

Predicting anatomic pathology turnaround time using machine learning models

Waanpa Kinnares

Sakon Nakhon Hospital, Sakon Nakhon

Abstract

Turnaround Time (TAT) is a critical quality indicator in anatomic pathology laboratories, yet its management remains a challenge, particularly in resource-limited public hospitals in Thailand. While recent work shows the potential of machine learning (ML) for operational intelligence, studies focusing on TAT prediction in this specific context are scarce. This study aimed to develop and evaluate ML models for predicting the TAT of anatomic pathology specimens at Sakon Nakhon Hospital and to identify the key factors influencing TAT. A retrospective dataset of 37,515 cleaned pathology records was utilized. Feature engineering was performed to create variables reflecting workload (e.g., daily case count) and temporality (e.g., week of the year). Several models, including Random Forest and LightGBM, were developed and compared, with Mean Absolute Error (MAE) as the primary performance metric. The optimized Random Forest model demonstrated the highest performance, achieving

an MAE of 1.85 days and an R-squared (R^2) of 0.91 on the test set. Feature importance analysis revealed that the assigned pathologist, week of the year, and daily case count were the most significant predictors, highlighting the impact of personnel, temporal, and workload factors. Machine learning models, particularly Random Forest, show high potential for accurately predicting TAT. The findings can be leveraged to develop a decision support tool for proactive laboratory management, aiming to enhance operational efficiency and improve the quality of patient care in resource-constrained healthcare environments.

Keywords: *Turnaround time; Anatomic pathology; Machine learning; Random forest, Predictive modeling*

Received: 29 September 2025, Revised: 25 October 2025, Accepted: 30 November 2025

Correspondence: Waanpa Kinnares, Sakon Nakhon Hospital, 1041 Rueang Sawat Rd, That Choeng Chum, Mueang Sakon Nakhon District, Sakon Nakhon 47000 Thailand, Tel.: 042 176 000, email: klopeme@gmail.com

การพยากรณ์ระยะเวลาของผลตรวจทางพยาธิวิทยาจาก ด้วยแบบจำลองการเรียนรู้

ว่านป้า กิณเรศ

โรงพยาบาลศูนย์สกลนคร สกลนคร

บทคัดย่อ

ระยะเวลาของผล (Turnaround Time: TAT) เป็นตัวชี้วัดคุณภาพที่สำคัญในห้องปฏิบัติการพยาธิวิทยา แต่การบริหารจัดการในโรงพยาบาลรัฐมีข้อจำกัดหลายด้าน งานวิจัยล่าสุดชี้ว่าแบบจำลองการเรียนรู้สามารถเพิ่มประสิทธิภาพการดำเนินงานได้ แต่การศึกษาที่มุ่งเน้น ปัจจุบันยังไม่มีข้อมูลเกี่ยวกับการพยากรณ์ระยะเวลาของผลมากนัก การศึกษานี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินประสิทธิภาพของแบบจำลองการเรียนรู้ในการพยากรณ์ TAT ของสิ่งส่งตรวจทางพยาธิวิทยาจาก ณ โรงพยาบาลศูนย์สกลนคร และระบุปัจจัยสำคัญที่มีอิทธิพลต่อ TAT ใช้ข้อมูลย้อนหลังที่ผ่านการทำความสะอาด จำนวน 37,515 รายการ มีการสร้างปัจจัยเพิ่มเติมเพื่อสะท้อนภาระงาน (เช่น จำนวนเคสรายวัน) และปัจจัยด้านเวลา (เช่น สัปดาห์ของปี) จากนั้นได้พัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลองหลายชนิด รวมถึง Random Forest และ LightGBM โดยใช้ค่า Mean Absolute Error (MAE) แบบจำลอง Random Forest ที่ผ่านการปรับปรุงให้ประสิทธิภาพสูงสุดสามารถพยากรณ์ TAT ด้วยค่าความคลาดเคลื่อนเฉลี่ย (MAE)

เพียง 1.85 วัน และมีค่า R-squared (R^2) 0.91 ปัจจัยที่มีอิทธิพลสูงสุดต่อการพยากรณ์คือพยาธิแพทย์ผู้รับผิดชอบ, สัปดาห์ของปี, และภาระงานรายวัน ซึ่งชี้ให้เห็นถึงผลกระทบของปัจจัยด้านบุคคล เวลา และปริมาณงาน แบบจำลอง Random Forest มีศักยภาพสูงในการพยากรณ์ TAT ได้อย่างแม่นยำ ผลลัพธ์ที่ได้สามารถนำไปพัฒนาเป็นเครื่องมือสนับสนุนการตัดสินใจสำหรับการบริหารจัดการเชิงรุกในห้องปฏิบัติการและเพิ่มประสิทธิภาพการดำเนินงานและปรับปรุงคุณภาพการดูแลผู้ป่วยในบริบทของโรงพยาบาลรัฐ

คำสำคัญ: ระยะเวลาดำเนินการ; พยาธิวิทยาทางกายวิภาค; การเรียนรู้ของเครื่อง; Random forest; การสร้างแบบจำลองเชิงทำนาย

วันที่รับต้นฉบับ: 29 กันยายน 2568, วันที่แก้ไข: 25 ตุลาคม 2568, วันที่ตอบรับ: 30 พฤศจิกายน 2568

บทนำ

พยาธิวิทยาทางกายวิภาค (Anatomic Pathology) เป็นสาขาวิชาทางการแพทย์ที่มีความสำคัญในกระบวนการวินิจฉัยโรค โดยเฉพาะโรคมะเร็งและโรคที่มีความซับซ้อน ซึ่งต้องอาศัยการตรวจชิ้นเนื้อทางจุลพยาธิวิทยาเพื่อยืนยันการวินิจฉัยและวางแผนการรักษาที่แม่นยำ [1, 2] ตัวชี้วัดคุณภาพและประสิทธิภาพที่สำคัญของห้องปฏิบัติการ คือ ระยะเวลาของผล (Turnaround Time: TAT) หมายถึงช่วงเวลาตั้งแต่ห้องปฏิบัติการได้รับสิ่งส่งตรวจจนถึงการออกรายงานผลฉบับสมบูรณ์ [3, 4] การลดระยะเวลา TAT ไม่เพียงแต่จะช่วยให้ผู้ป่วยได้รับการวินิจฉัยและเริ่มการรักษาที่เหมาะสมได้รวดเร็วยิ่งขึ้น ยังส่งผลต่อผลลัพธ์การรักษาและช่วยลดความวิตกกังวลของผู้ป่วย [5] แต่ยังคงสะท้อนถึงประสิทธิภาพในการบริหารจัดการทรัพยากรของโรงพยาบาลด้วย [6]

ปัจจัยที่มีผลต่อ TAT นั้นมีความซับซ้อนและหลากหลาย ตั้งแต่ชนิดและความเร่งด่วนของสิ่งส่งตรวจ ไปจนถึงกระบวนการภายในห้องปฏิบัติการ เช่น การทำสไลด์พยาธิวิทยาขณะผ่าตัด (Frozen section) ซึ่งราชวิทยาลัยพยาธิแพทย์แห่งประเทศไทยกำหนดเป้าหมาย TAT ไว้ที่ 30 นาที [7] หรืองานชิ้นเนื้อขนาดใหญ่และซับซ้อนที่ต้องผ่านการจัดการเป็นพิเศษ (Special Handling) เช่น การล้างแคลเซียม (Decalcification) หรือการย้อมอิมมูโนฮิสโตเคมี (Immunohistochemistry) ซึ่งอาจทำให้ TAT ยาวนานกว่ากรณีทั่วไปถึงสองเท่า [4] อย่างไรก็ตามการบรรลุเป้าหมาย TAT ตามมาตรฐานสากลมักเป็นความท้าทายโดยเฉพาะบริบทของโรงพยาบาลรัฐซึ่งอาจเผชิญกับข้อจำกัดด้านบุคลากรผู้เชี่ยวชาญและทรัพยากร [8, 9]

ความก้าวหน้าของเทคโนโลยีพยาธิวิทยาดิจิทัล (Digital Pathology) โดยเฉพาะการใช้เครื่องสแกนสไลด์ความละเอียดสูง (Whole Slide Imaging: WSI) ได้ปฏิวัติวงการพยาธิวิทยาโดยเปลี่ยนสไลด์แก้วให้เป็นข้อมูลดิจิทัล [10] การเปลี่ยนแปลงนี้ไม่เพียงแต่ช่วยอำนวยความสะดวกในการให้คำปรึกษาทางไกล (Telepathology) [11] แต่ยังเป็นรากฐานสำคัญที่เปิดโอกาสให้นา

ผู้พิมพ์ประสานงาน: ว่านป้า กิณเรศ โรงพยาบาลศูนย์สกลนคร 1041 ถ.เจริญเมือง ต.ธาตุเชิงชุม อ.เมืองสกลนคร จ.สกลนคร 47000, โทร.: 042 176 000, email: klopeme@gmail.com

เทคนิคปัญญาประดิษฐ์ (Artificial Intelligence: AI) และแบบจำลองการเรียนรู้ (Machine Learning: ML) มาประยุกต์ใช้ [12, 13] ปัจจุบัน AI/ML ได้ถูกนำมาใช้เป็นเครื่องมือสนับสนุนการตัดสินใจทางคลินิก (Clinical Decision Support: CDS) อย่างกว้างขวาง ภายใต้แนวคิด "ปัญญาเสริม" (Augmented Intelligence) ซึ่งเป็นการผสานความเชี่ยวชาญของมนุษย์เข้ากับความสามารถในการวิเคราะห์ของ AI เพื่อเพิ่มทั้งความแม่นยำและประสิทธิภาพ [14] โดยมีการนำไปช่วยวินิจฉัยโรคจากข้อมูลหลากหลายรูปแบบ ตั้งแต่ภาพถ่ายสไลด์ [15] ไปจนถึงการวิเคราะห์สัญญาณเสียงพูด [16]

AI และ ML นอกจากสนับสนุนด้านการวินิจฉัยแล้วยังมีศักยภาพในการวิเคราะห์ข้อมูลที่ถูกบันทึกไว้ในระบบสารสนเทศของห้องปฏิบัติการ (Laboratory Information System: LIS) ทำให้เข้าใจรูปแบบการทำงานและระบุปัจจัยที่เป็นคอขวดของกระบวนการได้ [17] งานวิจัยหลายชิ้นแสดงให้เห็นว่าการวิเคราะห์เชิงพยากรณ์ (Predictive Analytics) สามารถนำมาใช้เพื่อคาดการณ์ TAT และเพิ่มประสิทธิภาพการจัดลำดับความสำคัญของงานในห้องปฏิบัติการได้ [18] ทั้งนี้ การสร้างแบบจำลองที่น่าเชื่อถือจำเป็นต้องผ่านกระบวนการ การทวนสอบ (Verification) และ การตรวจสอบความถูกต้อง (Validation) ที่เข้มงวด [1] และต้องอาศัยเทคนิค Explainable AI (XAI) เพื่อให้มั่นใจว่าการตัดสินใจของแบบจำลองมีความโปร่งใสและสามารถอธิบายได้ในทางคลินิก [19]

การศึกษานี้มุ่งเน้นการพัฒนาแบบจำลองเชิงพยากรณ์สำหรับ TAT ในบริบทของโรงพยาบาลรัฐของไทยที่มีปัจจัยเฉพาะตัว เช่น วันหยุดราชการ, ความผันผวนของภาระงาน, และข้อจำกัดด้านทรัพยากรที่แตกต่างจากประเทศพัฒนาแล้ว [20] วัตถุประสงค์ของการศึกษานี้เพื่อพัฒนาแบบจำลองการเรียนรู้สำหรับพยากรณ์ TAT ของสิ่งส่งตรวจทางพยาธิวิทยาภาค และระบุปัจจัยสำคัญที่มีอิทธิพลต่อ TAT ในบริบทของโรงพยาบาลศูนย์สกลนคร ผลการศึกษานี้ได้จะเป็นการบุกเบิกการประยุกต์ใช้ ML เพื่อยกระดับคุณภาพและประสิทธิภาพของห้องปฏิบัติการพยาธิวิทยาในโรงพยาบาลรัฐ และอาจใช้เป็นต้นแบบสำหรับโรงพยาบาลอื่นๆ ในการนำเทคโนโลยีไปปรับใช้เพื่อพัฒนาระบบบริการสุขภาพต่อไป

วัสดุและวิธีการ

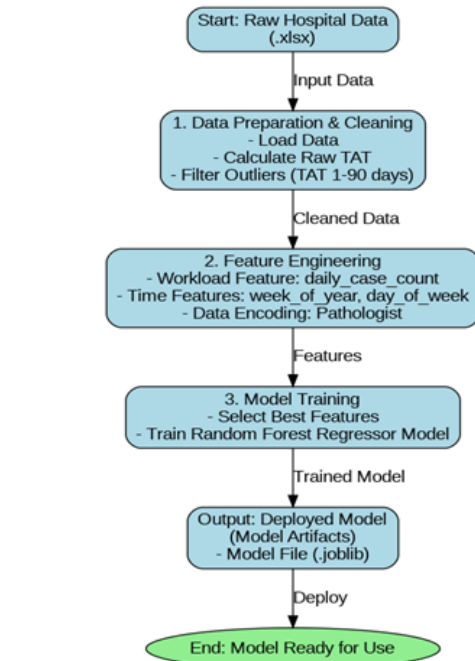
คำอธิบายชุดข้อมูล

การวิจัยนี้เป็นการศึกษาเชิงวิเคราะห์ย้อนหลัง โดยใช้ข้อมูลสิ่งส่งตรวจทางพยาธิวิทยาภาคของโรงพยาบาลศูนย์สกลนคร ระหว่างวันที่ 1 มกราคม พ.ศ. 2565 ถึงวันที่ 31 พฤษภาคม พ.ศ. 2568 รวมทั้งสิ้น 50,322 รายการ ข้อมูลที่รวบรวมประกอบด้วย 1. ข้อมูลเชิงเวลา คือ receive_date (วันที่รับส่งตรวจ), report_date (วันที่รายงานผล) 2. ข้อมูลประชากรผู้ป่วย คือ age (อายุ), sex (เพศ) 3. ข้อมูลการตรวจ คือ gitem (รหัสกรม

บัญชีกลาง), lab (inhouse/outsourse), labtype (surgical pathology/cytology), specimencity (simple/ intermediate/ complex), specimentype (routine/IHC+ special stain/ molecular studies) 4. ข้อมูลกระบวนการทำงาน คือ pathologist (รหัสพยาธิแพทย์ผู้รับผิดชอบ), departmentsent (แผนกที่ส่งตรวจ), Priority Level (ระดับความเร่งด่วน) ตัวแปรเป้าหมาย (Target Variable): TAT_days ซึ่งคำนวณจาก report_date และ receive_date และสร้างตารางวันหยุดนักขัตฤกษ์และวันหยุดราชการของไทยตั้งแต่ปี 2565-2568 ดังแสดงในตาราง 1

ตารางที่ 1 คำอธิบายตัวแปรในชุดข้อมูล

ตัวแปร (Variable)	คำจำกัดความ (Definition)	ประเภทข้อมูล (Data Type)	ค่าที่เป็นไปได้ / ตัวอย่าง (Values / Example)
receive_date	วันที่ห้องปฏิบัติการได้รับส่งตรวจ	วันที่ (Date)	DD/MM/YYYY (เช่น 2/5/2022)
report_date	วันที่พยาธิแพทย์ลงนามอนุมัติรายงานผลฉบับสมบูรณ์	วันที่ (Date)	DD/MM/YYYY (เช่น 2/5/2022)
age	อายุของผู้ป่วย ณ วันที่รับส่งตรวจ (ปี)	จำนวนเต็ม (Integer)	>= 0
sex	เพศของผู้ป่วย	ข้อความ (String)	'Male' = 1, 'Female' = 2
gitem	รหัสกรมบัญชีกลาง (ค่าตรวจวินิจฉัยทางเทคนิคการแพทย์และพยาธิวิทยา)	ข้อความ (String / Code)	รหัสเฉพาะ (เช่น '38001', '38002')
lab	สถานที่ทำการตรวจ (ภายในโรงพยาบาล หรือส่งตรวจภายนอก)	ข้อความ (String)	'inhouse' = 1, 'outsourse' = 2
labtype	ประเภทของงานตรวจทางพยาธิวิทยา (คือพยาธิวิทยา หรือเซลล์วิทยา)	ข้อความ (String)	'surgical pathology' = 1, 'cytology' = 2
specimencity	ระดับความซับซ้อนของสิ่งส่งตรวจ (อ้างอิงตามเกณฑ์ภายใน หรือลักษณะชิ้นเนื้อ) โดยอิงตามเกณฑ์ของกุ่มงานพยาธิวิทยาภาควิภาค โรงพยาบาลสกลนคร กำหนด	ข้อความ (String)	'simple' (ชิ้นเนื้อเล็ก ไม่ซับซ้อน) = 1, 'intermediate' (ชิ้นเนื้อขนาดกลาง) = 2, 'complex' (ชิ้นเนื้อขนาดใหญ่/หลายชิ้น/ซับซ้อน) = 3
specimentype	ประเภทการตรวจที่ทำกับสิ่งส่งตรวจ	ข้อความ (String)	'routine' (H&E ปกติ) = 1, 'IHC+ special stain' (มีการย้อมพิเศษ/อิมมูโนฯ) = 2, 'molecular studies' (มีการตรวจระดับโมเลกุล) = 3
pathologist	รหัสประจำตัวของพยาธิแพทย์ผู้รับผิดชอบหลักในการออกรายงาน	ข้อความ (String / Code)	รหัสเฉพาะ เช่น พยาธิแพทย์คนที่ 1 = 1, พยาธิแพทย์คนที่ 4 = 4,
departmentsent	แผนก/หน่วยงานภายในโรงพยาบาลที่ส่งส่งตรวจ	ข้อความ (String)	'แผนกศัลยกรรม' = 1, 'แผนกอายุรกรรม' = 2, 'เวชกรรมพิเศษ' = 17
Priority Level	ระดับความเร่งด่วนของสิ่งส่งตรวจ โดยอิงตามเกณฑ์ของกุ่มงานพยาธิวิทยาภาควิภาค โรงพยาบาลสกลนคร กำหนด	ข้อความ (String)	'Stat' (ด่วนที่สุด) = 1, 'Urgent' (ด่วน) = 2, 'Routine' (ปกติ) = 3
TAT_days	ตัวแปรเป้าหมาย: ระยะเวลาออก (วันทำการ) คำนวณจาก report_date - receive_date โดยไม่นับวันหยุด	จำนวนเต็ม (Integer)	>= 0
ตารางวันหยุด	ข้อมูลอ้างอิง: รายการวันหยุดนักขัตฤกษ์และวันหยุดราชการพิเศษของไทย พ.ศ. 2565-2568	รายการวันที่ (List of Dates)	ใช้ประกอบการคำนวณ TAT_days

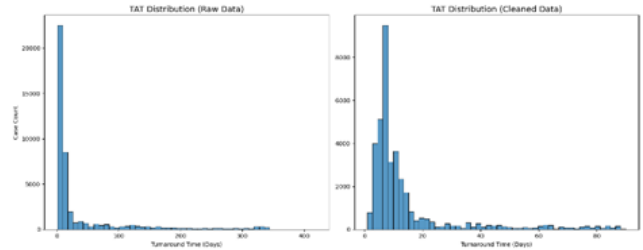


Workflow for Developing the TAT Prediction

รูปที่ 1 ขั้นตอนการดำเนินงานวิจัย เริ่มจากการทำความสะอาดข้อมูลดิบ (Data Cleaning), การสร้างปัจจัยใหม่ (Feature Engineering) ประกอบด้วย ปัจจัยด้านภาระงาน (Workload Features), ปัจจัยด้านเวลา (Time Features), และการเข้ารหัสพยาธิแพทย์ (Pathologist Encoding) จากนั้นนำข้อมูลไปฝึกสอนแบบจำลอง (Model Training) เพื่อค้นหาปัจจัยที่ดีที่สุด (Best Features) และแบบจำลองที่มีประสิทธิภาพสูงสุด (Best Model)

การเตรียมข้อมูล

ขั้นตอนการเตรียมข้อมูลประกอบด้วย การแปลงวันที่เป็นรูปแบบ datetime, การคำนวณตัวแปรเป้าหมาย TAT ($TAT_days = (report_date - receive_date).days + 1$), ข้อมูลหลังทำความสะอาดข้อมูล โดยเติมข้อมูลในช่องว่าง ลบข้อมูลที่ไม่ได้เติมในช่องวันรายงาน ไม่ใช่พยาธิแพทย์และ TAT ติดลบ เหลือข้อมูลสำหรับการวิเคราะห์จำนวน 43,390 รายการ การทำความสะอาดข้อมูลเพิ่มเติมเพื่อปรับการกระจายตัวของ TAT ดังภาพ 2 โดยกำหนดเกณฑ์ว่าข้อมูลที่ถูกต้องควรมี TAT อยู่ในช่วง 1 ถึง 90 วัน และการคัดกรองข้อมูลผิดปกติ (Outliers) ออก หลังจากการเตรียมข้อมูล ชุดข้อมูลสำหรับการวิเคราะห์มีจำนวน 37,515 รายการ (คัดออก 5,875 รายการ)



รูปที่ 2 การกระจายตัวของประชากรก่อนและหลังทำความสะอาดข้อมูล

สถิติเชิงพรรณนา (Descriptive Statistics)

ก่อนการสร้างแบบจำลอง ได้มีการวิเคราะห์สถิติเชิงพรรณนาเพื่อทำความเข้าใจลักษณะการกระจายตัวของข้อมูลในตัวแปรที่สำคัญ โดยเฉพาะตัวแปรเชิงปริมาณอย่างอายุของผู้ป่วยและระยะเวลาารผล (TAT) ซึ่งผลลัพธ์สรุปได้ดังตาราง 2

ตารางที่ 2 สถิติเชิงพรรณนาของตัวแปรเชิงปริมาณ (N = 37,515)

ตัวแปร	ค่าเฉลี่ย (Mean)	ส่วนเบี่ยงเบนมาตรฐาน (SD)	ค่าต่ำสุด (Min)	ควอร์ไทล์ที่ 1 (25%)	ค่ามัธยฐาน (50%)	ควอร์ไทล์ที่ 3 (75%)	ค่าสูงสุด (Max)
อายุ (Age)	54.46	18.22	0	43	57	68	101
TAT_days	14.21	16.65	1	6	8	13	90
daily_case_count	45.75	21.71	1	31	43	58	168

รายละเอียดแบบจำลองและอัลกอริทึม

ในการศึกษาชิ้นนี้ใช้ข้อมูลดิบ (Raw Data) เพียงอย่างเดียวโดยไม่ต้องทำความสะอาดข้อมูล เพื่อให้แบบจำลองการเรียนรู้สามารถเรียนรู้และพยากรณ์ TAT ได้อย่างมีประสิทธิภาพสูงสุด จึงสร้างปัจจัยใหม่ (Feature Engineering) ซึ่งเป็นขั้นตอนสำคัญในการสกัดข้อมูลเชิงลึก (Insights) ที่ซ่อนอยู่ในข้อมูลดิบ ปัจจัยเหล่านี้ถูกออกแบบมาเพื่อสะท้อนบริบทการทำงานจริงของห้องปฏิบัติการพยาธิวิทยา ซึ่งคาดว่าจะมีความสัมพันธ์โดยตรงกับระยะเวลาการทำงานในห้องปฏิบัติการและช่วยเพิ่มความแม่นยำในการพยากรณ์ TAT โดยปัจจัยที่สร้างขึ้นใหม่ (Engineered Features) มุ่งเน้นการจับรูปแบบที่ซับซ้อนซึ่งแบบจำลองอาจมองข้ามไป โดยแบ่งเป็นกลุ่มหลักๆ ดังนี้ 1. ปัจจัยด้านภาระงาน (Workload Features) เพื่อวัดปริมาณงานและความหนาแน่นของบริการ ณ เวลาหนึ่งๆ ซึ่งส่งผลโดยตรงต่อทรัพยากรที่มีจำกัด ตัวอย่างเช่น daily_case_count (จำนวนเคสทั้งหมดที่รับในวันนั้น) และ rolling_avg_workload_7d (ค่าเฉลี่ยภาระงานสะสมย้อนหลัง 7 วัน) 2. ปัจจัยด้านเวลาและปฏิทิน (Time-based Features) เพื่อจับรูปแบบ (Pattern) ที่เกี่ยวข้องกับวงจรการทำงานรายสัปดาห์หรือรายปี เช่น recv_day_of_week (วันที่รับส่งตรวจในสัปดาห์ ซึ่งอาจสะท้อนภาระงานสะสมจากสัปดาห์ก่อน), recv_week_of_year (สัปดาห์ที่รับของปี เพื่อดูผลกระทบเชิงฤดูกาล) และ is_holiday (ตัวแปรระบุว่าเป็นวันหยุดราชการ

หรือนักชดถุขหรือไม) ซึ่งเป็นปัจจัยสำคัญอย่างย้งในบริบทของโรงพยาบาลรัฐ โดยทำการเข้ารหัสข้อมูลเชิงคุณภาพ (Categorical Encoding) เพื่อแปลงข้อมูลที่ไม่ใช่ตัวเลข (เช่น รหัสพยาธิแพทย์ หรือประเภทการตรวจ) ให้อยู่ในรูปแบบที่แบบจำลองการเรียนรู้สามารถประมวลผลได้ เช่น pathologist_encoded และ specimen_type_encoded รายละเอียดคำจำกัดความและวิธีการคำนวณปัจจัยที่สร้างขึ้นใหม่ทั้งหมดได้สรุปไว้ในตาราง 3

ตารางที่ 3 คำอธิบายปัจจัยที่สร้างขึ้นใหม่ (Engineered Features)

ปัจจัยใหม่ (Engineered Feature)	คำอธิบาย (Description)	วิธีการคำนวณ / การเข้ารหัส (Calculation / Encoding Method)	วัตถุประสงค์ (Purpose)
daily_case_count	จำนวนเคสทั้งหมดที่ได้รับในวันเดียวกับเคสปัจจุบัน	นับจำนวนเคสทั้งหมดที่มี receive_date เดียวกัน	วัดภาระงาน ณ วันที่รับเคส
recv_day_of_week	วันของสัปดาห์ที่รับสิ่งส่งตรวจ (เช่น จันทร์, อังคาร)	แปลง receive_date เป็นวันของสัปดาห์ (อาทเข้าหัสเป็นตัวเลข 0-6 หรือ One-hot Encoding)	ดูผลกระทบของวันในสัปดาห์ (เช่น วันจันทร์อาจมีผลต่อ turnaround time)
recv_week_of_year	สัปดาห์ของปีที่รับสิ่งส่งตรวจ (1-52/53)	แปลง receive_date เป็นหมายเลขสัปดาห์ของปี	ดูผลกระทบของช่วงเวลาของปี (เช่น ใกล้เคียงปีงบประมาณ)
pathologist_encoded	รหัสพยาธิแพทย์ที่ถูกเข้ารหัสเป็นตัวเลข	Target Encoding, Label Encoding, หรือ One-hot Encoding จากค่า pathologist	ทำให้แบบจำลองการเรียนรู้สามารถใช้ข้อมูลพยาธิแพทย์ได้ (อาจสะท้อนประสบการณ์หรือภาระงาน)
specimentype_encoded	ประเภทการตรวจที่ถูกเข้ารหัสเป็นตัวเลข	Label Encoding หรือ One-hot Encoding จากค่า specimen_type	ทำให้แบบจำลองการเรียนรู้สามารถใช้ข้อมูลประเภทการตรวจได้ (สะท้อนความซับซ้อนของงาน)
rolling_avg_workload_7d	ค่าเฉลี่ยเคลื่อนที่ (7 วัน) ของจำนวนเคสที่ได้รับต่อวัน ก่อนถึงวันที่รับเคสปัจจุบัน	คำนวณค่าเฉลี่ยของ daily_case_count ในช่วง 7 วันก่อนหน้า receive_date	วัดภาระงานสะสมหรือแนวโน้มภาระงานในช่วงก่อนหน้า

ขั้นตอนถัดไปคือการคัดเลือกเฉพาะปัจจัยที่มีอิทธิพลต่อ TAT ที่มีนัยสำคัญทางสถิติ เพื่อลดความซับซ้อนของแบบจำลองและเพื่อป้องกันปัญหา Overfitting เนื่องจากตัวแปรเป้าหมาย (TAT) มีลักษณะเป็นข้อมูลประเภท Count Data (จำนวนวัน) ซึ่งมักพบการกระจายตัวแบบเบ้ขวา (Right-skewed) และมีแนวโน้มความแปรปรวนสูง (Overdispersion) จึงควรเลือกใช้เทคนิคที่เหมาะสมในการสร้างแบบจำลอง การใช้แบบจำลองการถดถอยเชิงเส้นตรง (Linear Regression) แบบปกติอาจไม่เหมาะสมในการศึกษานี้ จึงเลือกใช้ การวิเคราะห์การถดถอยแบบ Negative Binomial (Negative Binomial Regression) ซึ่งเป็นส่วนหนึ่งของตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Model - GLM) ในการประเมินความสัมพันธ์และนัยสำคัญทางสถิติของปัจจัยแต่ละตัว ปัจจัยทั้งหมดที่ถูกคัดเลือกเพื่อนำไปใช้ในการ

สร้างแบบจำลองการเรียนรู้ขั้นสุดท้าย ล้วนมีนัยสำคัญทางสถิติที่ระดับ $P\text{-value} < 0.05$ ($P > |z| < 0.05$) ดังผลลัพธ์ที่สรุปในตาราง 4

อัลกอริทึมที่ใช้ในการศึกษานี้ประกอบด้วย Random Forest Regressor ซึ่งเป็นอัลกอริทึมแบบ Ensemble ที่สร้างต้นไม้ตัดสินใจ (Decision Trees) จำนวนมากและนำผลลัพธ์มาหาค่าเฉลี่ย และ LightGBM Regressor ซึ่งเป็นอัลกอริทึมในกลุ่ม Gradient Boosting ที่มีประสิทธิภาพสูงและความเร็วในการฝึกสอนที่รวดเร็ว เหมาะสำหรับชุดข้อมูลขนาดใหญ่

ตารางที่ 4 ผลการวิเคราะห์นัยสำคัญทางสถิติของปัจจัยด้วย Negative Binomial Regression

Feature	Coefficient	P-value	95% Confidence Interval
Intercept	1.553	< 0.001	[0.979, 2.127]
C(pathologist)[T.2]	0.124	< 0.001	[0.098, 0.150]
C(pathologist)[T.3]	0.807	< 0.001	[0.767, 0.847]
C(pathologist)[T.4]	1.541	< 0.001	[1.507, 1.575]
C(specimentype)[T.2]	0.287	< 0.001	[0.251, 0.324]
C(specimentype)[T.3]	0.228	0.027	[0.026, 0.429]
C(departmentsent)[T.2]	-0.066	0.001	[-0.103, -0.029]
C(departmentsent)[T.3]	-0.021	0.216	[-0.054, 0.012]
C(departmentsent)[T.4]	-0.03	0.129	[-0.069, 0.009]
C(departmentsent)[T.5]	0.144	< 0.001	[0.099, 0.188]
C(departmentsent)[T.6]	-0.044	0.043	[-0.087, -0.001]
C(departmentsent)[T.7]	-0.11	< 0.001	[-0.157, -0.063]
C(departmentsent)[T.8]	-0.12	< 0.001	[-0.185, -0.055]
C(departmentsent)[T.9]	0.093	0.044	[0.003, 0.182]
C(departmentsent)[T.10]	0.178	0.004	[0.057, 0.300]
C(departmentsent)[T.11]	-0.172	0.037	[-0.333, -0.010]
C(departmentsent)[T.12]	0.039	0.788	[-0.244, 0.321]
C(departmentsent)[T.13]	-0.561	0.003	[-0.933, -0.189]
C(departmentsent)[T.14]	-0.041	0.877	[-0.564, 0.482]
C(departmentsent)[T.15]	0.374	0.374	[-0.451, 1.199]
C(departmentsent)[T.17]	0.255	0.808	[-1.801, 2.311]
C(labtype)[T.2]	-0.028	0.129	[-0.064, 0.008]
C(specimencity)[T.2]	-0.05	0.005	[-0.085, -0.015]
C(specimencity)[T.3]	0.014	0.497	[-0.027, 0.055]
C(priority_level)[T.2]	0.33	0.259	[-0.243, 0.902]
C(priority_level)[T.3]	0.39	0.181	[-0.181, 0.962]
daily_case_count	-0.001	0.075	[-0.001, 0.000]
recv_week_of_year	0.001	0.015	[0.000, 0.002]
recv_day_of_week	0.021	< 0.001	[0.014, 0.028]

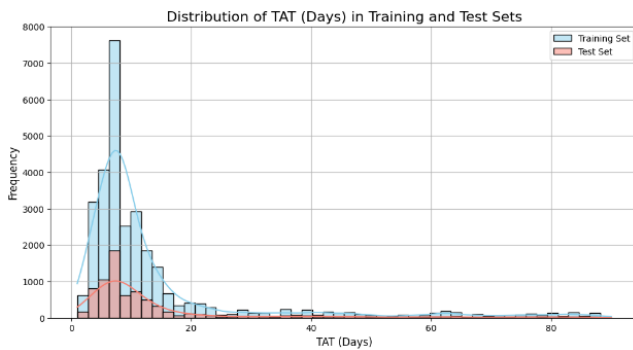
คำอธิบายในตาราง 4 Intercept (ค่าคงที่) หมายถึง ค่า Logarithm ของจำนวน TAT ที่คาดการณ์ เมื่อตัวแปรอิสระอื่นๆ ทั้งหมดมีค่าเป็นศูนย์ หรืออยู่ในกลุ่มฐานอ้างอิง (Reference Category) C(Variable)[T.x] เป็นสัญกรณ์จาก Library ทางสถิติ (เช่น statsmodels ใน Python) ที่ใช้แสดงผลของตัวแปรเชิงคุณภาพ หมายถึง ผลกระทบ (ค่าสัมประสิทธิ์) ของตัวแปร (เช่น specimen_type) ที่มีค่าเท่ากับ x โดยเปรียบเทียบกับ "ค่าฐานอ้างอิง" (Reference Level) ที่แบบจำลองกำหนดไว้โดยอัตโนมัติ

การตั้งค่าการทดลอง

ชุดข้อมูลที่ทำผ่านการทำความสะอาดแล้วถูกแบ่งออกเป็นชุดข้อมูลสำหรับฝึกสอน (Training Set) 80% และชุดข้อมูลสำหรับทดสอบ (Test Set) 20% และการแสดงการกระจายตัวของประชากรในชุดข้อมูลฝึกสอนและทดสอบ ดังตาราง 5 และภาพ 3

ตารางที่ 5 ค่าสถิติของ TAT ในชุดข้อมูลฝึกสอนและทดสอบ

ตัวแปร	Total	Avg	SD	Min	Q1 (25%)	Median (50%)	Q3 (75%)	Max
Training Set	30,012	14.19	16.64	1	6	8	13	90
Test Set	7,503	14.28	16.7	1	6	8	13	90



รูปที่ 3 การกระจายตัวของประชากรในชุดข้อมูลฝึกสอนและทดสอบ พบว่ามีการกระจายตัวแบบเบ้ขวา

ส่วนประสิทธิภาพของแบบจำลองถูกประเมินโดยใช้มาตรวัดมาตรฐาน 4 ชนิด ได้แก่ Mean Absolute Error (MAE) คือ ค่าเฉลี่ยของความคลาดเคลื่อนสัมบูรณ์ (หน่วยเป็นวัน) เป็นมาตรวัดหลักที่เข้าใจง่ายและสะท้อนถึงความผิดพลาดโดยเฉลี่ยในการพยากรณ์ Mean Squared Error (MSE) คือ ค่าเฉลี่ยของความคลาดเคลื่อนกำลังสอง (หน่วยเป็น วัน²) ซึ่งจะให้น้ำหนักกับความผิดพลาดขนาดใหญ่มากกว่า Root Mean Squared Error (RMSE) คือ รากที่สองของ MSE (หน่วยเป็นวัน) และ R-squared (R²) คือ ค่าสัมประสิทธิ์การตัดสินใจ ซึ่งบ่งชี้สัดส่วนของความแปรปรวนในตัวแปรตาม (TAT) ที่สามารถอธิบายได้ด้วยปัจจัยในแบบจำลอง (ค่าเข้าใกล้ 1 หมายถึงประสิทธิภาพสูง) นอกจากนี้ได้ทำการปรับจูนไฮเปอร์พารามิเตอร์ (Hyperparameter Tuning) สำหรับ Random Forest และ LightGBM โดยใช้เทคนิค GridSearchCV ร่วมกับ Cross-validation เพื่อค้นหาชุดค่าพารามิเตอร์ที่ให้ผลลัพธ์ที่ดีที่สุดดังแสดงในตาราง 6

ตารางที่ 6 Hyperparameter ของโมเดลที่ใช้ในการทดลอง

Model Type	Model Version	Hyperparameters
Random Forest Regressor	Baseline (ค่าเริ่มต้น)	• n_estimators: 100 • max_depth: None • min_samples_split: 2 • min_samples_leaf: 1 • random_state: 42 • n_jobs: -1
	Tuned - Best Features	• n_estimators: 200 • max_depth: None • min_samples_leaf: 2 • min_samples_split: 2 • max_features: 'sqrt' • random_state: 42 • n_jobs: -1
	Tuned - Statistically Significant Features	• n_estimators: 200 • max_depth: None • min_samples_leaf: 2 • min_samples_split: 2 • max_features: 'sqrt' • random_state: 42 • n_jobs: -1
LightGBM Regressor	Default (ค่าเริ่มต้น)	• boosting_type: 'gbdt' • n_estimators: 100 • learning_rate: 0.1 • num_leaves: 31 • max_depth: -1 • min_child_samples: 20 • subsample: 1.0 • colsample_bytree: 1.0 • random_state: 42 • n_jobs: -1
	Tuned - Statistically Significant Features	• learning_rate: 0.05 • max_depth: -1 • min_child_samples: 20 • n_estimators: 200 • num_leaves: 64 • subsample: 0.8 • colsample_bytree: 0.8 • random_state: 42 • n_jobs: -1

ผลการทดลอง

แบบจำลอง Random Forest ที่ผ่านการปรับจูนไฮเปอร์พารามิเตอร์และใช้ชุดปัจจัยที่คัดเลือกจากนัยสำคัญทางสถิติ (Tuned Random Forest with Statistically Significant Features) ให้ผลลัพธ์ที่ดีที่สุด โดยสามารถพยากรณ์ TAT บนชุดข้อมูลทดสอบได้ด้วยค่า Mean Absolute Error (MAE) เฉลี่ยเพียง 1.85 วัน ซึ่งหมายความว่าโดยเฉลี่ยแล้ว แบบจำลองสามารถพยากรณ์ระยะเวลาผลคลาดเคลื่อนไปไม่เกิน 2 วัน และมีค่า R-squared (R²) สูงถึง 0.91 บ่งชี้ว่าปัจจัยที่เลือกมาในแบบจำลองสามารถอธิบายความแปรปรวนของ TAT ได้ถึง 91% สะท้อนถึงความสามารถในการจับรูปแบบและความสัมพันธ์ที่ซับซ้อนในข้อมูลได้อย่างดี

การเปรียบเทียบประสิทธิภาพ

การเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ในตาราง 7 ได้แสดงให้เห็นถึงความสำคัญอย่างยิ่งของการเตรียมข้อมูล, การสร้างปัจจัย (Feature Engineering), และการปรับ

จูนไฮเปอร์พารามิเตอร์ (Hyperparameter Tuning) ซึ่งส่งผลโดยตรงต่อประสิทธิภาพของแบบจำลองในการพยากรณ์ TAT สรุปผลตามลำดับขั้นตอนการพัฒนาได้ 4 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 (Baseline) พบว่าแค่การทำความสะอาดข้อมูล (Cleaned Data) ก็ช่วยลดค่า MAE (ค่าเฉลี่ยความคลาดเคลื่อน) ของแบบจำลอง Random Forest (RF) ลงอย่างก้าวกระโดดจาก 27.91 วัน เหลือ 9.10 วัน ต่อมาในขั้นตอนที่ 2 (Simple Feature Engineering) ได้เพิ่มปัจจัย daily_case_count เพียงตัวเดียว ช่วยเพิ่มประสิทธิภาพของแบบจำลอง RF ต่อไปอีก โดยลด MAE ลงเหลือ 7.20 วัน จากนั้นในขั้นตอนที่ 3 (Full Feature Engineering) การใช้ชุดปัจจัยที่สร้างขึ้นใหม่ทั้งหมดซึ่งครอบคลุมบริบทการทำงาน (เช่น พยาธิแพทย์, เวลา, ภาระงาน, ความซับซ้อน) ได้เพิ่มประสิทธิภาพอย่างก้าวกระโดดทำให้ MAE ของ RF ลดลงเหลือเพียง 2.42 วัน (Rm เพิ่มเป็น 0.89) ซึ่งแสดงให้เห็นว่าปัจจัยที่สร้างขึ้นมีความสามารถในการพยากรณ์สูงมาก โดยผลการทดลองยังชี้ให้เห็นว่าการเพิ่มปัจจัยที่ซับซ้อนขึ้น (เช่น rolling_avg, holiday) ไม่ได้ช่วยให้แบบจำลองดีขึ้น (MAE เพิ่มขึ้น 2.50) และในขั้นตอนสุดท้าย (Final Model Tuning) การปรับจูนไฮเปอร์พารามิเตอร์โดยใช้ชุดปัจจัยที่มีนัยสำคัญทางสถิติ (จากตาราง 4) ช่วยเพิ่มความแม่นยำของแบบจำลอง RF ให้ดีที่สุด โดยลด MAE ลงจาก 2.42 วัน เหลือ 1.85 วัน (Rm = 0.91) เมื่อเปรียบเทียบกับอัลกอริทึมในขั้นสุดท้ายนี้ พบว่า Random Forest ให้ประสิทธิภาพที่ดีกว่า LightGBM อย่างมีนัยสำคัญ (MAE 1.85 วัน เทียบกับ 3.39 วันบน Test Set)

ตารางที่ 7 การเปรียบเทียบประสิทธิภาพของแบบจำลองในแต่ละขั้นตอน

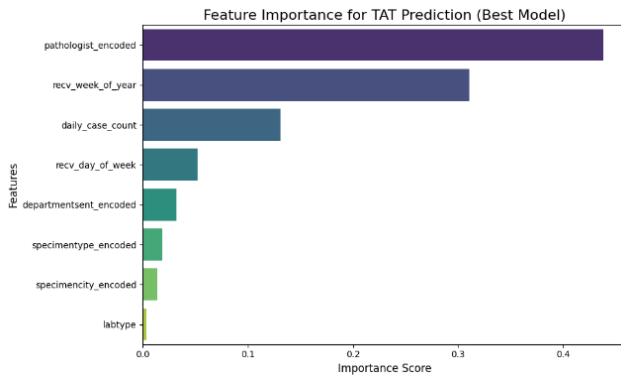
ขั้นตอนการทดลอง (Experiment Stage)	แบบจำลองและชุดปัจจัย (Model and Feature Set)	คำอธิบาย (Description)	M AE	MS E	RM SE	R ²
ขั้นตอนที่ 1: Baseline เบื้องต้น	Random Forest (Baseline on Raw Data)	แบบจำลอง Random Forest (RF) ที่ฝึกสอนด้วยข้อมูลดิบที่ยังไม่ผ่านการทำความสะอาด	27.91	-	-	-
	Random Forest (Baseline on Cleaned Data)	แบบจำลอง RF ที่ฝึกสอนด้วยข้อมูลที่ผ่านการทำความสะอาดแล้ว แต่ยังไม่มีสร้างปัจจัยใหม่	9.10	209.93	14.49	0.22
ขั้นตอนที่ 2: การสร้างปัจจัยอย่างง่าย	Random Forest +	แบบจำลอง RF ที่เพิ่มปัจจัยด้านภาระงาน	7.20	157.10	12.53	0.42

(Simple Feature Engineering)	daily_case_count	(daily_case_count) 1 ตัว				
	LightGBM + daily_case_count	แบบจำลอง LightGBM (LGBM) ที่ใช้ปัจจัยเหมือนกันเพื่อเปรียบเทียบอัลกอริทึม	7.81	162.42	12.74	0.40
ขั้นตอนที่ 3: การสร้างปัจจัยเต็มรูปแบบ (Full Feature Engineering)	Random Forest (BEST Before Tune)	แบบจำลอง RF ที่ใช้ชุดปัจจัยที่ดีที่สุด (พยาธิแพทย์, ภาระงาน, เวลา, ประเภทสิ่งส่งตรวจ) โดยยังไม่ปรับจูน Hyperparameters	2.42	31.21	5.59	0.89
	Random Forest (All Features incl. rolling_avg + holiday)	แบบจำลอง RF ที่ทดลองเพิ่มปัจจัยเชิงลึก (ค่าเฉลี่ยย้อนหลัง, วันหยุด)	2.50	33.07	5.75	0.88
	LightGBM (Default Params)	แบบจำลอง LGBM ที่ใช้ชุดปัจจัยที่ดีที่สุด (เหมือน RF BEST) โดยใช้ค่าพารามิเตอร์เริ่มต้น	3.95	46.73	6.84	0.83
ขั้นตอนที่ 4: การปรับจูนขั้นสุดท้าย (Final Model Tuning)	Random Forest (Tuned - Best Features)	แบบจำลอง RF จากขั้นตอนที่ 3 ที่ผ่านการปรับจูน Hyperparameters	1.92	24.49	4.95	0.91
	Random Forest (Tuned - Statistically Significant Features)	(แบบจำลองที่ดีที่สุด 1): RF ที่ปรับจูนโดยใช้เฉพาะชุดปัจจัยที่มีนัยสำคัญทางสถิติ	1.85	26.48	5.15	0.91
	LGBM (Tuned - Statistically Significant Features) (Test Set)	(แบบจำลองที่ดีที่สุด 2): LGBM ที่ปรับจูนโดยใช้ชุดปัจจัยที่มีนัยสำคัญทางสถิติ (แสดงผลบน Test Set)	3.39	37.25	6.10	0.87

การศึกษาความสำคัญของปัจจัย

จากการวิเคราะห์ความสำคัญของปัจจัย (Feature Importance) ของแบบจำลอง Random Forest ที่ดีที่สุด พบว่าปัจจัยที่มีอิทธิพลต่อการพยากรณ์ TAT มากที่สุด (MAE $\approx$ 1.85 วัน) พบว่าปัจจัยที่มีอิทธิพลต่อการพยากรณ์ TAT มากที่สุดเรียงตามลำดับจากมากไปน้อย ได้แก่ pathologist_encoded (พยาธิแพทย์ผู้รับผิดชอบ), recv_week_of_year (สัปดาห์ของปีที่ได้รับเคส), daily_case_count (จำนวนเคสที่ได้รับในแต่ละวัน), recv_day_of_week (วันในสัปดาห์ที่ได้รับเคส), department_sent_encoded (แผนกที่ส่งเคส), specimen_type_encoded

(ประเภทของสิ่งส่งตรวจ), specimencity_encoded (ความซับซ้อนของสิ่งส่งตรวจ) และ labtype (ประเภทของห้องปฏิบัติการ) ตามลำดับ (รูป 4) ผลลัพธ์นี้สอดคล้องกับการวิเคราะห์ทางสถิติเบื้องต้น และยืนยันว่าปัจจัยที่เกี่ยวข้องกับ บุคคล (พยาธิแพทย์), เวลา (สัปดาห์/วันของปี), และภาระงาน (จำนวนเคสรายวัน) เป็นตัวขับเคลื่อนหลักของความผันผวนในระยะเวลา รอผล



รูปที่ 4 ความสำคัญของปัจจัยต่อการพยากรณ์ TAT จากแบบจำลอง Random Forest ที่ดีที่สุด

อภิปรายผล

ผลการศึกษาพบว่าแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) โดยใช้ Random Forest สามารถพยากรณ์ระยะเวลาการรอผล (TAT) ในห้องปฏิบัติการพยาธิวิทยาภายในภาคได้อย่างแม่นยำ โดยแบบจำลองชี้ให้เห็นว่า TAT เป็นผลลัพธ์ของปฏิสัมพันธ์ที่ซับซ้อนระหว่างปัจจัยหลายมิติ ปัจจัยสำคัญที่มีอิทธิพลมากที่สุดคือ ปัจจัยด้านบุคคล โดยพยาธิแพทย์ผู้รับผิดชอบเป็นตัวแปรสำคัญ ซึ่งสะท้อนถึงความสำคัญของ "ปัจจัยมนุษย์" ในกระบวนการทำงาน ตามมาด้วยปัจจัยด้านเวลา ได้แก่ recv_day_of_week (วันของสัปดาห์) และ recv_week_of_year (สัปดาห์ของปี) ซึ่งมีผลต่อรูปแบบภาระงานอย่างชัดเจน โดยพบว่าวันอังคารและวันศุกร์มีค่าเฉลี่ย TAT สูงสุด อาจเกิดจากการสะสมงานก่อนและหลังวันหยุดสุดสัปดาห์ ขณะที่วันจันทร์มี TAT ต่ำที่สุด (ดังแสดงในตาราง 8) นอกจากนี้ ยังพบช่วงที่ TAT สูงขึ้นในสัปดาห์ที่ 19-23 (พฤษภาคม-มิถุนายน) และสัปดาห์ที่ 44 (ปลายตุลาคม) ซึ่งสัมพันธ์กับการมีวันหยุดต่อเนื่องหลายวัน (เช่น วันฉัตรมงคล, วันวิสาขบูชา, วันปิยมหาราช ดังแสดงในตาราง 9) และยังอาจสัมพันธ์กับบริบททางสังคมของจังหวัดสกลนครที่เป็นพื้นที่เกษตรกรรม ซึ่งตรงกับฤดูเพาะปลูกและเก็บเกี่ยว อันอาจส่งผลต่อพฤติกรรมกรรมการรับบริการหรือการลาของบุคลากร สุดท้ายคือ ปัจจัยด้านภาระงาน เช่น daily_case_count (จำนวนเคสที่ได้รับในแต่ละวัน) ซึ่งยืนยันอย่างเป็น

รูปธรรมว่าปริมาณงานมีผลต่อ TAT อย่างมีนัยสำคัญ ผลการศึกษานี้จึงบ่งชี้ว่าแบบจำลองการเรียนรู้ สามารถจับความสัมพันธ์ที่ซับซ้อนเหล่านี้ได้อย่างมีประสิทธิภาพ และมีศักยภาพสูงในการนำไปประยุกต์ใช้เพื่อปรับปรุงการบริหารจัดการห้องปฏิบัติการให้มีประสิทธิภาพมากขึ้นในอนาคต

ตารางที่ 8 ค่าเฉลี่ย TAT ในแต่ละวันของสัปดาห์

วันในสัปดาห์	ค่าเฉลี่ย TAT (วัน)
จันทร์	11.69
อังคาร	15.08
พุธ	13.3
พฤหัสบดี	14.2
ศุกร์	16.54
เสาร์	13.83
อาทิตย์	14.13

ตารางที่ 9 ค่าเฉลี่ย TAT สูงสุด 5 อันดับแรกแยกตามสัปดาห์ของปี

สัปดาห์ที่	ค่าเฉลี่ย TAT (วัน)
20	28.63
23	21.77
19	21.36
21	20.86
44	20.71

นัยสำคัญและการประยุกต์ใช้

แบบจำลองที่พัฒนาขึ้นมีศักยภาพในการนำไปประยุกต์ใช้เป็นเครื่องมือสนับสนุนการตัดสินใจในการบริหารจัดการห้องปฏิบัติการ โดยสามารถเปลี่ยนกระบวนการจากเชิงรับไปสู่เชิงรุกได้ สามารถใช้แบบจำลองนี้เพื่อพยากรณ์ภาระงานล่วงหน้า ระบุเคสที่มีความเสี่ยงล่าช้า และปรับปรุงประสิทธิภาพกระบวนการทำงาน การนำแบบจำลองไปใช้ในระบบสารสนเทศของโรงพยาบาล (Hospital Information System: HIS) จะช่วยยกระดับการบริหารจัดการที่ขับเคลื่อนด้วยข้อมูล (Data-Driven Management) และส่งผลต่อคุณภาพการบริการผู้ป่วย

ข้อจำกัดของการศึกษา

แม้ว่าผลการศึกษาเป็นที่น่าพอใจ แต่มีข้อจำกัดบางประการที่ต้องพิจารณา เช่น การศึกษาในสถาบันเดียว (Single-center study) การนำไปใช้กับโรงพยาบาลอื่นยังต้องได้รับการตรวจสอบเพิ่มเติม และปัจจัยอื่นๆ ที่อาจส่งผลต่อ TAT แต่ไม่ได้ถูกบันทึกไว้ในชุดข้อมูลนี้ เช่น จำนวนบล็อกชิ้นเนื้อ จำนวนวันลาของบุคลากร เครื่องมือชำรุด เหตุการณ์ที่ไม่คาดคิด เช่น ไฟไหม้ น้ำประปาไม่ไหล ยังไม่ได้ถูกนำมาพิจารณา

สรุป

การศึกษานี้ได้การพัฒนาแบบจำลอง Random Forest ซึ่งสามารถพยากรณ์ระยะเวลารอผล (TAT) ของสิ่งส่งตรวจทางพยาธิวิทยาภายในภาคได้อย่างแม่นยำ โดยมีค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) เพียง 1.85 วัน ผลการวิเคราะห์ปัจจัยสำคัญ (Feature Importance) ชี้ชัดว่าพยาธิแพทย์ผู้รับผิดชอบ, ช่วงเวลาของปี (recv_week_of_year), ภาระงานรายวัน (daily_case_count), และวันในสัปดาห์ที่รับส่งตรวจ (recv_day_of_week) เป็นปัจจัยที่มีอิทธิพลต่อ TAT มากที่สุด ซึ่งสะท้อนถึงความซับซ้อนของปฏิสัมพันธ์ระหว่างปัจจัยด้านบุคคล, เวลา, และปริมาณงานในกระบวนการทางพยาธิวิทยา แบบจำลองที่พัฒนาขึ้นนี้มีศักยภาพสูงในการนำไปประยุกต์ใช้เพื่อสนับสนุนการตัดสินใจและเพิ่มประสิทธิภาพการทำงานของห้องปฏิบัติการได้หลายด้าน ตัวอย่างเช่น การจัดลำดับความสำคัญของเคส โดยระบบ LIS สามารถใช้ค่าพยากรณ์ TAT เพื่อแจ้งเตือนเคสที่อาจล่าช้าให้ผู้รับผิดชอบพิจารณาเร่งรัด, การบริหารจัดการทรัพยากร โดยใช้แนวโน้ม TAT ที่พยากรณ์ได้มาวางแผนการจัดสรรกำลังคนและเครื่องมือให้เหมาะสมกับภาระงานที่คาดการณ์, การสื่อสารเชิงรุกโดยแจ้งค่า TAT โดยประมาณแก่แพทย์เจ้าของไข้ผ่านระบบ HIS เพื่อช่วยในการวางแผนการรักษาและลดการสอบถาม, และการระบุคอขวด โดยวิเคราะห์ปัจจัยสำคัญเพื่อชี้เป้าปัญหาในกระบวนการทำงานและวางแผนการปรับปรุงแก้ไขได้อย่างตรงจุด สำหรับงานวิจัยในอนาคต ควรมีการนำแบบจำลองนี้ไปผนวกเข้ากับระบบสารสนเทศของโรงพยาบาลเพื่อทดสอบประสิทธิภาพในสถานการณ์การทำงานจริง, ประเมินความสามารถในการนำไปปรับใช้ในโรงพยาบาลอื่น, และพิจารณารวบรวมข้อมูลที่มีความละเอียดมากขึ้นเพื่อพัฒนาความแม่นยำในการพยากรณ์ต่อไป

REFERENCES

- [1] M. G. Hanna et al., "Recommendations for performance evaluation of machine learning in pathology: A concept paper from the college of american pathologists," Arch Pathol Lab Med, vol. 148, no. 10, pp. e335–e361, Oct. 2024.
- [2] A. Frewing, A. B. Gibson, R. Robertson, P. M. Urie, and D. Della Corte, "A systematic review of machine learning for prostate cancer detection in pathology," Arch Pathol Lab Med, vol. 148, no. 5, pp. 603–612, May 2024.
- [3] B. Breil, F. Fritz, V. Thiemann, and M. Dugas, "Mapping turnaround times (TAT) to a generic timeline: A systematic review of TAT definitions in clinical domains," BMC Med Inform Decis Mak, vol. 11, no. 1, p. 34, May 2011.
- [4] K. E. Volmar, M. O. Idowu, R. J. Souers, D. S. Karcher, and R. E. Nakhleh, "Turnaround time for large or complex specimens in surgical pathology: A college of american pathologists Q-probes study of 56 institutions," Arch Pathol Lab Med, vol. 139, no. 2, pp. 171–177, Feb. 2015.
- [5] P. S. B. D. D. C. Martins, R. A. C. Leite, and G. C. D. S. Lamb, "Anxiety and depression in the preoperative period of cardiac surgery," Revista Brasileira de Cirurgia Cardiovascular, vol. 31, pp. 243–251, Jun. 2016.
- [6] T. O. S. Junior, R. de C. S. A. M. de Deus, and P. S. F. de Sousa, "A systematic review of the literature on the use of lean healthcare in the radiology sector," Journal of Health Management, vol. 26, no. 2, pp. 297–313, Apr. 2024.
- [7] พ. วรณไกรโรจน์, "ภาพรวมของมาตรฐานห้องปฏิบัติการพยาธิวิทยาภายในภาค," ราชวิทยาลัยพยาธิแพทย์แห่งประเทศไทย, 2560.
- [8] T. S. Ahluwalia et al., "The need for computational pathology in resource-limited settings," JCO Global Oncology, no. 8, p. e2200044, Aug. 2022.
- [9] G. Acs, "Telepathology for resource-limited settings," J. Am. Soc. Cytopathol., vol. 11, no. 5, pp. 291–304, Sep. 2022.

- [10] D. S. McClintock, "Informatics and digital pathology," *Pathobiology*, vol. 83, no. 2–3, pp. 73–82, 2016.
- [11] S. E. D'Alfonso and A. A. T. M. Reinders, "Artificial intelligence for digital pathology: a review of the literature and overview of a single-institution experience," *Seminars in Diagnostic Pathology*, vol. 40, no. 7, pp. 524–532, Nov. 2023.
- [12] J. H. Harrison et al., "Introduction to artificial intelligence and machine learning for pathology," *Arch Pathol Lab Med*, vol. 145, no. 10, pp. 1228–1254, Oct. 2021.
- [13] M. Z. Alom et al., "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, p. 292, Mar. 2019.
- [14] A. van der Byl et al., "Artificial intelligence in diagnostic pathology: a review of the literature and overview of a single-institution experience," *Seminars in Diagnostic Pathology*, vol. 40, no. 7, pp. 524–532, Nov. 2023.
- [15] A. B. Nassif, O. Al-Dahidi, S. Al-Ali, and M. A. Talib, "The battle of machine learning and deep learning in healthcare: A survey," *Journal of Big Data*, vol. 11, no. 1, p. 95, Dec. 2024.
- [16] N. M. Abdul Latiff et al., "Voice pathology detection using machine learning algorithms based on different voice databases," *Results in Engineering*, vol. 25, p. 103937, Mar. 2025.
- [17] R. M. T. Gebru, J. A. M. da Silva, A. M. da S. A. Filho, J. M. da S. A. Junior, and S. R. L. da Silva, "Turnaround time analysis using data mining from a laboratory information system," *The Journal of Health Care Organization, Provision, and Financing*, vol. 62, Jan. 2025.
- [18] I. Karaca, "A comprehensive review of the evolution of deep learning: a recent survey," *Multimedia Tools and Applications*, vol. 83, no. 17, pp. 50921–50974, Jun. 2024.
- [19] S. Reddy et al., "Predicting chronic kidney disease progression using small pathology datasets and explainable machine learning models," *Comput Methods Programs Biomed Updat*, vol. 6, p. 100160, 2024.
- [20] A. van der Byl, D. van der Byl, P. Cooper, A. van der Merwe, J. Schneider, and J. M. Warwick, "The operationalisation of artificial intelligence in anatomical pathology," *J Clin Pathol*, vol. 76, no. 4, pp. 225–231, Apr. 2023.