เทคนิคทาง Machine Learning เพื่อการพัฒนางานการตรวจปัสสาวะสำหรับเพาะเชื้อ

สมพร ทาบัว*

บทคัดย่อ

การตรวจวิเคราะห์ทางปัสสาวะ (UA) เป็นการตรวจที่ทำได้ง่ายและราคาไม่แพง การพบเม็ดเลือดขาวจะเป็นข้อบ่งชี้ว่ามีการอักเสบติดเชื้อแบคทีเรีย การนำเทคนิคทาง Machine learning มาใช้ควบคุมคุณภาพปัสสาวะก่อนส่งตรวจเพาะเชื้อ สามารถลดภาระงานและลดค่าใช้จ่ายได้

วัตถุประสงค์: เพื่อพัฒนาและเปรียบเทียบประสิทธิภาพแบบจำลองที่เหมาะสมที่สุดสำหรับการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะโดยใช้เทคนิคทาง Machine learning

วิธีการศึกษา: ศึกษาข้อมูลย้อนหลังการส่งตรวจปัสสาวะสำหรับการเพาะเชื้อ ระหว่างปี 2565 และใช้ข้อมูลผลตรวจปัสสาวะ(UA) และผลเพาะเชื้อทางเดินปัสสาวะ (U/C) ทางห้องปฏิบัติการ จำนวน 345 ราย การทำแบบจำลองในโปรแกรม Orange: เทคนิคการสร้างแบบจำลองต่าง ๆ จะถูกเลือกเปรียบเทียบเป็นค่าที่เหมาะสมที่สุดและมีการประเมินโมเดลด้วยสถิติ ความไว ความจำเพาะ ค่าทำนายผลบวก ค่าทำนายผลลบและค่าความถูกต้อง

ผลการศึกษา: โมเดลต้นไม้ตัดสินใจ (Decision tree) เป็นแบบจำลองที่มีประสิทธิภาพและเหมาะสมที่สุดสำหรับการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะ ซึ่งแสดงจำนวน WBC $\geq$ 3-5 cells/HPF ขึ้นไป มีความสอดคล้องกับการเกิดการติดเชื้อในทางเดินปัสสาวะ เมื่อประเมินแบบจำลองด้วยสถิติ พบว่าการใช้เกณฑ์การตรวจพบ WBC $\geq$ 3-5 cells/HPF ขึ้นไป มาคัดกรองการตรวจปัสสาวะสำหรับเพาะเชื้อ พบว่ามีความไวเท่ากับร้อยละ 96.0 ความจำเพาะร้อยละ 80.0 ค่าทำนายผลบวกเท่ากับร้อยละ 81.5 ค่าทำนายผลลบเท่ากับร้อยละ 96.0 และค่าความถูกต้องเท่ากับร้อยละ 88.0

สรุป: โมเดลต้นไม้ตัดสินใจ (Decision tree) เป็นแบบจำลองที่มีประสิทธิภาพและเหมาะสมที่สุดสำหรับการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะ การใช้ผลการตรวจปัสสาวะที่พบเม็ดเลือดขาวตั้งแต่ 3-5 cells/HPF ขึ้นไป สามารถนำมาเป็นแนวทางการตรวจคัดกรองการตรวจปัสสาวะสำหรับเพาะเชื้อได้

คำสำคัญ: การตรวจวิเคราะห์ทางปัสสาวะ, การเพาะเชื้อทางเดินปัสสาวะ, Machine learning

ส่งเรื่องตีพิมพ์: 21 ตุลาคม 2567
อนุมัติตีพิมพ์: 27 ธันวาคม 2567

*วท.บ. (เทคนิคการแพทย์) กลุ่มงานเทคนิคการแพทย์และพยาธิวิทยาคลินิก โรงพยาบาลภูเขียวเฉลิมพระเกียรติ จังหวัดชัยภูมิ

Machine Learning techniques for work development: Testing urine analysis for urine culture

Somporn Thabua*

ABSTRACT

Background: Urine analysis (UA) is a simple and cost-effective test. The presence of WBC indicates bacterial infection and inflammation. Applying machine learning techniques to pre-screen urine samples before sending them for culture can reduce workload and costs.

***Objective*:** To develop and compare the performance of the most suitable machine learning model for predicting urinary tract infections (UTIs) using urine analysis data.

***Methods*:** A retrospective study was conducted using urine culture data from 2022. Data from 345 urine samples, including UA results and culture results, were analyzed. Various machine learning models were developed and evaluated using Orange software. Model performance was assessed using metrics such as sensitivity, specificity, positive predictive value, negative predictive value, and accuracy.

***Results*:** The decision tree model demonstrated the highest performance in predicting UTIs. A WBC count of $\geq$ 3-5 cells/HPF was significantly associated with UTIs. When evaluated using the decision tree model, a WBC count cutoff of $\geq$ 3-5 cells/HPF yielded a sensitivity of 96.0%, specificity of 80.0%, positive predictive value of 81.5%, negative predictive value of 96.0%, and an overall accuracy of 88.0%.

***Conclusion*:** The decision tree model is a highly effective tool for predicting UTIs. Using a WBC count of $\geq$ 3-5 cells/HPF as a cutoff for urine culture can serve as a guideline for screening urine samples.

Keywords: Urinalysis, Urine Culture, Machine Learning

Submission: 21 October 2024

Publication: 27 December 2024

* B.Sc. (Medical technology) Department of Medical Technology and Clinical Pathology, Phukhieo Chalermprakiat hospital, Chaiyaphum province

ความสำคัญของปัญหา

ในปัจจุบัน กระทรวงสาธารณสุขได้มีนโยบายส่งเสริมการตรวจวิเคราะห์ทางห้องปฏิบัติการอย่างสมเหตุผล (Rational Lab Use: RLU) เพื่อมุ่งหวังลดจำนวนรายการตรวจที่ไม่จำเป็น ลดภาระงานและค่าใช้จ่ายในระบบการรักษา โดยเฉพาะในผู้ป่วยติดเชื้อระบบทางเดินปัสสาวะ (Urinary Tract Infections; UTI) ที่มักมีอาการแตกต่างกันตามประเภทของการติดเชื้อ เช่น อาการปัสสาวะขัด ปัสสาวะบ่อย และอาการไข้เมื่อเกิดกรวยไตอักเสบ[1, 2]

การตรวจปัสสาวะ (Urinalysis) ถือเป็นการยืนยันการวินิจฉัย UTI โดยเฉพาะในผู้ป่วยที่มีอาการ อย่างไรก็ตาม ผลการตรวจปัสสาวะ เช่น การตรวจพบเม็ดเลือดขาว หรือแบคทีเรียในปัสสาวะนั้น อาจไม่สามารถยืนยันการติดเชื้อได้เสมอไป นอกจากนี้ การเพาะเชื้อและการทดสอบความไวต่อยาปฏิชีวนะอาจไม่จำเป็นในกรณีของ acute uncomplicated cystitis แต่จำเป็นในกรณีที่มี pyelonephritis หรือ UTI ที่ซับซ้อน เพื่อให้การรักษามีประสิทธิภาพสูงสุด[3, 4]

การใช้การตรวจทางห้องปฏิบัติการแบบเบื้องต้น รวมถึงการเพาะเชื้อและการทดสอบความไวต่อสารต้านจุลชีพ จะช่วยเพิ่มความไวและความจำเพาะในการวินิจฉัย ทำให้แพทย์สามารถวางแผนการรักษาได้อย่างเหมาะสม อย่างไรก็ตาม โรงพยาบาลภูเขียวเฉลิมพระเกียรติมีผู้ป่วยเข้ารับบริการในด้านการวินิจฉัย UTI เป็นจำนวนมาก ส่งผลให้มีการใช้ทรัพยากรเกินความจำเป็น และเพิ่มภาระงานในห้องปฏิบัติการ[1, 3, 5]

ในการวิจัยนี้ ผู้วิจัยจึงมีจุดมุ่งหมายเพื่อศึกษาและประยุกต์ใช้เทคนิค Machine Learning ในการประมวลผลและวิเคราะห์ข้อมูล เพื่อพัฒนาระบบคัดกรองเบื้องต้นสำหรับผู้ป่วย UTI โดยเฉพาะในกรณีที่มีการตรวจปัสสาวะพื้นฐานที่สามารถช่วยในการระบุความเสี่ยงและแนวโน้มการเกิด UTI ได้อย่างมีประสิทธิภาพและแม่นยำมากขึ้น ซึ่งจะนำไปสู่การลดค่าใช้จ่ายและภาระงานที่เกินความจำเป็นในระบบการรักษา ส่งเสริมการใช้ทรัพยากรทางการแพทย์อย่างมีประสิทธิผล และยกระดับคุณภาพการรักษาผู้ป่วยในอนาคต การนำเทคโนโลยี Machine Learning มาใช้ในงานวิจัยนี้ ไม่เพียงแต่เป็นการประยุกต์ใช้เทคโนโลยีที่ทันสมัย แต่ยังสามารถสร้างประโยชน์ในการวิเคราะห์ข้อมูลจำนวนมาก เพื่อมุ่งสู่การการตรวจคัดกรองที่มีคุณภาพและแม่นยำยิ่งขึ้นในการดูแลรักษาผู้ป่วย UTI ในสถานพยาบาลต่าง ๆ ได้อย่างมีประสิทธิภาพ

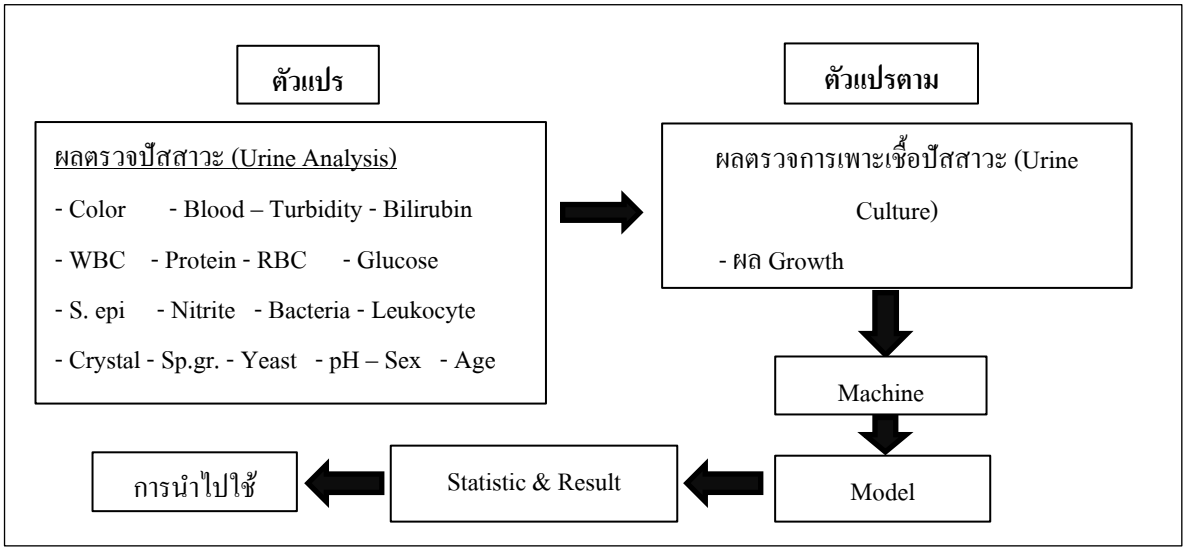
วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาแบบจำลองการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะ (Urinary Tract Infection) ด้วยผลตรวจทางห้องปฏิบัติการโดยใช้เทคนิคทาง Machine Learning
2. เพื่อเปรียบเทียบประสิทธิภาพแบบจำลองที่เหมาะสมที่สุดสำหรับการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะ

กรอบแนวคิดในการวิจัย

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบประสิทธิภาพแบบจำลองที่เหมาะสมในการวินิจฉัยโรคทางเดินปัสสาวะ โดยเทคนิค Machine Learning ตัวแปรที่นำมาวิเคราะห์ ได้แก่

อายุ เพศ ประวัติการติดเชื้อทางเดินปัสสาวะ โรคประจำตัว และผลการตรวจทางห้องปฏิบัติการอื่น โดยใช้เทคนิค Decision Tree ในการสร้างโมเดล



รูปภาพที่ 1 กรอบแนวคิด (Concept framework)

ระเบียบวิธีวิจัย

รูปแบบการวิจัย

เป็นการศึกษาจากข้อมูลย้อนหลัง (Retrospective study) และพัฒนางานประจำสู่การวิจัย (Research to research: R2R) ผ่านการพิจารณาจริยธรรมการวิจัย จากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ของสำนักงานสาธารณสุขจังหวัดชัยภูมิ หมายเลข 37/2567 ผ่านวันที่ 4 มิถุนายน 2567

ประชากรและกลุ่มตัวอย่าง

ประชากรที่ศึกษา (Study Population) คือ ผู้ป่วยที่มีผลการส่งตรวจวิเคราะห์ทางปัสสาวะและส่งเพาะเชื้อทางเดินปัสสาวะของโรงพยาบาลภูเขียวเฉลิมพระเกียรติ และกลุ่มตัวอย่าง เก็บข้อมูลในช่วงวันที่ 1 มกราคม – 31 ธันวาคม 2565 จำนวน 345 ราย

การพัฒนาโมเดลแบบจำลอง (Modeling)

ประกอบด้วยขั้นตอน ดังต่อไปนี้

1. กำหนดปัญหา (Problem Definition) เข้าใจและกำหนดเป้าหมายของปัญหาในการศึกษา ก่อนเริ่มทำแบบจำลอง

2. การเตรียมข้อมูล (Data Preparation) ข้อมูลที่ใช้ต้องมีคุณภาพเพียงพอและเหมาะสมสำหรับการฝึกแบบจำลอง ได้แก่ การเก็บรวบรวมข้อมูล (Data Collection): รวมข้อมูลจากฐานข้อมูล การทำความสะอาดข้อมูล (Data Cleaning): จัดการข้อมูลผิดปกติและข้อมูลที่ไม่สมบูรณ์ การแปลงข้อมูล (Data Transformation): การแปลงข้อมูลให้เป็นตัวเลขและจัดมาตรฐานข้อมูล การแบ่งชุดข้อมูล (Data Splitting): แบ่งข้อมูลออกเป็นชุด Train, Validation และ Test

3. การเลือกแบบจำลอง (Model Selection) เลือกประเภทของอัลกอริทึมตามลักษณะของปัญหาและข้อมูล เช่น ได้แก่ (1) Supervised Learning (การเรียนรู้แบบมีผู้สอน) ได้แก่ Linear Regression, Logistic Regression, Decision Trees, Random Forest และ Neural Networks (2) Unsupervised Learning (การเรียนรู้แบบไม่มีผู้สอน) และ (3) Reinforcement Learning (การเรียนรู้เสริมกำลัง) ซึ่งการศึกษานี้ทำการศึกษาแบบ Supervised Learning

4. การฝึกแบบจำลอง (Training the Model) ได้แก่ การกำหนดฟังก์ชันเป้าหมาย เช่น การลดค่าความผิดพลาด การใช้ข้อมูล Training Set เพื่อสอนให้แบบจำลองเรียนรู้รูปแบบของข้อมูล

5. การประเมินแบบจำลอง (Model Evaluation) เป็นการตรวจสอบประสิทธิภาพของแบบจำลองโดยใช้ข้อมูล Validation Set หรือ Test Set เช่น Accuracy, Precision, Recall, F1-score, ROC-AUC เป็นต้น

6. การปรับปรุงแบบจำลอง (Model Optimization) ได้แก่ ปรับพารามิเตอร์ของอัลกอริทึม เช่น Learning Rate, Depth ของ Decision Tree, จำนวน Neurons, Feature Engineering: เพิ่ม, ลบ, หรือแปลงฟีเจอร์เพื่อปรับปรุงประสิทธิภาพ และการเลือกแบบจำลองใหม่: ลองใช้โมเดลอื่น ๆ หากผลลัพธ์ยังไม่ดี

7. การนำไปใช้งาน (Model Deployment) เมื่อได้แบบจำลองที่เหมาะสมแล้ว สามารถนำไปใช้งานจริง ติดตามผลและปรับปรุงแบบจำลองเมื่อข้อมูลเปลี่ยนไป (Model Monitoring)

ขั้นตอนการทำวิจัย

1. เก็บข้อมูลจากระบบ HIS (Hosxp) และ MLAB

2. การเตรียมข้อมูล (Data Preparation): มีผลตัวแปรต้น คือผลการตรวจวิเคราะห์ปัสสาวะ ได้แก่ Color, Blood, Turbidity, Bilirubin, WBC, Protein, RBC, Glucose, S. epi, Nitrite, Bacteria, Leukocyte, Crystal, Sp.gr., Yeast, pH, Sex และ Age เป็นต้น ตัวแปรตาม คือ ผลการเพาะเชื้อ ได้แก่ ผล Growth และผล No growth เตรียมชุดข้อมูลในรูปแบบ Excel จำนวน 345 ราย

3. ชุดข้อมูลในการสอนระบบทดสอบแบ่งทดสอบเป็น ชุดเรียนรู้: ชุดทดสอบ เท่ากับ 70: 30

4. การทำแบบจำลอง (Modeling) ในโปรแกรม Orange: เทคนิคการสร้างแบบจำลองต่าง ๆ จะถูกเลือกและนำไปใช้และพารามิเตอร์เปรียบเทียบเป็นค่าที่เหมาะสมที่สุด

5. การประเมินผล (Evaluation): ทำการตรวจสอบประสิทธิภาพของแบบจำลองโดยใช้ข้อมูล Test Set ต้องประเมินโมเดลอย่างละเอียดและทบทวนขั้นตอนที่ดำเนินการ มีเกณฑ์การตัดสินใจผลการทำเหมืองข้อมูล

6. การนำไปใช้ (Deployment): ใช้พารามิเตอร์หาค่าความสัมพันธ์ของผลการตรวจ UA ที่สัมพันธ์กับผลการเพาะเชื้อปัสสาวะ U/C

ข้อมูลและการเก็บรวบรวม

เก็บข้อมูลโดยโปรแกรม Orange - 3.35.0-Miniconda-x86_64.exe (64 bit) ประกอบด้วย ข้อมูลผลตรวจปัสสาวะ (Urine analysis) Attribute ของการทดสอบตามตัวแปรจากผลตรวจทางปัสสาวะ และผลเพาะเชื้อ

ทางเดินปัสสาวะ (Urine culture) ทางห้องปฏิบัติการผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะจากระบบฐานข้อมูลจากโรงพยาบาลภูเขียวเฉลิมพระเกียรติ ดังตารางที่ 1

ตารางที่ 1 ตัวแปรการตรวจปัสสาวะและช่วงที่กำหนด

No.	ตัวแปร	ช่วงที่กำหนด	No.	ตัวแปร	ช่วงที่กำหนด
1.	Age	0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, >=70	11.	Protein	Negative, Positive
2.	Sex	Male, Female	12.	Glucose	Negative, Positive
3.	Color Urine	Yellow, Pale yellow, Deep yellow, Red, Brown	13.	Nitrite	Negative, Positive
4.	Turbidity	Turbid, Slightly turbid, Clear	14.	Leukocyte	Negative, Positive
5.	WBC	0, 0-1, 1-2, 3-5, 5-10, 20-30, 30-50, 50-100, >100 (<3-5, >=3-5)	15.	Blood	Negative, Positive
6.	RBC	0, 0-1, 1-2, 3-5, 5-10, 20-30, 30-50, 50-100, >100	16.	pH	5.0 – 9.0
7.	Squamous epithelial cell	0, 0-1, 1-2, 3-5, 5-10, 20-30	17.	Specific gravity	1.005 – 1.030
8.	Yeast	Found, Not found	18.	Crystal	Found, Not found
9.	Bacteria	Few, Moderate, Numerous			
10.	Bilirubin	Negative, Positive			

สถิติวิเคราะห์

สถิติที่ใช้ ได้แก่ หาความไว ความจำเพาะ ค่าทำนายผลบวก ค่าทำนายผลลบและค่าความถูกต้อง และแจกแจงเป็น จำนวน ร้อยละ

ความไว (Sensitivity) = $\frac{TP}{TP+FN}$

ความจำเพา (Specificity) = $\frac{TN}{TN+FP}$

ค่าทำนายผลบวก (PPV) = $\frac{TP}{TP+FP}$

ค่าทำนายผลลบ (NPV) = $\frac{TN}{TN+FN}$

ค่าความถูกต้อง (Accuracy) = $\frac{TP+TN}{TP+TN+FP+FN}$

TP คือ ค่าผลบวกจริง (true positive), FP คือ ค่าผลบวกปลอม (false positive), TN คือ ค่าผลลบจริง (true negative), FN คือ ค่าผลลบปลอม (false negative)

ผลการศึกษา

จากการพัฒนาโมเดล และการทดสอบ พบว่า แบบจำลองที่เหมาะสมที่สุดสำหรับการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะ คือ โมเดลต้นไม้ตัดสินใจ (Decision tree) ซึ่งเป็นโมเดลที่ใช้ในการจัดกลุ่มข้อมูลหรือทำนายค่าในรูปแบบของ tree diagram แบ่งข้อมูลเป็นส่วนย่อยเรื่อย ๆ ประกอบด้วย Decision Nodes และ leaf nodes Decision Nodes จะตรวจสอบ attribute และแสดงค่า attribute ที่ทดสอบ จนหา Final rule ดังรูปที่ 2 – 4

จากการศึกษาจากการใช้โปรแกรม Orange Version 3.35.0-Miniconda-x86 (64 bit) พบว่า Neural Network Decision tree และ Logistic Regression มีค่ามากที่สุดตามลำดับ เมื่อตรวจสอบค่าทางสถิติตามรูปพบว่าค่า AUC CA

F1 Precision Recall และ MCC ของ Neural Network ได้ค่าเท่า 1.00 ทั้งหมด Decision Tree ได้ค่าเท่ากับ 0.994, 0.926, 0.919, 0.927, 0.926 และ 0.882 ตามลำดับ ส่วน Logistic Regression ได้ค่าเท่ากับ 0.993, 0.914, 0.905, 0.918, 0.914 และ 0.862 ตามลำดับ ดังรูปที่ 2

Evaluation results for target (None, show average over classes)

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	1.000	1.000	1.000	1.000	1.000	1.000
Tree	0.994	0.926	0.919	0.927	0.926	0.882
Logistic Regression	0.993	0.914	0.905	0.918	0.914	0.862
SVM	0.883	0.877	0.842	0.893	0.877	0.803
kNN	0.957	0.802	0.778	0.773	0.802	0.678
Naive Bayes	0.954	0.728	0.765	0.883	0.728	0.651

รูปที่ 2 การประเมินโมเดลตามสถิติ

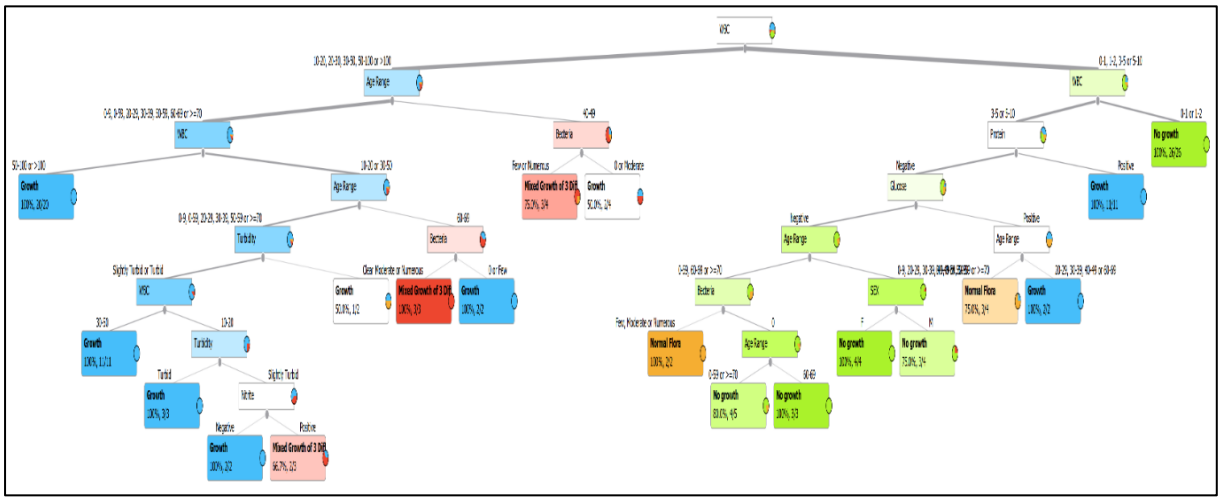
พบว่าออกอริทึมต่าง ๆ จากผลการตรวจปัสสาวะตัวที่มีความสัมพันธ์กับผลการตรวจเพาะเชื้อทางปัสสาวะมากที่สุด ได้แก่ WBC Bacteria Turbidity และ Leukocyte ตามลำดับ โดย WBC ได้ค่าความสัมพันธ์มากที่สุด ดังนี้ ค่า Info. Gain เท่ากับ 0.904 ค่า Gain ratio เท่ากับ 0.324 ค่า Gini เท่ากับ 0.331 ค่า X^2 เท่ากับ 256.067 และค่า ReliefF เท่ากับ 0.591 ตามลำดับ ดังรูปที่ 3

		#	Info. gain	Gain ratio	Gini	X^2	ReliefF
1	WBC	9	0.904	0.324	0.331	256.067	0.591
2	Becteria	4	0.785	0.412	0.278	120.872	0.441
3	Turbidity	3	0.497	0.320	0.165	70.156	0.370
4	Leukocyte	2	0.341	0.342	0.135	54.644	0.098
5	Protein	2	0.321	0.321	0.125	52.634	0.106
6	Blood	2	0.315	0.315	0.122	51.977	0.098
7	Nitrite	2	0.269	0.270	0.099	46.981	0.227
8	Glucose	2	0.218	0.230	0.103	45.845	0.295
9	Age Range	8	0.151	0.066	NA	3.359	0.208
10	Yeast	2	0.035	0.412	0.012	0.271	0.021
11	SEX	2	0.018	0.019	0.004	4.298	0.104

รูปที่ 3 แสดงความสัมพันธ์ของอัลกอริทึมที่มีความสัมพันธ์กัน

Decision tree ซึ่งเป็นโมเดลที่จัดกลุ่มข้อมูลหรือทำนายค่าในรูปแบบของ Decision tree diagram ซึ่งโมเดลแสดงจำนวน WBC มีผลต่อการตรวจปัสสาวะสำหรับเพาะเชื้อ โดยพบว่าการตรวจ WBC $\geq$ 3-5 cells/HPF มีความสอดคล้องกับการเกิดภาวะการณ์ติดเชื้อในทางเดินปัสสาวะ

อย่างมีนัยสำคัญดังรูปที่ 4



รูปที่ 4 แสดงโมเดล Decision Tree

พบว่าการตรวจคัดกรองปัสสาวะ (UA) โดยใช้ WBC ≥ 3-5 cells/HPF ขึ้นไป มีความสัมพันธ์กับผลการเพาะเชื้อทางเดินปัสสาวะที่ให้ผลเพาะเชื้อบวก (Growth) โดยพบว่าจากตารางที่ 2 สามารถหาความไวได้เท่ากับ ร้อยละ 96.0 ความจำเพาะเท่ากับร้อยละ 80.0 ค่าทำนายผลบวกเท่ากับ ร้อยละ 81.5 ค่าทำนายผลลบ ร้อยละ 96.0 และค่าความถูกต้องอยู่ที่ ร้อยละ 88.0

ตารางที่ 2 แสดงการประเมินโมเดลโดยใช้สถิติ

Method	**Urine culture (gold standard)**	
Urine analysis	Growth	No Growth
WBC ≥ 3 -5	159	36
WBC ≤ 3 - 5	6	144

ผลการศึกษาดังกล่าว นำไปใช้เป็นเกณฑ์คัดกรองการตรวจปัสสาวะสำหรับการเพาะเชื้อ โดยต้องผ่านการตรวจ Urine analysis ทุกราย และพบ WBC ≥ 3- 5/HPF ขึ้นไปจึงจะได้รับการพิจารณาต่อ การดำเนินงานดังกล่าว สามารถลดปริมาณการส่งตรวจปัสสาวะสำหรับเพาะเชื้อ ปี 2566 ลงได้ร้อยละ 65.8 และลดค่าใช้จ่าย ลงได้ 62,280 บาท

ตารางที่ 3 ข้อมูลค่าตรวจปัสสาวะสำหรับเพาะเชื้อ ปีงบประมาณ 2566

รายการ	ปริมาณงาน (ราย)	ร้อยละ	ค่าตรวจปัสสาวะ (บาท)	ค่าตรวจเพาะเชื้อปัสสาวะ (บาท)	ค่าใช้จ่ายรวม (บาท)
ปัสสาวะที่ผ่านเกณฑ์	180	34.2	5,400	27,000	32,400
ปัสสาวะที่ไม่ผ่านเกณฑ์	346	65.8	10,380	51,900	62,280
รวม	526	100	15,780	78,900	94,680

อภิปรายผล

ผลการศึกษาพบว่า โมเดลต้นไม้ตัดสินใจ (Decision Tree) ให้ประสิทธิภาพในการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะดีที่สุด เมื่อพิจารณาจากค่า AUC, F1-score, Precision, Recall และ MCC ที่สูงใกล้เคียง 1.00 ซึ่งบ่งชี้ถึงความสามารถในการจำแนกผู้ป่วยที่เป็นโรคและไม่เป็นโรคได้อย่างแม่นยำ ผลลัพธ์นี้สอดคล้องกับงานวิจัยของ Alqahtani และคณะ (2023) ที่พบว่า Random Forest, Naive Bayes และ K-NN ให้ผลแม่นยำสูงในการทำนายผลของโรค[8]

เมื่อเปรียบเทียบกับโมเดลอื่นๆ เช่น Logistic Regression, SVM และ K-NN พบว่าโมเดลต้นไม้ตัดสินใจให้ผลลัพธ์ที่ดีกว่า ซึ่งสอดคล้องกับการศึกษาของวนิดา พงษ์สงวนและคณะ (2560) ที่ศึกษาการพัฒนาแบบจำลองปัจจัยที่มีผลต่อการเป็นโรคเบาหวาน และชลธิศา พลทองมาก และพุธษดี ศิริแสงตระกูล (2560)[9] ที่ศึกษาการวิเคราะห์ความเสี่ยงการเป็นโรคไวรัสตับอักเสบซี โดยทั้งสองการศึกษาพบว่าโมเดลต้นไม้ตัดสินใจให้ผลลัพธ์ที่ดีในการจำแนกประเภท และสอดคล้องกับการศึกษาของชลธิศา พลทองมาก และพุธษดี ศิริแสงตระกูล[10] ที่ศึกษาการวิเคราะห์ความเสี่ยงการเป็นโรคไวรัสตับอักเสบซี โดยต้นไม้การตัดสินใจ (Decision Tree) และทฤษฎีเบย์เซี่ยน (Bayesian Theorem) พบว่าวิธีการต้นไม้ตัดสินใจ (Decision tree) ให้ความแม่นยำเท่ากับ 92.9% แต่ทฤษฎีเบย์เซี่ยน (Bayesian Theorem) ให้ความแม่นยำ (Accuracy) มากที่สุดคือ 93.9% ซึ่งการศึกษาดังกล่าวใช้โปรแกรม Wega ต่างจากการศึกษานี้ที่ใช้โปรแกรม Orange

นอกจากนี้ การศึกษานี้ ยังพบว่าการตรวจปัสสาวะโดยใช้เกณฑ์พบ WBC $\geq$ 3-5 cell/HPF มีความสัมพันธ์กับการเกิดการติดเชื้อในทางเดินปัสสาวะ ซึ่งสอดคล้องกับการศึกษาของอาทิตยา ปุ๊กมะดัน (2558)[11] และพุทรา ชลสวัสดิ์และคณะ (2563)[12] ที่ศึกษาการใช้ Urine dipstick ในการตรวจปัสสาวะ อย่างไรก็ตาม การศึกษาของ Little P และคณะ (2009)[13] แนะนำให้ใช้ผลตรวจปัสสาวะที่พบเม็ดเลือดขาวมากกว่า >10 WBC/HPF ซึ่งแตกต่างจากการศึกษานี้

สรุปและข้อเสนอแนะ

โมเดลต้นไม้ตัดสินใจ (Decision tree) เป็นแบบจำลองที่มีประสิทธิภาพและเหมาะสมที่สุดสำหรับการพยากรณ์ผู้ป่วยโรคติดเชื้อทางเดินปัสสาวะ โดยใช้การตรวจปัสาวะที่พบเม็ดเลือดขาวตั้งแต่ 3-5 cells/HPF ขึ้นไป สามารถนำมาเป็นแนวทางการตรวจคัดกรองการตรวจการ

เพาะเชื้อทางปัสสาวะได้อย่างมีประสิทธิภาพ และสามารถลดภาระงาน รวมทั้งสามารถลดค่าใช้จ่ายที่เกินความจำเป็นได้ หากมีการใช้ Leukocyte esterase (LE) และ Nitrite ในเกณฑ์ร่วมด้วยจะทำให้มีความแม่นยำมากยิ่งขึ้น

เพื่อเป็นการพัฒนาต่อยอดในอนาคต การพิจารณาใช้เทคนิค Feature Selection เพื่อเลือกเฉพาะตัวแปรที่มีความสำคัญในการทำนายผลลัพธ์ ซึ่งจะช่วยลดความซับซ้อนของโมเดลและเพิ่มประสิทธิภาพ รวมถึงการวิเคราะห์ความสำคัญของแต่ละตัวแปรที่ใช้ในการสร้างโมเดล เพื่อทำความเข้าใจกลไกการเกิดโรคและพัฒนาวิธีการป้องกันและรักษาโรคได้ดียิ่งขึ้น

กิตติกรรมประกาศ

ขอขอบพระคุณนายแพทย์สุภาพ สำราญวงษ์ ผู้อำนวยการโรงพยาบาลภูเขียวเฉลิมพระเกียรติ และทนพญ.กชกร ทองสุขแก้ง หัวหน้ากลุ่มงานเทคนิคการแพทย์และพยาธิวิทยาคลินิก โรงพยาบาลภูเขียวเฉลิมพระเกียรติ ที่สนับสนุนและอนุญาตให้ทำการศึกษาในครั้งนี้

เอกสารอ้างอิง

1. Chang R, Greene MT, Chenoweth CE, Kuhn L, Shuman E, Rogers MA, et al. Epidemiology of hospital-acquired urinary tract-related bloodstream infection at a university hospital. Infect Control Hosp Epidemiol 2011;32(11):1127-9.
2. Wilson ML, Gaido L. Laboratory diagnosis of urinary tract infections in adult patients. Clin Infect Dis 2004;38(8):1150-8.
3. Schmiemann G, Kniehl E, Gebhardt K, Matejczyk MM, Hummers-Pradier E. The diagnosis of urinary tract infection: A systematic review. Dtsch Arztebl Int 2010;107(21):361-7.
4. Mercaldo ND, Lau KF, Zhou XH. Confidence intervals for predictive values with an emphasis on case-control studies. Stat Med 2007;26(10):2170-83.
5. ปิยะ ศิริลักษณ์, นิศารัตน์ โอภาสเกียรติกุล, สุรศักดิ์ หมื่นพล, [บรรณาธิการ]. แนวทางการตรวจทางห้องปฏิบัติการอย่างสมเหตุผล (Rational Laboratory Use, RLU). นนทบุรี : สำนักมาตรฐานห้องปฏิบัติการ กรมวิทยาศาสตร์การแพทย์ กระทรวงสาธารณสุข, 2566.
6. Hay AD, Birnie K, Busby J, Delaney B, Downing H, Dudley J, et al. The diagnosis of urinary tract infection in young children (DUTY): A diagnostic prospective observational study to derive and validate a clinical algorithm for the diagnosis of urinary tract infection in children presenting to primary care with an acute illness. Health Technol Assess 2016;20(51):291-294.
7. ศุภสิทธิ์ พรรณารุโณทัย, ทรงยศ พิลาสันต์, วิไลลักษณ์ แสงศรี, ปฤษฐพร กิ่งแก้ว, [บรรณาธิการ]. คู่มือการประเมินเทคโนโลยีด้านสุขภาพในประเทศไทย ฉบับปรับปรุง พ.ศ. 2564. นนทบุรี : โครงการประเมินเทคโนโลยีและนโยบายด้านสุขภาพ, 2564.

8. Alqahtani A, Alsubai S, Binbusayyis A, Sha M, Gumaei A, Zhang YD. Prediction of urinary tract infection in IoT-fog environment for smart toilets using modified attention-based ANN and machine learning algorithms. Appl Sci 2023;13:5859-60.
9. วนิดา พงษ์สงวน, ทิพยา ถินสูงเนิน, มาโนช ถินสูงเนิน. การพัฒนาแบบจำลองปัจจัยที่มีผลต่อการเป็นโรคเบาหวานด้วยเทคนิคต้นไม้ตัดสินใจ. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม 2561;1(1):1-8.
10. ชลธิศา พลทองมาก, พุธษดี ศิริแสงตระกูล. การวิเคราะห์ความเสี่ยงการเป็นโรคไวรัสตับอักเสบซี โดยต้นไม้การตัดสินใจและทฤษฎีเบย์เซี่ยน. (ออนไลน์). 2553. (เข้าถึงเมื่อวันที่ 17 มกราคม 2566: http://202.28.94.51/SAR/sar52/files/25540001_Paper%20NCCIT2011-168-%E0%B8%8A%E0%B8%A5%E0%B8%98%E0%B8%B4%E0%B8%A9%E0%B8%B2.pdf)
11. อาทิตยา ปุ๊กมะดัน. การประเมินประสิทธิภาพการใช้แถบน้ำยาสำเร็จรูปตรวจสารเคมีในปัสสาวะเปรียบเทียบกับการเพาะเชื้อในปัสสาวะเพื่อวิจัยการติดเชื้อในระบบทางเดินปัสสาวะ ณ โรงพยาบาลสิรินธร. (ออนไลน์). 2558. (เข้าถึงเมื่อวันที่ 17 มกราคม, 2566 :http://www.msdbangkok.go.th/dowload%20file/Personal/Succeed/290663/2.pdf)
12. พุทรา ชลสวัสดิ์, เบญจมาภรณ์ วงษ์พันธุ์, จันทวรรณ สัตยารักษ์, อัสรินดา อับดุลกานาน. การเปรียบเทียบผลการตรวจวิเคราะห์ปัสสาวะด้วยแถบทดสอบกับการเพาะเชื้อเพื่อวินิจฉัยการติดเชื้อทางเดินปัสสาวะ. วารสารเทคนิคการแพทย์ 2563;48(3):751-61.
13. Little P, Turner S, Rumsby K, Warner G, Moore M, Lowes JA, et al. Dipsticks and diagnostic algorithms in urinary tract infection: Development and validation, randomised trial, economic analysis, observational cohort and qualitative study. Health Technol Assess 2009;13(19): iii-iv, ix-xi, 1-73.