

ระบบคัดกรองความเสี่ยงของโรคธาลัสซีเมียเบื้องต้นด้วยปัญญาประดิษฐ์โดยการประมวลผลจากภาพเซลล์เม็ดเลือด เพื่อเพิ่มการเข้าถึงบริการสาธารณสุข ในระดับท้องถิ่น
An AI-Based Preliminary Thalassemia Risk Screening System Using Blood Cell Image Analysis to Improve Local Healthcare Accessibility

ชวัลรัตน์ สุธรรมา สิริญาณรณ การันต์
ณัฐภัตสร เหล่าเนตร์¹ ณัฐชนน รัตนพิทักษ์กุล²

¹โรงเรียนวิทยาศาสตร์จุฬาราชวิทยาลัย พิษณุโลก 86 หมู่ 4 ตำบลมะขามสูง อำเภอเมือง จังหวัดพิษณุโลก 65000

²โรงพยาบาลพุทธชินราช พิษณุโลก 90 ถนนศรีธรรมไตรปิฎก ตำบลในเมือง อำเภอเมืองพิษณุโลก จังหวัดพิษณุโลก 65000

Email: chawanrat.sut@pccpl.ac.th

บทคัดย่อ

โครงการเรื่องระบบคัดกรองความเสี่ยงของโรคธาลัสซีเมียเบื้องต้นด้วยปัญญาประดิษฐ์ โดยการประมวลผลจากภาพเซลล์เม็ดเลือด เพื่อเพิ่มการเข้าถึงบริการสาธารณสุขในระดับท้องถิ่น มีวัตถุประสงค์เพื่อพัฒนาและทดสอบประสิทธิภาพระบบปัญญาประดิษฐ์บูรณาการในการคัดกรองผู้ป่วยกลุ่มเสี่ยงของโรคธาลัสซีเมีย สืบเนื่องมาจาก โรคธาลัสซีเมียเป็นโรคที่เกิดจากการถ่ายทอดทางพันธุกรรม ที่เซลล์เม็ดเลือดแดงทำงานผิดปกติ หากไม่ได้รับการรักษาจะส่งผลให้ผู้ป่วยมีภาวะโลหิตจางเรื้อรังหรือเกิดภาวะแทรกซ้อนรุนแรง การได้รับการคัดกรองโรคอย่างรวดเร็วจึงมีความจำเป็นอย่างยิ่ง โดยคณะผู้จัดทำมีวิสัยทัศน์การ ดังนี้ 1) การพัฒนาระบบการเก็บข้อมูลสถิติ ประกอบด้วยการใช้ปัญญาประดิษฐ์สำหรับจำแนกชนิดความผิดปกติของเม็ดเลือดแดง 9 ชนิดและปัญญาประดิษฐ์สำหรับการคัดกรองคุณภาพภาพโดยเปรียบเทียบระหว่าง YOLOv8/11n-cls และ DenseNet-121 ซึ่งพบว่า YOLOv11n-cls ในโมเดลคัดกรองคุณภาพของภาพมีความแม่นยำร้อยละ 99.18 และโมเดลจำแนกชนิดความผิดปกติของเม็ดเลือดแดงได้ความแม่นยำสูงสุด ร้อยละ 97.57 รองลงมา คือ YOLOv8n-cls และ DenseNet-121 ตามลำดับ รวมถึงใช้การประมวลผลภาพสำหรับข้อมูลเชิงปริมาณอื่นๆ 2) การพัฒนาโมเดลจำแนกกลุ่มโรคจากข้อมูลสถิติโดยเปรียบเทียบประสิทธิภาพของ 4 โมเดล ได้แก่ XGBoost, KNN, SVM และ Random Forest ซึ่งผลลัพธ์พบว่าโมเดล XGBoost ให้ค่าความแม่นยำรวม (Accuracy) สูงสุดที่ ร้อยละ 98.00 และให้ค่าความระลึก (Recall) ในกลุ่มเสี่ยงสูงถึง ร้อยละ 97.00 รองลงมาคือ Random Forest, SVM และ KNN ตามลำดับ 3) การทดสอบระบบรวมบนโปรแกรมต้นแบบ (Pipeline Application Prototype) สามารถทำนายและคัดกรองความเสี่ยงโรคธาลัสซีเมียในภาพรวมได้ความแม่นยำร้อยละ 96.00 และมีค่า Recall เท่ากับร้อยละ 97.00 ซึ่งจะช่วยเสริมศักยภาพในการคัดกรองโรคธาลัสซีเมีย โดยเฉพาะอย่างยิ่งในพื้นที่ห่างไกลที่ขาดแคลนทรัพยากรทางการแพทย์ ให้สามารถเข้าถึงการวินิจฉัยที่มีประสิทธิภาพและครอบคลุมทุกภูมิภาค เพื่อยกระดับคุณภาพชีวิตของประชาชนอย่างยั่งยืน

คำสำคัญ: โรคธาลัสซีเมีย, ภาวะโลหิตจางเรื้อรัง, ปัญญาประดิษฐ์

1. บทนำ

โรคธาลัสซีเมีย (Thalassemia) เป็นโรคที่เกิดจากการถ่ายทอดทางพันธุกรรม ที่เซลล์เม็ดเลือดแดงทำงานผิดปกติซึ่งหากไม่ได้รับการรักษาจะส่งผลให้

ผู้ป่วยมีภาวะโลหิตจางเรื้อรังหรือเกิดภาวะแทรกซ้อนรุนแรงได้ (Medparkhospital, 2566) ในปัจจุบันมีผู้ป่วยโรคธาลัสซีเมียในประเทศไทยมากกว่า 800,000 คนรวมถึงมีผู้ที่เป็นพาหะประมาณ ร้อยละ 30-40 ของ

ประชากรทั้งหมดยังต้องการตรวจวินิจฉัยโรคธาลัสซีเมีย (วารสารโรงพยาบาลนครพิงค์, 2568) ซึ่งในปัจจุบันมีวิธีการที่ซับซ้อนทำให้ประชาชนในพื้นที่ห่างไกลเข้าถึงการตรวจคัดกรองได้ยาก

คณะผู้วิจัยจึงมีแนวคิดในการพัฒนาระบบคัดกรองความเสี่ยงโรคธาลัสซีเมียเบื้องต้นด้วยปัญญาประดิษฐ์ โดยแบ่งการดำเนินงานออกเป็น 3 ส่วนหลัก ส่วนแรกคือการพัฒนากระบวนการเก็บข้อมูลสถิติเบื้องต้นจากภาพถ่ายเซลล์เม็ดเลือดจากกล้องจุลทรรศน์ ซึ่งมีการประยุกต์ใช้เทคโนโลยีการประมวลผลภาพในการปรับสีภาพด้วยภาพอ้างอิงเพื่อปรับภาพจากต่างแหล่งที่มาให้เหมาะสมกับโมเดลร่วมกับการตรวจจับขอบเพื่อระบุตำแหน่งและตัดภาพ (Crop) เซลล์เม็ดเลือด ตลอดจนใช้โมเดลปัญญาประดิษฐ์ประเภทการจำแนกข้อมูลเพื่อคัดกรองคุณภาพเม็ดเลือดและจำแนกความผิดปกติของเซลล์เม็ดเลือดแดง สำหรับส่วนที่ 2 เป็นการพัฒนาระบบจำแนกผู้ป่วยออกจากคนปกติโดยนำข้อมูลสถิติจากส่วนแรกมาทำวิเคราะห์ และใช้เทคนิคการคัดเลือกคุณลักษณะด้วยการกำหนดน้ำหนักเพื่อหาตัวแปรที่มีผลต่อการตัดสินใจอย่างแท้จริง พร้อมทั้งเปรียบเทียบประสิทธิภาพของ 4 โมเดล ได้แก่ XGBoost, KNN, SVM และ Random Forest และส่วนที่ 3 จะนำระบบทั้งหมดมารวมกันเพื่อทดสอบประสิทธิภาพ

2. วิธีการดำเนินงาน

2.1 การเตรียมวัสดุและอุปกรณ์

นำชุดข้อมูลภาพสำหรับพัฒนาโมเดลจำแนกผู้ป่วยโรคธาลัสซีเมียประกอบด้วยข้อมูลสองส่วนหลัก โดยส่วนแรกเป็นการเตรียมข้อมูลภาพเซลล์เม็ดเลือดแดงของผู้ป่วยโรคธาลัสซีเมียและผู้ที่ไม่ได้โรค ซึ่งจัดเตรียมโดยผู้เชี่ยวชาญและกลุ่มนักเทคนิคการแพทย์จากโรงพยาบาลพุทธชินราชผ่านการใช้อัลกอริทึม และส่วนที่สองเป็นการนำเข้าชุดข้อมูลภาพเม็ดเลือดที่มีความผิดปกติจากงานวิจัย Erythrocyte (Red Blood Cell) Dataset in Thalassemia case ของ Dyah

Aruming Tyas และคณะ เพื่อใช้สำหรับการจำแนกชนิดความผิดปกติของเซลล์เม็ดเลือด

2.2 การพัฒนาระบบเก็บข้อมูลสถิติในภาพเม็ดเลือดด้วยการใช้ image processing และ การใช้โมเดล AI แบบ classification

2.2.1 พัฒนาระบบปรับสีภาพ

เตรียมชุดภาพและภาพอ้างอิงจากงานวิจัยแปลงปริภูมิสี BGR เป็น CIE L* A* B* เพื่อแยกความสว่างกับสี แล้วคำนวณค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของแต่ละ Channel ทั้งภาพอ้างอิงและภาพเป้าหมายเพื่อปรับปรุงค่าสีตามความสัมพันธ์ทางสถิติ จากนั้นจำกัดค่าสีในช่วง 0-255 แปลงกลับเป็น BGR ส่งออกเป็นภาพที่ถูกแปลง และนำระบบนี้ไปใช้กับชุดข้อมูลโรงพยาบาลพุทธชินราชทั้งหมด

2.2.2 พัฒนาระบบจับตำแหน่งเม็ดเลือดโดยใช้ image processing ด้วยการ edge และ โมเดลปัญญาประดิษฐ์ในการคัดกรองภาพเม็ดเลือด

เตรียมภาพเม็ดเลือดจากโรงพยาบาลพุทธชินราช 10 ภาพ ใช้ Canny Edge Detection ร่วมกับการเติมพื้นที่ปิดทำ Binary Mask เพื่อ crop และ save เซลล์เม็ดเลือด จากนั้นแบ่งชุดข้อมูลเป็น 2 class (ใช้ได้/ใช้ไม่ได้) สัดส่วน 70:20:10 (เทรน:ทดสอบ:ตรวจสอบ) เพื่อเปรียบเทียบประสิทธิภาพโมเดล YOLOv8-cl, YOLO11-cl และ densenet-121 แล้วนำระบบตรวจจับข้างต้นมาพัฒนาต่อโดยการเพิ่มการเก็บขนาดพื้นที่เม็ดเลือดเพื่อคำนวณค่า cv, sd และใช้โมเดลที่เลือกมาเป็นเงื่อนไขในการคัดกรองเพื่อ save

2.2.3 พัฒนาระบบการแบ่งระดับของ Hypochromic จากพื้นที่สีขาวส่วนกลางของเม็ดเลือด

นำภาพเม็ดเลือดที่คัดกรองแล้วมาแปลงเป็นระดับสีเทา ลดสัญญาณรบกวนด้วย Gaussian Blur แล้วใช้ Otsu's Thresholding แยกเซลล์ออกจากพื้นหลังเพื่อคำนวณพื้นที่ทั้งหมด จากนั้นตรวจจับพื้นที่สว่างส่วนกลางด้วยเกณฑ์ Adaptive จากค่าเฉลี่ยและค่าต่ำสุดในเซลล์ เพื่อคำนวณอัตราส่วนพื้นที่มืดต่อพื้นที่

เซลล์ และแบ่งระดับความรุนแรงเป็น Grade 0–4 ตามเกณฑ์ที่กำหนด

2.2.4 การพัฒนาโมเดลจำแนกความผิดปกติของเม็ดเลือด

นำชุดข้อมูลจากงานวิจัยมาแบ่งสัดส่วนแต่ละชนิดเป็น ร้อยละ 70 สำหรับการเทรน, ร้อยละ 20 สำหรับการทดสอบ และร้อยละ 10 สำหรับการตรวจสอบความถูกต้อง เพื่อนำไปใช้เปรียบเทียบประสิทธิภาพในการเทรนด้วยโมเดล YOLOv8n-cl, YOLOv11n-cl และ densenet-121

2.2.5 พัฒนาระบบเก็บข้อมูลสถิติ

บูรณาการอัลกอริทึมทุกขั้นตอนเข้าด้วยกัน โดยเริ่มจากปรับมาตรฐานสีและตรวจจับตำแหน่งเพื่อสกัดภาพเม็ดเลือดรายวัตถุ แล้วนำภาพที่ผ่านการคัดกรองด้วย AI มาวิเคราะห์จัดเก็บค่าสถิติ ได้แก่ ค่า SD และ CV ของขนาดเซลล์เพื่อดูความคงที่, สัดส่วนร้อยละของภาวะ Hypochromic (Grade 0–4) เพื่อประเมินความรุนแรง และสัดส่วนร้อยละของเม็ดเลือดผิดปกติทั้ง 9 ชนิด (Elliptocyte, SPERO bulat, TARGETSEL, Acanthocyte, Hypochromic, Normal, Pencil, Stomatocyte, Teardrop) เพื่อเป็นฐานข้อมูลสนับสนุนการวินิจฉัยทางการแพทย์

2.3 การพัฒนาระบบคัดกรองความเสี่ยงของโรคธาลัสซีเมียเบื้องต้น

2.3.1 การ data augmentation และการจำลองเพื่อสร้างข้อมูลเสมือนด้วยโครงข่าย CTGAN

แก้ปัญหาข้อมูลไม่สมดุลจากโรงพยาบาลพุทธชินราชที่มีผู้ป่วย 57 รายและคนปกติ 125 ราย โดยใช้ data augmentation หมุนภาพ 90, 180, 270 องศาให้ได้ภาพผู้ป่วย 228 ภาพและคนปกติ 500 ภาพ แล้วใช้ระบบสถิติแปลงเป็นข้อมูลตารางขนาด 228 และ 500 แถวตามลำดับ จากนั้นนำไปเทรนโมเดล CTGAN (2,000 Epochs, Batch size 64) เพื่อจำลองข้อมูลเสมือนใหม่ 1,000 แถว พร้อมตรวจสอบคุณภาพโดยเทียบสถิติรายคอลัมน์และตัวชี้วัดสากลกับข้อมูลจริง เมื่อพบว่าการกระจายตัวและความสัมพันธ์ใกล้เคียงกัน

จึงรวมชุดข้อมูลทั้งหมดเข้าด้วยกันเพื่อใช้เทรนโมเดลต่อไป

2.3.2 การคัดเลือกคุณลักษณะสำคัญและการวิเคราะห์ค่าน้ำหนัก

นำชุดข้อมูลที่สมดุลมารวมกัน คัดเลือกเฉพาะคุณลักษณะตัวเลขมาผ่าน StandardScaler เพื่อปรับสเกล จากนั้นใช้ Logistic Regression ร่วมกับเทคนิค RFECV แบบ 5-Fold คัดเลือกคุณลักษณะอัตโนมัติโดยตัดตัวแปรที่ส่งผลน้อยออกเพื่อหาตัวแปรที่ให้ค่า Accuracy สูงสุด แล้วนำค่าน้ำหนักสัมประสิทธิ์มาคิดค่าสัมบูรณ์เพื่อจัดลำดับและแปลงเป็นเปอร์เซ็นต์การส่งผลต่อโมเดล

2.3.3 การพัฒนาโมเดลประเมินความเสี่ยงของโรคธาลัสซีเมียเบื้องต้น

แบ่งชุดข้อมูลเป็นชุดทดสอบ 400 แถว และชุดเทรน 1600 แถว (อัตราส่วน 1:4) แล้วนำข้อมูลเข้าฝึกสอนบนสถาปัตยกรรมโมเดล 4 ชนิด ได้แก่ XGBoost, KNN, SVM และ Random Forest จากนั้นนำโมเดลทั้งหมดมาทดสอบจริงกับชุดทดสอบเพื่อเปรียบเทียบประสิทธิภาพหลังการเทรน

2.4 การพัฒนาโปรแกรมต้นแบบและระบบบูรณาการเพื่อการคัดกรองโรคธาลัสซีเมีย

2.4.1 การออกแบบสถาปัตยกรรมระบบและโครงสร้างโปรแกรม (System Architecture Design)

พัฒนาโปรแกรมต้นแบบ (Application Prototype) บูรณาการระบบประมวลผลภาพและประเมินความเสี่ยงเชิงสถิติเข้าด้วยกันในรูปแบบ Pipeline Architecture แบ่งเป็น 3 ส่วนหลักคือ ส่วนติดต่อผู้ใช้งาน สำหรับอัปโหลดภาพสไลด์เม็ดเลือดแดงขนาดใหญ่และแสดงภาพสกัด ตารางสถิติ กับเปอร์เซ็นต์ความเสี่ยง, ส่วนประมวลผลหลัก ทำหน้าที่ปรับมาตรฐานสีภาพในปริภูมิ CIE L* A* B* ใช้ Canny Edge สกัดเม็ดเลือดเดี่ยว คัดกรองภาพด้วย Classification คำนวณพื้นที่ที่ขีด และจำแนกความผิดปกติ 9 ชนิด, และส่วนสรุปผลและรายงานสถิติ ที่รวบรวมค่าสถิติตัวเลขทั้งหมด (SD, CV ขนาดเซลล์,

ร้อยละ Hypochromic เกเรด 0-4 และร้อยละ เม็ดเลือดที่ผิดปกติ 9 ชนิด) มาปรับสเกลด้วย StandardScaler แล้วส่งให้โมเดลประเมินความเสี่ยง เพื่อกำหนดผลสรุปการคัดกรอง

3. ผลการทดลอง

3.1 ผลการพัฒนากระบวนการระบุตำแหน่ง (crop) และการคัดกรองคุณภาพภาพเม็ดเลือดด้วย โมเดลปัญญาประดิษฐ์

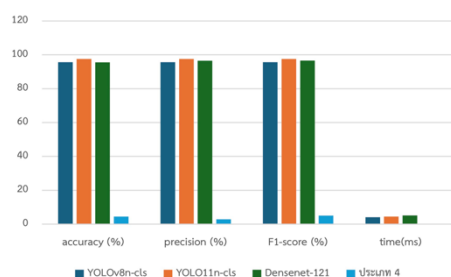
จากการทดลองเปรียบเทียบประสิทธิภาพ พบว่าโมเดล YOLO11n-cls ให้ประสิทธิภาพ สูงสุด โดยมีค่าความแม่นยำอยู่ที่ ร้อยละ 99.2 ค่า Precision ร้อยละ 99.2 และค่า F1-score ร้อยละ 99.1 ดังตารางที่ 1

ตารางที่ 1 เปรียบเทียบประสิทธิภาพของโมเดล

โมเดล	accuracy (%)	precision (%)	F1-score (%)	Time (ms)
YOLOv8n-cls	98.8	98.8	98.8	3.62
YOLO11n-cls	99.2	99.2	99.1	4.31
Densenet-121	96.6	96.5	96.6	4.34

3.2 ผลการพัฒนาโมเดล AI สำหรับการ จำแนกชนิดความผิดปกติของเม็ดเลือด

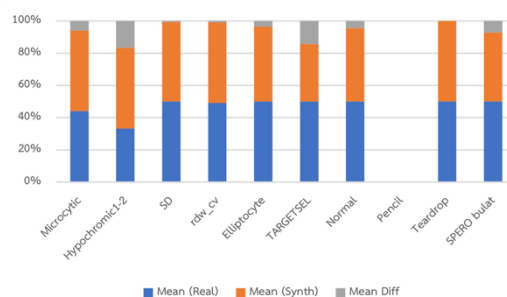
จากการศึกษาพบว่า YOLO11n-cls เป็น โมเดลที่มีประสิทธิภาพสูงสุด โดยมี Accuracy อยู่ที่ ร้อยละ 97.57 และ F1-score ร้อยละ 97.56 ซึ่งตรง ตามมาตรฐานทางการแพทย์ในขณะที่ยังคงความเร็ว ในการประมวลผลอยู่ที่ 4.42 ms ดังแผนภูมิที่ 1



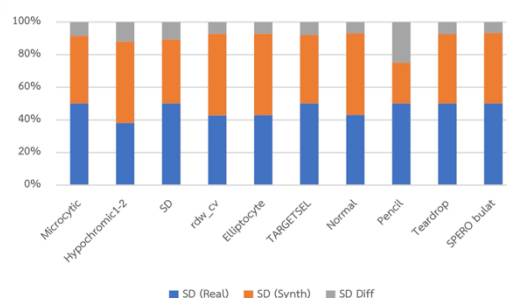
แผนภูมิที่ 1 ผลการเปรียบเทียบประสิทธิภาพของ โมเดลในการจำแนกความผิดปกติของเม็ดเลือด

3.3 ผลการตรวจสอบคุณภาพข้อมูลเสมือน ของชุดข้อมูลผู้ป่วยและคนปกติหลังการสร้างด้วย CTGAN

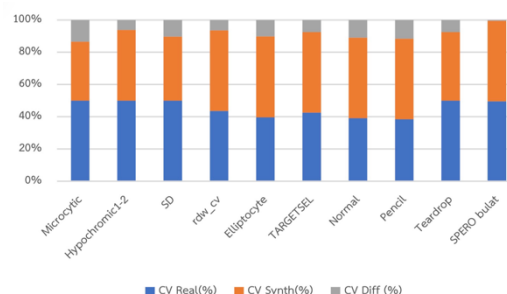
จากการทดลอง พบว่าโมเดลสามารถจำลอง รูปทรงการกระจายตัวของข้อมูลรายคอลัมน์ ได้แม่นยำ ร้อยละ 79.81 และรักษาแนวโน้มความสัมพันธ์ระหว่าง คู่ตัวแปรได้ที่ร้อยละ 83.56 ส่งผลให้ได้คะแนนความ คล้ายคลึงโดยเฉลี่ยร้อยละ 81.69 ดังแผนภูมิที่ 2, 3 และ 4



แผนภูมิที่ 2 การเปรียบเทียบค่าเฉลี่ยของชุดข้อมูลจริง และข้อมูลเสมือน



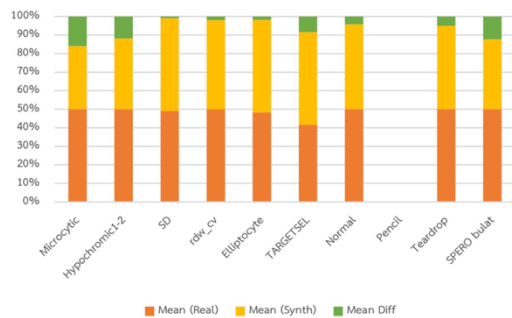
แผนภูมิที่ 3 การเปรียบเทียบส่วนเบี่ยงเบนมาตรฐาน ของชุดข้อมูลจริงและข้อมูลเสมือน



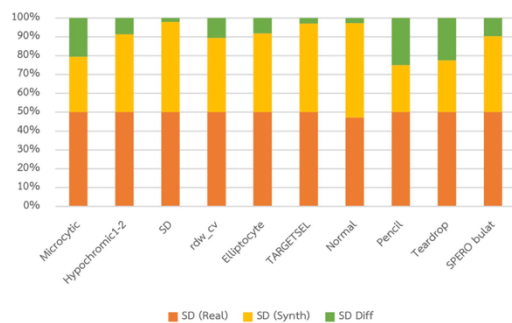
แผนภูมิที่ 4 การเปรียบเทียบสัมประสิทธิ์ความแปรผัน ของชุดข้อมูลจริงและข้อมูลเสมือน

3.4 การเปรียบเทียบค่าทางสถิติเชิงพรรณนาระหว่างชุดข้อมูลจริงและชุดข้อมูลเสมือนที่สร้างด้วย CTGAN ในกลุ่มผู้ป่วยโรคธาลัสซีเมีย

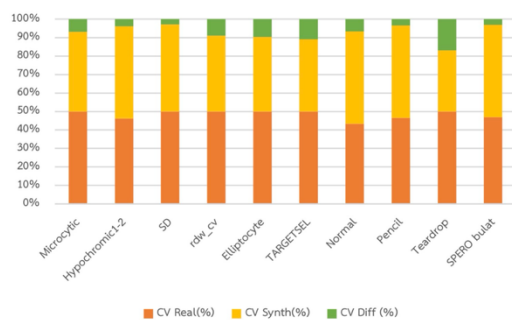
จากการศึกษา พบว่า โมเดลสามารถจำลองรูปทรงการกระจายตัวของข้อมูลรายคอลัมน์ ได้แม่นยำร้อยละ 82.89 และรักษาแนวโน้มความสัมพันธ์ระหว่างคู่ตัวแปรได้ร้อยละ 91.66 ส่งผลให้ได้คะแนนความคล้ายคลึงโดยรวมเฉลี่ย ร้อยละ 87.28 ดังแผนภูมิที่ 5, 6, 7 และภาพที่ 1



แผนภูมิที่ 5 ผลการเปรียบเทียบค่าเฉลี่ยข้อมูลผู้ป่วยระหว่างข้อมูลจริงและเสมือน



แผนภูมิที่ 6 ผลการเปรียบเทียบส่วนเบี่ยงเบนมาตรฐานของผู้ป่วยในข้อมูลจริงและเสมือน



แผนภูมิที่ 7 ผลการเปรียบเทียบสัมประสิทธิ์ความแปรผันของชุดข้อมูลจริงและเสมือน

```
(1/2) Evaluating Column Shapes: 16/16 [00:00:00, 665.44it/s]
Column Shapes Score: 79.81%
(2/2) Evaluating Column Pair Trends: 120/120 [00:00:00, 275.26it/s]
Column Pair Trends Score: 83.56%
Overall Score (Average): 81.69%
```

ภาพที่ 1 ผลการประเมินความคล้ายคลึงของชุดข้อมูลเสมือนด้วยตัวชี้วัดสากล (SDMetrics)

3.5 ผลการคัดเลือกคุณลักษณะและความสำคัญของฟีเจอร์

จากผลการทดลองวิเคราะห์ที่ฟีเจอร์พบว่าแบบจำลองให้ความสำคัญสูงสุดกับลักษณะเม็ดเลือดแดงขนาดเล็ก ซึ่งส่งผลต่อการทำนายมากที่สุดถึงร้อยละ 43.79 ควบคู่ไปกับค่าความผันผวนของขนาดเม็ดเลือด (sd และ rdw_cv) และลักษณะรูปร่างเซลล์ที่ผิดปกติจำเพาะ ดังตารางที่ 2

ตารางที่ 2 ผลการคัดเลือกคุณลักษณะและความสำคัญของตัวแปร

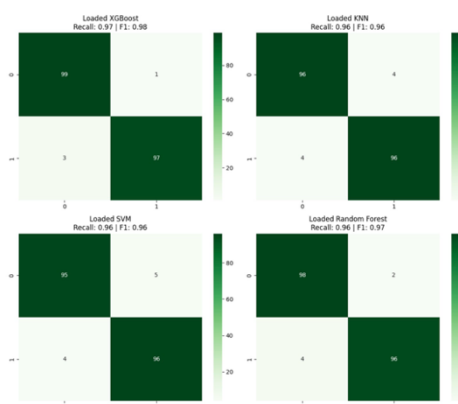
คุณลักษณะ	การคัดเลือก	ค่าความสำคัญ	การจัดอันดับ	เปอร์เซ็นต์การส่งผล
Microcytic	True	3.285957	1	43.794542
SD	True	0.844484	1	11.255099
rdw_cv	True	0.534306	1	7.121114
Hypochromic1-2	True	0.509472	1	6.790132
Stomatocyte	True	0.373122	1	4.972885

3.6 ผลการพัฒนาโมเดล AI สำหรับคัดกรองความเสี่ยงของโรคธาลัสซีเมีย

จากการเปรียบเทียบประสิทธิภาพของแบบจำลองในการจำแนกข้อมูล พบว่าแบบจำลอง XGBoost ให้ประสิทธิภาพสูงที่สุด โดยมีค่าความแม่นยำร้อยละ 98.00 และค่า F1-Score อยู่ที่ร้อยละ 97.98 รองลงมาคือแบบจำลอง Random Forest ในขณะที่ KNN และ SVM ให้ผลลัพธ์ใกล้เคียงกันตามลำดับ ดังตารางที่ 3 และ ภาพที่ 2

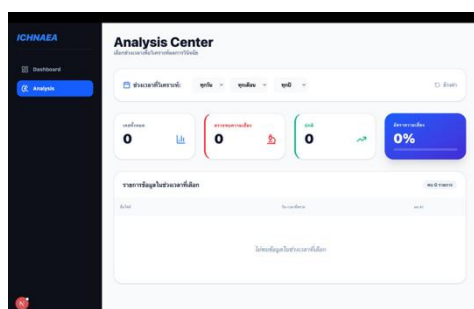
ตารางที่ 3 การเปรียบเทียบประสิทธิภาพของโมเดล
 จำแนกชนิดเม็ดเลือดระหว่างการเทรนด้วย XGBoost,
 KNN, SVM, และ Random Forest

Model	Accuracy	Precision	Recall	F1 score
XGBoost	0.9800	0.9898	0.9700	0.9798
Random forest	0.9700	0.9796	0.9600	0.9697
KNN	0.9600	0.9600	0.9600	0.9600
SVM	0.9550	0.9505	0.9600	0.9552

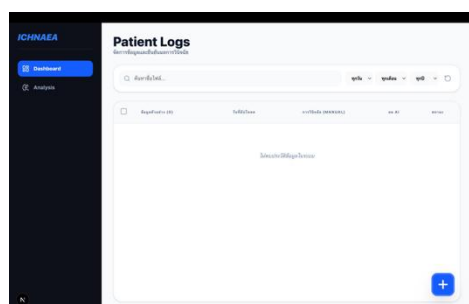


ภาพที่ 2 ตาราง confusion matrix ของผลการทดลอง

3.7 ข้อมูลภาพตัวอย่างในเว็บไซต์



ภาพที่ 3 ตัวอย่างหน้าเว็บไซต์



ภาพที่ 4 ตัวอย่างหน้าเว็บไซต์ (2)

4. สรุปและอภิปรายผลการทดลอง

ระบบนี้ประสบความสำเร็จในการคัดกรอง
 โรคธาลัสซีเมีย โดยใช้ Color Normalization ปรับสี
 สไลด์เลือด และใช้ YOLO11n-cls คัดกรองคุณภาพ
 ภาพร่วมกับ Canny Edge ได้แม่นยำ ร้อยละ 99.18
 โมเดลสามารถจำแนกความผิดปกติของเม็ดเลือดแดง 9
 ชนิด ได้แม่นยำสูงถึง ร้อยละ 97.57 ด้วยเวลาเพียง
 4.42 มิลลิวินาทีต่อภาพมีการใช้ Otsu's Thresholding
 วิเคราะห์สัญญาณวิทยาเชิงปริมาณเพื่อสร้าง Tabular
 Data และแก้ปัญหาความไม่สมดุลของข้อมูลด้วย
 CTGAN (2,000 Epochs) จำนวน 1,000 แถวได้เสมือน
 จริงเมื่อนำข้อมูลมาผ่าน StandardScaler และ RFECV
 5-Fold พบว่าโมเดล XGBoost ให้ประสิทธิภาพดีที่สุด
 ในการทำนายแยกกลุ่มโรคด้วยความแม่นยำ ร้อยละ
 98.00 เมื่อรวมระบบ พบว่ามีความแม่นยำในการคัด
 กรองโรคภาพรวมถึง ร้อยละ 96.00 พร้อมเป็น
 เครื่องมือสนับสนุนงานวินิจฉัยในพื้นที่ห่างไกลได้อย่างมี
 ประสิทธิภาพ

กิตติกรรมประกาศ

ขอขอบพระคุณ อาจารย์ที่ปรึกษา โรงเรียน
 วิทยาศาสตร์จุฬาภรณราชวิทยาลัย พิษณุโลก และ
 โรงพยาบาลพุทธชินราช พิษณุโลก ที่ให้การสนับสนุน
 การทำงานวิจัยครั้งนี้

เอกสารอ้างอิง

- Dyah Tyas. (2565). Red Blood Cell Dataset in
 Thalassemia Case. (ออนไลน์). สืบค้นเมื่อ
 20 มิถุนายน 2568, จาก [https://www.
 researchgate.net/publication/3626780
 36_Red_Blood_Cell_Dataset_in_Thala
 ssemia_case](https://www.researchgate.net/publication/362678036_Red_Blood_Cell_Dataset_in_Thalassemia_case).
- Korranat. (2563). Red blood cell segmentation
 and classification from microscopic
 images using machine learning.
 วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต
 (วิทยาการคอมพิวเตอร์). กรุงเทพฯ:
 จุฬาลงกรณ์มหาวิทยาลัย. ถ่ายเอกสาร.