Evaluating machine learning models for triage in the emergency room: A case study at Bang Khla Hospital

Phanikarn Phuttharaksa

Bang Khla Hospital, Chachoengsao

Abstract

Emergency room triage is a critical process that significantly impacts patient outcomes and resource allocation. The application of Machine Learning (ML) in this domain has garnered substantial interest due to its potential to enhance decision-making accuracy and efficiency. This study aims to develop and evaluate an ML model for classifying patient urgency levels in an emergency room setting. We conducted a retrospective analysis using data from patients admitted to Bang Khla Hospital, Chachoengsao, Thailand, between October 1, 2021, and August 30, 2024. From an initial dataset of 61,602 records, 29,389 were retained after data cleaning and preprocessing. These records were categorized into four urgency levels according to the Emergency Severity Index (ESI): resuscitate, emergent, urgent, and semi-urgent. The Random Forest model developed in this study achieved an overall accuracy of 98.67%, with an F1-Score of 1.00 for both resuscitate and emergent levels. However, the model faced challenges in classifying the urgent level, achieving an F1-Score of 0.86. These findings demonstrate the potential of ML in improving triage accuracy in emergency rooms, which could lead to more efficient patient prioritization and resource management. Nonetheless, further research is necessary to enhance the model's performance in classifying complex cases and to evaluate its implementation in real-world clinical settings. This study contributes to the growing body of evidence supporting the integration of ML technologies in emergency healthcare services.

Keywords: *RMachine learning; Triage classification; Emergency room, Random Forest; Retrospective study; Clinical decision support systems; Emergency severity index*

Received: 10 January 2025, Revised: 25 April 2025, Accepted: 1 May 2025

Correspondence: Phanikarn Phuttharaksa, Bang Khla Hospital, 62 Moo 1 Pak Nam, Bang Khla District, Chachoengsao 24110, Tel: 038 541 009, E-mail: sachikorainy@gmail.com

การประเมินโมเดลการเรียนรู้ของเครื่องสำหรับการจำแนกระดับความเร่งด่วนในห้องฉุกเฉิน: กรณีศึกษาโรงพยาบาลบางคล้า

ภณิกาญจน์ พุทธรักษา

โรงพยาบาลบางคล้า จังหวัดฉะเชิงเทรา

บทคัดย่อ

การจำแนกระดับความเร่งด่วน (Triage) ในห้องฉุกเฉินเป็นกระบวนการสำคัญที่ส่งผลต่อประสิทธิภาพการรักษาและการจัดการทรัพยากร การประยุกต์ใช้ Machine Learning (ML) ในกระบวนการนี้ จึงได้รับความสนใจอย่างมาก การศึกษานี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินประสิทธิภาพของโมเดล ML ในการจำแนกระดับความเร่งด่วนของผู้ป่วยในห้องฉุกเฉิน โดยใช้ข้อมูลย้อนหลังจากผู้ป่วยที่เข้ารับบริการ ณ โรงพยาบาลบางคล้า จังหวัดฉะเชิงเทรา ระหว่างวันที่ 1 ตุลาคม 2564 ถึง 30 สิงหาคม 2567 จากข้อมูลทั้งหมด 61,602 รายการ หลังจากการคัดกรองและทำความสะอาดข้อมูล เหลือข้อมูลที่นำมาใช้ในการพัฒนาโมเดลจำนวน 29,389 รายการ แบ่งออกเป็น 4 ระดับตามระบบ Emergency Severity Index (ESI) ได้แก่ resuscitate, emergent, urgent, และ semi-urgent โมเดล Random Forests ที่พัฒนาขึ้นสามารถทำนายระดับความเร่งด่วนได้อย่างแม่นยำถึง 98.67% โดยมีค่า F1-Score สูงถึง 1.00 สำหรับกลุ่ม resuscitate และ emergent อย่างไรก็ตาม โมเดลยังมีความท้าทายในการจำแนกผู้ป่วยระดับ urgent ซึ่งมีค่า F1-Score อยู่ที่ 0.86 ผลการศึกษานี้แสดงให้เห็นถึงศักยภาพของ ML ในการปรับปรุงความแม่นยำของการจำแนกระดับความเร่งด่วนในห้องฉุกเฉิน ซึ่งอาจนำไปสู่การพัฒนาระบบสนับสนุนการตัดสินใจที่มีประสิทธิภาพมากขึ้นในอนาคต อย่างไรก็ตาม การวิจัยเพิ่มเติมยังจำเป็นเพื่อปรับปรุงความแม่นยำในการจำแนกผู้ป่วยที่มีอาการซับซ้อน และเพื่อประเมินการนำโมเดลไปใช้ในสถานการณ์จริง

คำสำคัญ: *การเรียนรู้ของเครื่อง; การจำแนกระดับความเร่งด่วน; ห้องฉุกเฉิน; Random Forest; การศึกษาย้อนหลัง; ระบบสนับสนุนการตัดสินใจทางคลินิก*

วันที่รับต้นฉบับ: 10 มกราคม 2568, วันที่แก้ไข: 25 เมษายน 2568, วันที่ตอบรับ: 1 พฤษภาคม 2568

บทนำ

การประเมินระดับวิกฤตผู้ป่วย (Triage) เป็นสิ่งสำคัญในการจัดลำดับความเร่งด่วนของผู้ป่วยที่มาเข้ารับการรักษาในห้องฉุกเฉิน โดยมีเป้าหมายเพื่อคัดกรองผู้ป่วยฉุกเฉินให้ได้รับการประเมิน รักษา อย่างรวดเร็วและถูกต้อง เหมาะสม ให้ทรัพยากรทางการแพทย์ถูกนำไปใช้กับผู้ป่วยที่มีความจำเป็นที่สุดก่อน การประเมินนี้ช่วยลดความเสี่ยงในการเกิดเหตุการณ์ร้ายแรงที่อาจเกิดขึ้นหากผู้ป่วยที่อยู่ในภาวะวิกฤตไม่ได้รับการรักษาอย่างทันท่วงทีตรงกับความจำเป็นทางการแพทย์ อันจะนำไปสู่การลดโอกาสอัตราการเสียชีวิตหรือความพิการ ทั้งนี้ยังช่วยเพิ่มความพึงพอใจของผู้ป่วยที่เข้ารับบริการ เนื่องจากได้รับการดูแลตามระดับความเจ็บป่วย [1]

Emergency Severity Index (ESI) เป็นเครื่องมือที่ถูกนำมาใช้ในการจำแนกระดับความเร่งด่วนของผู้ป่วยในห้องฉุกเฉิน ซึ่งแบ่งออกเป็น 5 ระดับตามความรุนแรงและความเร่งด่วนของอาการ ESI ได้รับการพัฒนาเพื่อช่วยเจ้าหน้าที่ทางการแพทย์ในการตัดสินใจอย่างรวดเร็วว่า ผู้ป่วยคนใดจำเป็นต้องได้รับการรักษาเป็นอันดับแรก และทรัพยากรทางการแพทย์ใดที่ควรถูกจัดสรรให้เหมาะสมกับระดับความเร่งด่วน การใช้ ESI ช่วยให้การจัดลำดับผู้ป่วยในห้องฉุกเฉินมีความเป็นระบบและเป็นธรรมมากยิ่งขึ้น [2] 3 คำถามหลักที่ทำให้รอการรักษาไม่ได้หรือไม่ คือ ภาวะเสี่ยง ซีม ปวด [3]

ห้องฉุกเฉินของโรงพยาบาลบางคล้า จังหวัดฉะเชิงเทรา ให้บริการผู้ป่วยอุบัติเหตุและฉุกเฉินตลอด 24 ชั่วโมง โดยเฉพาะอย่างยิ่งในกรณีอุบัติเหตุและผู้ป่วยฉุกเฉินที่ต้องการการดูแลรักษาอย่างเร่งด่วน อย่างไรก็ตาม พบว่ามีปัญหาการเกิด under triage ซึ่งเป็นการประเมินระดับความเร่งด่วนของผู้ป่วยต่ำกว่าความเป็นจริง โดยเฉพาะในผู้ป่วยกลุ่ม emergency ที่มีอาการทรุดลงขณะรอตรวจรักษา นอกจากนี้ ยังพบกรณีที่ผู้ป่วยที่มีสัญญาณชีพผิดปกติตั้งแต่แรกรับถูกส่งกลับจากแผนกผู้ป่วยนอก (OPD) กลับมารับการรักษาที่ห้องฉุกเฉิน ซึ่งชี้ให้เห็นถึงปัญหาการประเมินความรุนแรงของผู้ป่วยที่ไม่แม่นยำ

ผู้นิพนธ์ประสานงาน: ภณิกาญจน์ พุทธรักษา, โรงพยาบาลบางคล้า, 62 หมู่ 1 ปากน้ำ อำเภอบางคล้า ฉะเชิงเทรา 24110, โทร. 038 541 009, E-mail: sachikorainy@gmail.com

การนำ Machine Learning (ML) มาใช้ในการทำ Triage เป็นหนึ่งในแนวทางที่ได้รับความสนใจเพิ่มขึ้นในปัจจุบัน ML มีศักยภาพในการช่วยปรับปรุงความแม่นยำในการจำแนกระดับความรุนแรงของผู้ป่วย โดยอาศัยการวิเคราะห์ข้อมูลจำนวนมากและการเรียนรู้จากข้อมูลเหล่านั้น ทำให้สามารถประเมินอาการของผู้ป่วยได้อย่างแม่นยำและรวดเร็วมากขึ้น ส่งผลให้การตัดสินใจในการจัดลำดับความสำคัญเป็นไปอย่างมีประสิทธิภาพ อีกทั้งยังช่วยลดความเสี่ยงในการเกิด under triage ซึ่งอาจนำไปสู่การจัดการทรัพยากรทางการแพทย์ที่ไม่เหมาะสมและเพิ่มความเสี่ยงต่อผู้ป่วยที่มีอาการรุนแรง [4] ในยุคที่เทคโนโลยีมีความก้าวหน้าเช่นปัจจุบัน ผู้จัดทำจึงสนใจศึกษาว่า สามารถนำ Machine Learning (ML) มาใช้เพื่อลดปัญหา under triage ได้อย่างไร โดยการพัฒนาและประเมินผลโมเดล Machine Learning สำหรับการทำ Triage ในห้องฉุกเฉิน

การศึกษาจำนวนมากแสดงให้เห็นว่าโมเดล Machine Learning สามารถเพิ่มความแม่นยำในการประเมินผลลัพธ์ทางคลินิกในห้องฉุกเฉินได้ดีกว่าวิธีการแบบดั้งเดิม [5], [6] การศึกษานี้จึงมีวัตถุประสงค์เพื่อพัฒนาและประเมินผลโมเดลการเรียนรู้ของเครื่อง (ML) ที่สามารถนำมาใช้ในการจำแนกระดับความเร่งด่วนของผู้ป่วยในห้องฉุกเฉินได้อย่างมีประสิทธิภาพ โดยการใช้ข้อมูลย้อนหลังจากผู้ป่วยที่เข้าใช้บริการในโรงพยาบาลบางคล้า จังหวัดฉะเชิงเทรา การวิจัยนี้จะสำรวจว่าการใช้ ML สามารถปรับปรุงการทำ Triage ในห้องฉุกเฉินได้อย่างไร และสามารถนำไปใช้ในสภาพแวดล้อมจริงเพื่อเพิ่มประสิทธิภาพในการให้บริการแก่ผู้ป่วยได้หรือไม่

ทบทวนวรรณกรรม

ในปัจจุบัน การประยุกต์ใช้เทคโนโลยีปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง (Machine Learning; ML) ในห้องฉุกเฉิน (Emergency Department; ED) ได้รับความสนใจเพิ่มขึ้นอย่างมากเพื่อเพิ่มประสิทธิภาพในการคัดกรองและการพยากรณ์ผลลัพธ์ของผู้ป่วย โดยมีการพัฒนาโมเดลต่างๆ ที่ช่วยให้สามารถจัดลำดับความสำคัญของผู้ป่วยได้อย่างแม่นยำและลดความคลาดเคลื่อนในการตัดสินใจของเจ้าหน้าที่ [4],[7],[8]

โมเดล ML เช่น Random Forest, Logistic Regression, และ Gradient Boosting ถูกนำมาใช้บ่อยในงานวิจัยหลายชิ้น เนื่องจากมีประสิทธิภาพสูงในการทำนายผลลัพธ์ เช่น การรับเข้ารักษาในโรงพยาบาลและการคาดการณ์ผลลัพธ์ในผู้ป่วยฉุกเฉิน [4], [9]. Raita และคณะ [4] ใช้ข้อมูลจาก National Hospital and Ambulatory Medical Care Survey เพื่อพัฒนาโมเดล ML โดยมีการทดสอบ 4 โมเดลได้แก่ Lasso regression, random forest, gradient boosted decision tree, และ deep neural network ซึ่งพบว่า ML model ทั้งหมดมีค่า AUC สูงกว่าค่า AUC ของโมเดลดั้งเดิม ESI ในการทำนายผลลัพธ์ของผู้ป่วย เช่น ICU admission และการรับเข้ารักษาในโรงพยาบาล

การเลือก features สำหรับการพัฒนาโมเดล ML นั้นมักประกอบด้วยข้อมูลพื้นฐานของผู้ป่วย (เช่น อายุ เพศ และอาการสำคัญ) รวมถึงค่าชีวเคมีในเลือดที่สามารถบ่งบอกถึงความรุนแรงของอาการ [9] Feretzakis และคณะ [10] ได้ทดสอบ features ต่างๆ เช่น ค่าชีวเคมี (เช่น Urea, Creatinine, Lactate Dehydrogenase) และค่าการแข็งตัวของเลือด (เช่น D-Dimer, INR) โดยพบว่าโมเดล ML ที่ใช้ข้อมูลเหล่านี้มีความแม่นยำสูงในการทำนายการรับเข้ารักษาในโรงพยาบาล ซึ่งช่วยสนับสนุนการตัดสินใจทางคลินิกได้อย่างมีประสิทธิภาพ

การประยุกต์ใช้ ML เพื่อสนับสนุนระบบคัดกรองผู้ป่วยในห้องฉุกเฉินยังช่วยแก้ปัญหาที่เกิดจากการพึ่งพาการตัดสินใจของบุคลากรโดยตรง งานวิจัยล่าสุดแสดงให้เห็นว่า ML สามารถลด bias ในการประเมินและเพิ่มความสม่ำเสมอในการคัดกรอง [7], [11] ตัวอย่างเช่น, Tyler และคณะ [11] พบว่า ML-based triage สามารถพยากรณ์ผลลัพธ์ของผู้ป่วยได้แม่นยำและรวดเร็วกว่า ESI ซึ่งช่วยให้ทีมแพทย์สามารถจัดการทรัพยากรได้ดียิ่งขึ้น และลดการใช้ทรัพยากรที่ไม่จำเป็นในผู้ป่วยที่ไม่วิกฤต

จากงานวิจัยเหล่านี้จะเห็นได้ว่าการใช้ ML สามารถยกระดับคุณภาพของการคัดกรองผู้ป่วยในห้องฉุกเฉินได้อย่างมีนัยสำคัญ อย่างไรก็ตาม การพัฒนาโมเดลเหล่านี้ยังคงมีความท้าทาย เช่น ความต้องการข้อมูลจำนวนมากและการจัดการปัญหาด้าน ethical ที่เกิดขึ้นจากการใช้ข้อมูลทางการแพทย์ในปริมาณมาก [8], [10]

วิธีการศึกษา

Study design, settings, and data source

การศึกษาเชิงเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่องเพื่อหาวิธีการที่มีประสิทธิภาพดีที่สุดในการประเมินผลลัพธ์การคัดกรองผู้ป่วยในห้องฉุกเฉินของโรงพยาบาลบางคล้า จังหวัดฉะเชิงเทรา ข้อมูลถูกเก็บรวบรวมย้อนหลังจากฐานข้อมูลโรงพยาบาล (Electronic Medical Records: EMRs) ระหว่างวันที่ 1 ตุลาคม 2564 ถึง 30 สิงหาคม 2567 จากข้อมูลทั้งหมด 61,602 รายการ ข้อมูลส่วนบุคคลที่สามารถระบุตัวตน ไม่ได้เป็นส่วนหนึ่งของชุด

ข้อมูล โดยมุ่งเน้นที่ข้อมูลระดับความเร่งด่วนตามมาตรฐาน Emergency Severity Index (ESI) ซึ่งแบ่งเป็น 4 ระดับหลัก และจำกัดเฉพาะผู้ป่วยที่มีอายุ 18 ปีขึ้นไป (18 years) เพื่อให้ได้ข้อมูลที่สอดคล้องกับการประเมินในกลุ่มผู้ใหญ่เท่านั้น การประมวลผลและการวิเคราะห์ข้อมูลได้ดำเนินการโดยใช้ภาษาโปรแกรม Python และไลบรารีทางวิทยาศาสตร์ที่เกี่ยวข้อง ได้แก่ Pandas, NumPy และ scikit-learn

Data analysis approach

การศึกษานี้ดำเนินการตามทดลองดังที่แสดงใน ภาพที่ 1 เพื่อสร้างแบบจำลองการคัดกรองผู้ป่วย (triage model) เราได้ประเมินแบบจำลองการเรียนรู้ของเครื่อง (ML) จำนวน 4 แบบ ประกอบด้วย Logistic Regression, Decision Tree, Random Forest, และ Gradient Boosting ทั้งนี้ การประเมินประสิทธิภาพของแต่ละโมเดลถูกดำเนินการโดยใช้เทคนิค Cross-Validation เพื่อให้ได้ผลลัพธ์ที่แม่นยำและเชื่อถือได้

Data preparation

การเตรียมข้อมูลเป็นสิ่งจำเป็นในกระบวนการเรียนรู้ของ ML รวมถึงการเข้ารหัสป้ายกำกับ (label encoding) ข้อมูลบางส่วนถูกตัดออกเพื่อลดความซับซ้อนและเพิ่มความแม่นยำในการจำแนก โดยข้อมูลที่ถูกตัดออก ได้แก่ ผู้ป่วยในระดับ 5 (non-urgent) เนื่องจากผู้ป่วยกลุ่มนี้เป็นผู้ป่วยตามนัดหมายและไม่ถือว่าเป็นกรณีฉุกเฉิน นอกจากนี้ ผู้ป่วยที่เสียชีวิตก่อนถึงโรงพยาบาลก็ถูกคัดออกจากการศึกษาเนื่องจากข้อมูลเหล่านี้ไม่สามารถใช้ในการประเมินระดับความเร่งด่วนของผู้ป่วยในห้องฉุกเฉินได้ หลังจากเก็บรวมรวมข้อมูล ได้ทำการตรวจสอบข้อมูลค่าผิดปกติ (Outlier Detection) ดังภาพที่ 2 โดยใช้เกณฑ์ในการตัดค่าผิดปกติออกจากข้อมูล เนื่องจากค่าที่ผิดปกติอาจเกิดจากการป้อนข้อมูลที่ไม่ถูกต้องหรือการวัดค่าที่ไม่สมเหตุสมผล ซึ่งอาจทำให้โมเดล Machine Learning เรียนรู้ผิดพลาดและทำให้ผลการจำแนกมีความแม่นยำน้อยลง ในขั้นตอนนี้ ค่าที่สูงกว่าหรืออยู่นอกขอบเขตที่กำหนดไว้จะถูกลบออก ได้แก่ ค่าความดันโลหิตตัวบน (Systolic Blood Pressure) > 300 mmHg, ค่าความดันโลหิตตัวล่าง (Diastolic Blood Pressure) > 200 mmHg, อัตราการเต้นของหัวใจ (Pulse Rate) > 300 ครั้งต่อนาที, อัตราการหายใจ (Respiratory Rate) > 80 ครั้งต่อนาที, และค่าออกซิเจนในเลือด (Oxygen Saturation) > 100% อาจเกิดขึ้นได้จากหลายสาเหตุ เช่น ข้อผิดพลาดในการวัด, ความผิดพลาดในการป้อนข้อมูล หลังจากการทำความสะอาดข้อมูล ทำให้เหลือข้อมูลที่พร้อมใช้ในการพัฒนาโมเดลจำนวน 29,389 รายการ จากข้อมูลทั้งหมดที่รวบรวมได้ 61,602 รายการ แบ่งตามระดับ 1-4 ได้ชุดข้อมูล ดังภาพที่ 3 จัดการความไม่สมดุลของข้อมูล (Class Imbalance) ในชุดข้อมูลเทคนิค Oversampling ด้วย SMOTE (Synthetic Minority Oversampling Technique)

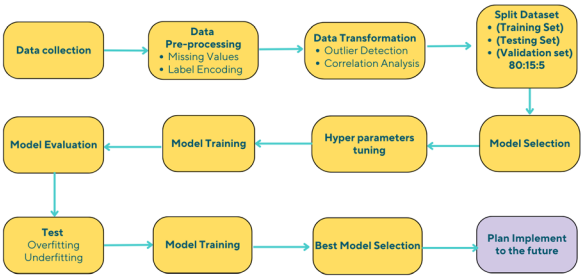
Features

ตัวแปรทางคลินิกที่ใช้ในการทดลองนี้ ประกอบด้วยข้อมูลลักษณะผู้ป่วย (อายุและเพศ), สัญญาณชีพ (อัตราการหายใจ, อุณหภูมิร่างกาย, ค่าอิ่มตัวของออกซิเจน, ความดันโลหิต และอัตราชีพจร) และระดับการประเมินภาวะฉุกเฉิน ESI (1-4) ได้ทำการตรวจสอบและปรับปรุงข้อมูลระดับ ESI ที่มาจากการประเมินโดยพยาบาลในครั้งแรก ซึ่งอาจมีข้อผิดพลาดทั้งการประเมินที่ต่ำกว่า (under-triage) หรือสูงกว่า (over-triage) ระดับที่ควรจะเป็น ก่อนนำข้อมูลดังกล่าวไปใช้ train โมเดล เช่น ผู้ป่วยที่มีสัญญาณชีพผิดปกติแต่ถูกประเมินต่ำกว่าความเป็นจริงจะได้รับการปรับให้สูงขึ้น ในทางกลับกัน กรณี over-triage เช่น ผู้ป่วยที่มีอาการไม่รุนแรงถูกจัดอยู่ในระดับเร่งด่วน จะถูกปรับลงมา การปรับปรุงระดับการประเมินนี้ดำเนินการโดยพยาบาลเฉพาะทางเวชปฏิบัติฉุกเฉิน และได้มีการชี้แจงอธิบายเหตุผลแก่พยาบาลผู้ปฏิบัติงานเพื่อความเข้าใจที่ตรงกัน

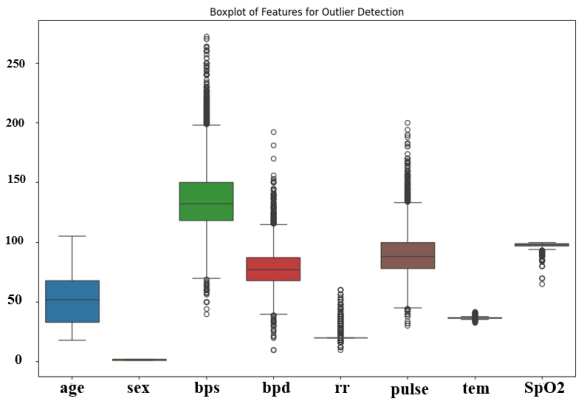
Training and Test Data

ข้อมูลถูกแบ่งออกเป็น 3 ชุดได้แก่ ชุดข้อมูลฝึกอบรม (Training Set) ชุดข้อมูลทดสอบ (Testing Set) และชุดตรวจสอบ (Validation Set) โดยใช้สัดส่วน 70:15:5 โมเดล Machine Learning ที่เลือกใช้ในการศึกษาครั้งนี้ ได้จากการทดสอบว่าโมเดลใดทำงานได้ดีที่สุดกับข้อมูลที่มีอยู่ โดยโมเดลที่ถูกทดสอบประกอบด้วย Logistic Regression, Decision Tree, Random Forest, และ Gradient Boosting อัลกอริธึมเหล่านี้ถูกเลือกเนื่องจากแสดงความสามารถในการพยากรณ์ที่สูงกว่าในปัญหาการคัดกรองผู้ป่วย (triage classification) ตามที่พบในงานวิจัยก่อนหน้า [12].

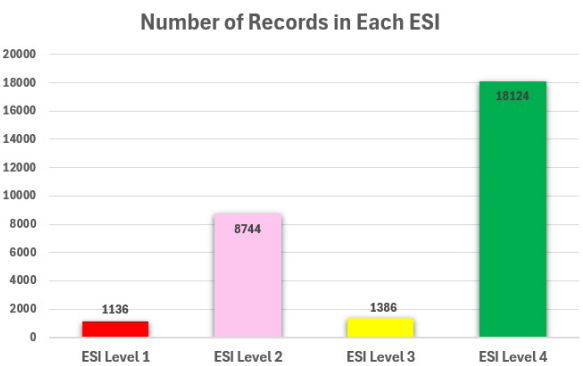
การประเมินประสิทธิภาพของแต่ละโมเดลถูกดำเนินการโดยใช้เทคนิค K-Fold Cross-Validation แบ่งข้อมูลทดสอบ 5 รอบ ให้แน่ใจว่าโมเดลไม่เกิดการ overfitting เพื่อให้ได้ผลลัพธ์ที่แม่นยำและเชื่อถือได้ จากการเปรียบเทียบพบว่าโมเดล Random Forest ให้ผลลัพธ์ที่ดีที่สุดในแง่ของความแม่นยำ (Accuracy) และ F1-Score ทำให้โมเดลนี้ถูกเลือกเป็นโมเดลหลักในการพัฒนา ตารางที่ 1 ปรับแต่งพารามิเตอร์เพิ่มเติม (Hyperparameter Tuning) เช่น จำนวนต้นไม้ในป่า(n_estimators) และความลึกสูงสุดของต้นไม้ (max_depth) เพื่อเพิ่มประสิทธิภาพของโมเดลในการจำแนกระดับความเร่งด่วน



ภาพที่ 1 ขั้นตอนที่ใช้ในการออกแบบและดำเนินการทดลอง



ภาพที่ 2 Boxplot of Features for Outlier Detection



ภาพที่ 3 Number of Records in Each ESI

ตารางที่ 1 ตารางการเปรียบเทียบผลทดสอบของโมเดล

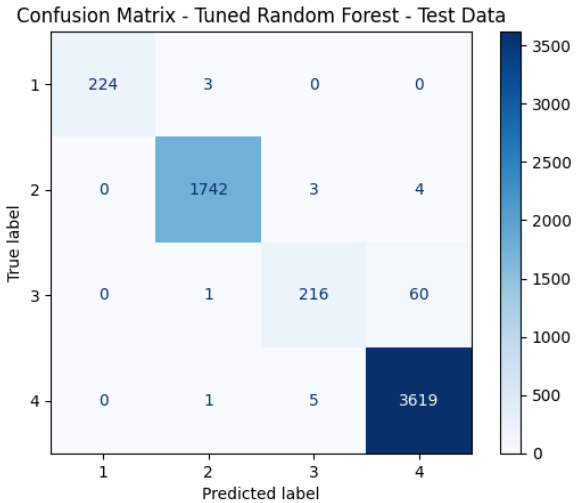
Class	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.88	0.24	0.06	0.10
Decision Tree	0.98	0.88	0.86	0.87
Random Forest	0.98	0.97	0.81	0.88
Gradient Boosting	0.98	0.95	0.79	0.86

ผลการศึกษาและการอภิปรายผล

1. ผลการศึกษา

1.1 ประสิทธิภาพของโมเดลที่ได้รับการปรับแต่ง (Tuned Model Performance)

โมเดล Random Forest ที่ได้รับการปรับแต่งแสดงความแม่นยำ (Accuracy) สูงถึง 98.69% บนชุดข้อมูลทดสอบ ซึ่งแสดงให้เห็นว่าโมเดลสามารถจำแนกระดับความเร่งด่วนของผู้ป่วยได้อย่างแม่นยำในกรณีส่วนใหญ่ ผลการทดสอบโมเดลแสดงให้เห็นว่า โมเดลมีค่า Precision, Recall, และ F1-Score สูงในระดับความเร่งด่วนที่สำคัญ โดยเฉพาะในระดับ 1 และระดับ 2 แต่ประสิทธิภาพในระดับ 3 Recall ยังค่อนข้างต่ำ (0.78) ซึ่งหมายความว่ามีผู้ป่วยในระดับ 3 บางส่วนที่ถูกจัดประเภทผิดไปเป็นระดับอื่น ทำให้ F1-Score ในระดับ 3 อยู่ที่ 0.86 ดังภาพที่ 4



ภาพที่ 4 Confusion matrix แสดงผลจากประสิทธิภาพของโมเดลที่ได้รับการทดสอบผ่านชุดข้อมูลทดสอบที่ไม่เคยเห็น

1.2 การวิเคราะห์ Confusion Matrix

โมเดล Random Forest ทำงานได้ดีในระดับ 1 และ 2 แต่ยังมีข้อผิดพลาดเล็กน้อยในการจำแนกระดับ 3 (Urgent) ซึ่งมี Recall ต่ำกว่าระดับอื่น ๆ แม้ว่าโมเดลจะมี Precision สูง (0.96) สำหรับระดับ 3 แต่ Recall ยังค่อนข้างต่ำ (0.78) ซึ่งหมายความว่ามีผู้ป่วยในระดับ 3 บางส่วนที่ถูกจัดประเภทผิดไปเป็นระดับอื่น ทำให้ F1-Score ในระดับ 3 อยู่ที่ 0.86 ดังตารางที่ 2 false positive คือผู้ป่วยถูกจำแนกให้มีความเร่งด่วนสูงกว่าที่เป็นจริง ซึ่งอาจทำให้เกิดการจัดลำดับความสำคัญที่ไม่จำเป็นและส่งผลต่อการจัดสรรทรัพยากรที่ไม่ถูกต้อง เช่น

การจัดการผู้ป่วยที่มีอาการน้อยเกินความจำเป็น อาจส่งผลให้ผู้ป่วยที่วิกฤตจริงไม่ได้รับการดูแลทันเวลา และ false negative เกิดขึ้นเมื่อผู้ป่วยถูกจำแนกให้อยู่ในระดับความเร่งด่วนต่ำกว่าที่ควรจะเป็น ซึ่งอาจส่งผลให้ผู้ป่วยที่มีอาการวิกฤตไม่ได้รับการรักษาอย่างทันท่วงที เพิ่มความเสี่ยงต่อการเสื่อมอาการหรือเสียชีวิต Random Forest อาจลดผลลบลวงได้ดีในกลุ่มผู้ป่วยระดับ resuscitate และ emergent แต่ยังมีความท้าทายในการจำแนกผู้ป่วยระดับ urgent ซึ่งส่งผลต่อค่า F1-Score ในระดับนี้ (0.86)

ตารางที่ 2 Confusion Matrix Model Random Forest แสดงให้เห็นผลการจำแนกในแต่ละระดับความเร่งด่วน

Class	Precision	Recall	F1-Score	Support
1.Resuscitate	1.00	0.99	0.99	227
2.Emergent	1.00	1.00	1.00	1749
3.Urgent	0.96	0.78	0.86	277
4.Seme-Urgent	0.98	1.00	0.99	3625

1.3 การเปรียบเทียบโมเดล (Model Comparison)

ในระหว่างการพัฒนาโมเดล มีการทดสอบและเปรียบเทียบโมเดลหลายแบบ เช่น Logistic Regression, Decision Tree, Random Forest และ Gradient Boosting จากการเปรียบเทียบพบว่า Random Forest มีประสิทธิภาพดีที่สุด โดยให้ค่า Accuracy และ F1-Score สูงกว่ารูปแบบอื่น ๆ และสามารถจัดการกับข้อมูลที่มีความซับซ้อนได้ดีกว่า ดังตารางที่ 1

2. อภิปรายผล

ผลการศึกษานี้แสดงให้เห็นว่าโมเดล Random Forest ที่พัฒนาขึ้นมีความแม่นยำสูงในการจำแนกระดับความเร่งด่วนของผู้ป่วยในห้องฉุกเฉิน โดยเฉพาะในระดับ 1 (Resuscitate) และระดับ 2 (Emergent) ซึ่งสอดคล้องกับงานวิจัยอื่น ๆ ที่แสดงให้เห็นว่าโมเดล Random Forest มีประสิทธิภาพในการจัดการข้อมูลที่มีความซับซ้อนและไม่สมดุลได้ดี เช่นเดียวกับที่พบในงานวิจัยของ Goto et al. (2019) ที่แสดงว่า Random Forest มีความสามารถสูงในการทำนายผลลัพธ์ทางคลินิกในสถานการณ์ฉุกเฉิน [6].

การศึกษานี้ยังชี้ให้เห็นถึงข้อดีของการใช้ Random Forest เมื่อเทียบกับโมเดลอื่น ๆ เช่น Logistic Regression และ Decision Tree ซึ่งได้รับการสนับสนุนจากงานวิจัยโดย Wright et al. (2017) ที่ระบุว่า Random Forest มักมีประสิทธิภาพสูงกว่าในด้านความแม่นยำและการจัดการกับข้อมูลที่มีหลายมิติ [13]. อย่างไรก็ตาม การจำแนกระดับ 3 (Urgent) ที่พบว่า Recall ต่ำกว่าในระดับอื่น ๆ แสดงให้เห็นว่ามีความท้าทายในการแยกแยะผู้ป่วยที่ต้องการการดูแลเร่งด่วนจากกลุ่มที่มีความเร่งด่วนต่ำกว่า ซึ่งอาจสอดคล้องกับข้อค้นพบในงานวิจัยของ Handel et al. (2021) ที่ระบุว่าความไม่สมดุลของข้อมูลในระดับต่าง ๆ อาจส่งผลต่อประสิทธิภาพของโมเดล [13], [14] จากการวิเคราะห์อาจเนื่องมาจากจำนวนข้อมูลที่มีอยู่ในระดับนี้น้อยกว่าหรือมีความคล้ายคลึงกับระดับอื่น ๆ ทำให้โมเดลแยกแยะได้ยากขึ้น และจากการพิจารณาพบว่าคุณลักษณะที่จะแยกระดับ 3 และระดับ 4 ออกจากกันอย่างชัดเจน ขึ้นอยู่กับกิจกรรมทางการแพทย์ที่ต้องทำเพิ่มเติมกับผู้ป่วย 2 กิจกรรมขึ้นไป เช่น ตรวจทางห้องปฏิบัติการ ฉีดยาบรรเทาอาการ ปรึกษาแพทย์เฉพาะทาง

ความสำคัญของผลลัพธ์ที่ได้รับในงานวิจัยนี้อยู่ที่การพิสูจน์ว่าโมเดล Machine Learning สามารถถูกนำมาใช้เพื่อเพิ่มประสิทธิภาพในการทำ Triage ในห้องฉุกเฉิน ซึ่งเป็นการสนับสนุนให้มีการนำเทคโนโลยีขั้นสูงเข้ามาใช้ในการตัดสินใจทางการแพทย์ที่ต้องการความรวดเร็วและแม่นยำสูง [15]

ผลการศึกษาชี้ให้เห็นว่า โมเดล Machine Learning ที่พัฒนาขึ้นมีประสิทธิภาพสูงในการจำแนกระดับความเร่งด่วนของผู้ป่วยในห้องฉุกเฉิน โดยเฉพาะในระดับ 1 และ 2 ที่มีความแม่นยำสูงมาก นอกจากนี้ การทดสอบและเปรียบเทียบโมเดลหลายรูปแบบช่วยยืนยันว่า Random Forest เป็นตัวเลือกที่เหมาะสมที่สุดในการใช้งานจริง การตรวจสอบ Overfitting/Underfitting เป็นส่วนสำคัญของการพัฒนาโมเดล Machine Learning เพื่อให้แน่ใจว่าโมเดลที่พัฒนาไม่เรียนรู้ข้อมูลมากเกินไป (Overfitting) หรือเรียนน้อยเกินไป (Underfitting) การตรวจสอบนี้ถูกดำเนินการโดยการเปรียบเทียบผลการทำงานของโมเดลบนชุดข้อมูลฝึกอบรมและชุดข้อมูลทดสอบ ผลการทดสอบโมเดล Random Forest ที่ถูกปรับแต่ง (Tuned Random Forest) บนชุดข้อมูลทดสอบแสดงให้เห็นว่าโมเดลมีความแม่นยำ (Accuracy) สูงถึง 98.69% ซึ่งบ่งชี้ว่าโมเดลสามารถจำแนกระดับความเร่งด่วนของผู้ป่วยได้อย่างแม่นยำในกรณีส่วนใหญ่

ข้อจำกัดของงานวิจัย

ข้อมูลจากโรงพยาบาลเพียงแห่งเดียว การวิจัยนี้ใช้ข้อมูลจากโรงพยาบาลบางคล้า จังหวัดฉะเชิงเทราเพียงแห่งเดียว ดังนั้นผลลัพธ์อาจไม่สามารถนำไปใช้อ้างอิงในโรงพยาบาลอื่นหรือในบริบทที่มีทรัพยากรและปัจจัยด้านสิ่งแวดล้อมที่แตกต่างกันได้

คุณภาพและความครบถ้วนของข้อมูล ข้อมูลที่ใช้ในการพัฒนาโมเดลมาจากการบันทึกในระบบ EMRs ซึ่งอาจมี

ความคลาดเคลื่อนหรือไม่สมบูรณ์ ส่งผลให้โมเดลมีความเสี่ยงต่อความไม่แม่นยำในการพยากรณ์ การตรวจสอบและการทำความสะอาดข้อมูลจึงเป็นสิ่งสำคัญและอาจส่งผลกระทบต่อผลลัพธ์หากมีข้อมูลที่ไม่สอดคล้องกับข้อกำหนด

การเลือก features สำหรับโมเดล เนื่องจากการศึกษานี้ใช้เพียง features ที่มีอยู่ในฐานข้อมูลของโรงพยาบาลและเป็นข้อมูลที่บันทึกได้โดยทั่วไปเท่านั้น จึงอาจขาดคุณสมบัติที่สำคัญอื่น ๆ ซึ่งอาจช่วยเพิ่มความแม่นยำในการพยากรณ์ของโมเดล เช่น การประเมินระดับความรู้สึกตัว (Consciousness Level) และการวัดระดับความเจ็บปวด (Pain Score)

ข้อจำกัดในการนำไปใช้งานจริง แม้ว่าโมเดล ML จะมีความแม่นยำสูงในกรณีศึกษา แต่การนำไปใช้งานในสภาพแวดล้อมจริงอาจมีข้อจำกัดเพิ่มเติม เช่น ความแตกต่างของลักษณะผู้ป่วยและการตอบสนองของระบบที่หลากหลาย ซึ่งจำเป็นต้องมีการทดสอบเพิ่มเติมในบริบทอื่น ๆ เพื่อยืนยันความแม่นยำของโมเดล

สรุปผลการศึกษา

การศึกษานี้ได้แสดงให้เห็นถึงความสามารถที่ดีในการคัดกรอง (triage) ในแผนกฉุกเฉิน แบบจำลองการเรียนรู้ของเครื่องเป็นเทคโนโลยีใหม่ในการเพิ่มประสิทธิภาพการตัดสินใจของแพทย์และพยาบาลในการคัดกรองผู้ป่วย ซึ่งจะช่วยปรับปรุงกระบวนการดูแลผู้ป่วยฉุกเฉิน ลดความแออัดในห้องฉุกเฉิน ลดระยะเวลารอคอยและช่วยจัดการการใช้ทรัพยากรของโรงพยาบาลได้อย่างมีประสิทธิภาพมากขึ้น สามารถลดจำนวนผู้ป่วยที่ได้รับการคัดกรองต่ำเกินไป (under triage) จากวิธีการแบบเดิม นอกจากนี้ แบบจำลองยังช่วยลดการคัดกรองเกิน (over triage) ที่นำไปสู่การจัดสรรทรัพยากรให้กับผู้ป่วยที่มีอาการไม่รุนแรงเกินความจำเป็น ผลลัพธ์ของการศึกษานี้จะเป็นประโยชน์ต่อการวิจัยในอนาคต

ผลการศึกษา พบว่ามีจำนวนข้อมูลทั้งหมด 61,602 รายการ แต่เมื่อทำการคัดกรองและทำความสะอาดข้อมูล พบว่าเหลือเพียง 29,389 รายการที่สามารถนำมาใช้ได้ในโมเดล Machine Learning ซึ่งหมายถึงมีข้อมูลที่ต้องถูกตัดทิ้งไปมากกว่าครึ่งหนึ่งของข้อมูลทั้งหมด สาเหตุของการตัดข้อมูลออก เกิดจากข้อมูลที่ไม่ครบถ้วน ข้อมูลที่ผิดพลาด หรือข้อมูลที่ไม่สอดคล้องกับเกณฑ์ที่ตั้งไว้ และเพื่อให้การคัดกรองผู้ป่วยด้วย ESI Triage มีความแม่นยำมากขึ้น แนะนำให้มีการเก็บข้อมูลเพิ่มเติมที่สำคัญ เช่น ระดับความรู้สึกตัว (Consciousness Level), ระดับความเจ็บปวด (Pain Score), และอาการสำคัญที่ทำให้ผู้ป่วยมารับการรักษาในโรงพยาบาล (Chief Complaint) เพิ่มเข้าไปในชุดข้อมูลที่ทำการ Train model ข้อมูลเหล่านี้เป็นปัจจัยสำคัญที่สามารถช่วยให้โมเดล Machine Learning สามารถเรียนรู้ได้ดีคัดกรองผู้ป่วยได้แม่นยำขึ้น

สำหรับการพัฒนาในอนาคต ผู้วิจัยแนะนำแนวทางการพัฒนาการบันทึกข้อมูลดังนี้ 1.การพัฒนาแบบฟอร์มบันทึกข้อมูลที่เป็นมาตรฐาน 2.จัดฝึกอบรมเจ้าหน้าที่ที่มีหน้าที่บันทึกข้อมูล เพื่อให้มีความเข้าใจในการบันทึกข้อมูลถูกต้องและสมบูรณ์เป็นไปในแนวทางเดียวกัน 3.ปรับปรุงระบบในการบันทึกข้อมูล เช่นหากบันทึกข้อมูลค่าที่ผิดไปจากปกติ ให้แจ้งเตือนและบันทึกข้อมูลไม่ได้จนกว่าจะแก้ไข และนอกจากการปรับปรุงคุณภาพการลงข้อมูลแล้ว ควรมีการทดลองใช้งานโมเดลในสภาพแวดล้อมทางคลินิกจริง เพื่อทดสอบผลลัพธ์ในสถานการณ์ที่มีความซับซ้อนมากขึ้น

REFERENCES

[1] C. B. Soster, F. Anschau, N. H. Rodrigues, L. G. A. da Silva, and A. Klafke, "Advanced triage protocols in the emergency service: A systematic review and meta-analysis," Rev. Latino-Am. Enfermagem, vol. 30, no. e3511, 2022, doi: 10.1590/1518-8345.5479.3511.

[2] N. Gilboy, P. Tanabe, D. Travers, and A. M. Rosenau, "Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4," Implementation Handbook. Agency for Healthcare Research and Quality, 2012.

[3] Department of Health & Human Services (USA), Emergency Severity Index (ESI) A Triage Tool for Emergency Department Care, 4th ed. USA: AHRQ Publishing, 2011.

[4] Y. Raita, T. Goto, M. Faridi, D. F. M. Brown, C. A. Camargo, and K. Hasegawa, "Emergency department triage prediction of clinical outcomes using machine learning models," Critical Care, vol. 23, no. 1, p . 64, 2019.

[5] O. Salman, Z. Taha, M. Alsabah, Y. Soubhi, et al., "A review on utilizing machine learning technology for emergency department triage," Computer Methods and Programs in Biomedicine, vol. 203, pp. 106357, 2021.

[6] T. Goto, C. Camargo, M. Faridi, R. Freishtat, et al., "Machine Learning–Based Prediction of Clinical Outcomes for Emergency Department Triage," JAMA Network Open, vol. 2, no. 9, p. e187354, 2019.

[7] S. Tyler, M. Olis, N. Aust, L. Patel, L. Simon, C. Triantafyllidis, V. Patel, D. W. Lee, B. Ginsberg, H. Ahmad, and R. J. Jacobs, "Use of Artificial Intelligence in Triage in Hospital Emergency Departments: A Scoping Review," Cureus, vol. 16, no. 5, p. e59906, May 2024.

[8] N. Shafaf and H. Malek, "Applications of Machine Learning Approaches in Emergency Medicine: A Review Article," Archives of Academic Emergency Medicine, vol. 7, no. 1, p. e34, July 2019.

[9] F. Gao, B. Boukebous, M. Pozzar, E. Alaoui, B. Sano, and S. Bayat-Makoei, "Predictive Models for Emergency Department Triage using Machine Learning: A Review," Obstetrics and Gynecology Research, vol. 5, no. 2, pp. 107–121, June 2022.

[10] G. Feretzakis, G. Karlis, E. Loupelis, D. Kalles, R. Chatzikyriakou, N. Trakas, E. Karakou, A. Sakagianni, L. Tzelves, S. Petropoulou, A. Tika, I. Dalainas, and V. Kaldis, "Using Machine Learning Techniques to Predict Hospital Admission at the Emergency Department," The Journal of Critical Care Medicine, vol. 8, no. 2, pp. 107–116, 2022.

[11] J. Y. Yu, G. Y. Jeong, O. S. Jeong, D. K. Chang, and W. C. Cha, "Machine Learning and Initial Nursing Assessment-Based Triage System for Emergency Department," Healthc. Inform. Res., vol. 26, no. 1, pp. 13–19, Jan. 2020.

[12] Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, Dugas A, Linton B, Kirsch T, Kelen G. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. Ann Emerg Med 2018;71:565–574. e2. https://doi.org/ 10.1016/j.annemergmed.2017. 08.005.

[13] A. Wright, M. McCoy, M. Henkin, et al., "Use of Machine Learning for Predicting Hospital Admissions and ED Resource Use," Acad. Emerg. Med., vol. 24, no. 8, pp. 893-902, 2017.

[14] R. W. Handel, J. A. Ma, P. A. Hacke, K. B. Heller, and S. J. Kellermann, "Impact of a Machine Learning Model on Emergency Department Triage Accuracy: A Systematic Review," J. Emergency Med., vol. 56, no. 3, pp. 324-330, 2021.

[15] M. J. Aljubran, H. J. Aljubran, M. Aljubran, M. Alkhalifah, M. Alkhalifah, and T. Alabdullah, "The utilizing of machine learning algorithms to improve triage in emergency departments: a retrospective observational study," Saudi J. Emerg. Med., vol. 4, no. 2, pp. 112–119, 2023. [Online]. Available: https://doi.org/10.24911/SJEMed/72-1673885230&#8203