการพยากรณ์ล่วงหน้า 120 นาทีที่แยกออกเป็น 2 กลุ่มคือ

1.) ความสามารถในการตรวจจับฝน ที่จะเห็นได้ว่าการพยากรณ์ในกรณี H นั้นจะให้ประสิทธิภาพ 0.9-0.95 จำนวน 6 สถานี ในทิศทางตอนล่างของกรุงเทพชั้นใน ต่ำกว่าการเรียนรู้ในกรณี I ที่มีจำนวน 10 สถานี ในทางตอนล่างของกรุงเทพชั้นในและในทางทิศตะวันออก ดังแสดงในรูปที่ 19

2.) ความแม่นยำในการพยากรณ์เนื่องจากตำแหน่งและลักษณะในการนำเข้าที่แตกต่างกันพบว่าความแม่นยำของกรณี E ที่มีการนำเข้าแบบสถานีใกล้เคียงในให้จำนวนประสิทธิภาพความแม่นยำสูงกว่า 13 สถานี และกรณี I ที่ให้ความแม่นยำสูงกว่า 0.99 นั้นจำนวน 12 สถานี ดังแสดงในรูปที่ 20



รูปที่ 19 ประสิทธิภาพในการตรวจจับฝนของการเรียนรู้ที่ดีที่สุด กรณี H และ L



รูปที่ 20 ประสิทธิภาพในการตรวจจับฝนของการเรียนรู้ที่ดีที่สุด กรณี E และ I

6. บทสรุป

งานวิจัยนี้ได้ทำการทดสอบประสิทธิภาพของการเรียนรู้ในรูปแบบต่าง ออกเป็น 12 กรณีเพื่อการพยากรณ์ล่วงหน้า 120 นาที ระยะเวลาย้อนหลัง ที่ดีที่สุดคือ 540 นาที สำหรับในทั้ง 12 กรณีการเรียนรู้ ความสามารถในการพยากรณ์ล่วงหน้า 120 นาที ในกรณีที่ 1 นั้นไม่สามารถพยากรณ์ล่วงหน้า 120 นาทีได้เนื่องจากความสามารถในการตรวจจับฝนนั้นน้อยกว่า เกณฑ์ขั้นต่ำที่ยอมรับได้โดยแบ่งความสามารถในการพยากรณ์ออกเป็น 2 รูปแบบที่ประกอบด้วยความสามารถในการตรวจจับที่ดีที่สุดและความแม่นยำที่สุดความสามารถในการตรวจจับฝนได้ดีที่สุดนี้เป็นการพยากรณ์ โดยใช้ข้อมูลในช่วงฤดูฝนในช่วงเที่ยงวันถึงเที่ยงคืน ของทั้งกรณีที่ 2 และ 3 นั้นจะให้ประสิทธิภาพที่ 120 นาที ถึง 95% ความแม่นยำในการพยากรณ์ นั้นจะเป็นการเรียนรู้ข้อมูลทั้งหมดเพื่อการพยากรณ์ล่วงหน้า 120 นาที ซึ่ง จะให้ความแม่นยำสูงถึง 98% ทั้งในกรณีที่ 2 และ 3 ผลของการนำเข้าสู่ สถานีแตกต่างกันนั้นจะให้ประสิทธิภาพสูงสุดคือความแม่นยำ คือ การ เรียนรู้ในรูปแบบ E ที่มีการนำเข้าสู่แบบ 30 สถานีที่ใกล้ที่สุด และการ พยากรณ์ที่มีการตรวจจับได้ดีที่สุดเป็นการพยากรณ์ในกรณีที่ I ที่มีการสุ่ม สถานีในการนำเข้าสู่

พบว่าปัจจัยสำคัญที่มีผลต่อประสิทธิภาพของแบบจำลองนั้นสามารถ ได้แก่

ระยะเวลาย้อนหลังของข้อมูลฝน: การใช้ข้อมูลย้อนหลัง 540 นาที ให้ผลการพยากรณ์ที่แม่นยำและสามารถตรวจจับฝนได้ดีที่สุดสำหรับการ พยากรณ์ล่วงหน้า 120 นาที

รูปแบบการเลือกสถานีนำเข้าสู่: การรวมข้อมูลจาก 30 สถานีที่ใกล้ที่สุด หรือการสุ่มเลือกสถานีข้างเคียง ช่วยเพิ่มความสามารถในการตรวจจับฝน ได้ดีกว่าการใช้ข้อมูลจากสถานีเดียวอย่างชัดเจน

ช่วงเวลาของข้อมูลที่น่ามาเรียนรู้: การใช้ข้อมูลเฉพาะช่วงฤดูฝน และ ช่วงเวลา 12:00-24:00 น. ช่วยให้แบบจำลองมีประสิทธิภาพในการ พยากรณ์สูงกว่าการใช้ข้อมูลตลอดทั้งปี

7. ข้อเสนอแนะ

เนื่องจากการทำการพยากรณ์ฝนตกหรือไม่ตกนั้นสามารถนำไปใช้ในการ ต่อยอดเพื่อบอกว่าการพยากรณ์ฝนตกนั้นมีความรุนแรงของฝนได้มาก น้อยเพียงใดเมื่อเป็นการพยากรณ์และเนื่องจากการศึกษานั้นนั้นมีการเรียนรู้

ข้อมูลฝนสถานีเพียงอย่างเดียวนั้นสามารถเพิ่มข้อมูลทางอุตุนิยมวิทยาอื่นๆ เพื่อให้ได้การพยากรณ์ฝนที่มีความแม่นยำสูงขึ้น

การนำวิธีการดังกล่าวไปใช้กับการพยากรณ์ในพื้นที่เมืองที่มีลักษณะ คล้ายคลึงกับพื้นที่ดังกล่าวเพื่อความเร็วสำหรับการใช้งานแจ้งเตือนได้ อย่างรวดเร็ว

เอกสารอ้างอิง

- [1] Subrahmanyam, V., & Ghosh, S. (2013). "Statistical Models in Weather Forecasting." *Climate Dynamics*, 41 (9), 2001-2015
- [2] Meng, J., & Liu, B. (2019). "A review of physics-based weather forecasting models." *Atmospheric Research*, 228, 81-98.
- [3] สิริธรร, น. (2561) การศึกษาการเปลี่ยนแปลงตามพื้นที่ของโค้งความ เข้มฝน-ช่วงเวลาและรอบปีการเกิดซ้ำและรูปแบบฝนในเขต กรุงเทพมหานคร.,วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต ,คณะวิศวกรรมศาสตร์,มหาวิทยาลัยเกษตรศาสตร์.
- [4] สำนักการระบายน้ำ(2566). แผนปฏิบัติการป้องกันและแก้ไขปัญหาน้ำท่วม. สำนักการระบายน้ำ
- [5] Srithagon, S., Phisanbut, N., Piamsa-nga, N., & Piamsa-nga, P. (2021). Rainfall nowcasting based on neighboring rain gauge stations using learning machines Paper presented at the 2021 25th International Computer Science and Engineering Conference (ICSEC). 18-20 Nov. 2021, pp.33-38
- [6] กรมอุตุนิยมวิทยา. (2567). ภูมิอากาศของประเทศไทย. *กรมอุตุนิยมวิทยา*.<https://www.tmd.go.th>
- [7] สำนักการระบายน้ำ(2566). แผนปฏิบัติการป้องกันและแก้ไขปัญหาน้ำท่วม. สำนักการระบายน้ำ
- [8] Hamill, T. M., & Nehrkorn, T. (1993). A short-term cloud forecast scheme using cross correlations. AIR FORCE MATERIEL COMMAND HANSCOM AIR FORCE BASE, MA 01731-3010, 346.M
- [9] Thong, P. H., & Fujita, H. (2016, July). Interpolative picture fuzzy rules: A novel forecast method for weather

- nowcasting. In 2016 IEEE international conference on fuzzy systems (FUZZ-IEEE) (pp. 86-93). IEEE.
- [10] Son, L. H., & Thong, P. H. (2017). Some novel hybrid forecast methods based on picture fuzzy clustering for weather nowcasting from satellite image sequences. *Applied Intelligence*, 46, 1-15.
- [11] Wetchayont, P., Ekkawatpanit, C., Rueangrit, S., & Manduang, J. (2023). Improvements in rainfall estimation over Bangkok, Thailand by merging satellite, radar, and gauge rainfall datasets with the geostatistical method. *Big Earth Data*, 7(2), 251–275.
- [12] [11] Kessler, E. (1966). Computer program for calculating average lengths of weather radar echoes and pattern bandedness. *Journal of Atmospheric Sciences*, 23(5), 569-574.
- [13] Hering, A. M., Morel, C., Galli, G., Sénési, S., Ambrosetti, P., & Boscacci, M. (2004, September). Nowcasting thunderstorms in the Alpine region using a radar based adaptive thresholding scheme. In *Proceedings of ERAD* (Vol. 1, No. 6).
- [14] Jung, S. H., & Lee, G. (2015). Radar-based cell tracking with fuzzy logic approach. *Meteorological Applications*, 22(4), 716-730.
- [15] Breiman, L. (2544). RANDOM FOREST. Statistics Department University of California Berkeley, CA 94720
- [16] Aditya, G., Sri Suryani, P., & Yuliant, S. (2565). Performance Analysis of Hybrid Machine Learning Methods on Imbalanced Data (Rainfall Classification). *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 6 (3) . doi:10.29207/resti.v6i3.4142
- [17] R. H. Kumar, B. Vishnu Prasad, H. Yadav and M. Venugopalan, "Rainfall Prediction Enhancement Using SMOTE and Machine Learning Algorithms," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-6
- [18] Irshada, M., Kumar, V. SMOTE and ExtraTreesRegressor based random forest technique for predicting Australian rainfall. *Int. j. inf. tecnol.* 15, 1679–1687 (2023)