

บทคัดย่อ

วัตถุประสงค์: เพื่อพัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง (machine learning) เพื่อทำนายโรคหลอดเลือดสมองชนิดหลอดเลือดสมองแตก โดยใช้เพียงข้อมูลทางคลินิกพื้นฐานแรกรับ ณ ห้องฉุกเฉิน เพื่อเป็นเครื่องมือสนับสนุนการตัดสินใจในการบริหารทรัพยากรและการส่งต่อผู้ป่วย

วิธีการศึกษา: การศึกษาแบบย้อนหลัง (retrospective cohort study) รวบรวมข้อมูลผู้ป่วยที่ได้รับการวินิจฉัยโรคหลอดเลือดสมองเฉียบพลัน ณ ห้องฉุกเฉิน โรงพยาบาลสิรินธร จังหวัดขอนแก่น จำนวน 406 ราย (มีชนิดหลอดเลือดสมองแตก 71 ราย หรือร้อยละ 17.5) ข้อมูลถูกแบ่งเป็นชุดฝึกสอน (70%) และชุดทดสอบ (30%) และพัฒนาแบบจำลอง 3 รูปแบบ ได้แก่ logistic regression (LR), random forest (RF) และ XGBoost โดยมีการคัดเลือกตัวแปรด้วยเทคนิค LASSO และประเมินผลผ่านค่า AUROC, AUPRC, ความสอดคล้องของการทำนาย (calibration) และวิเคราะห์อิทธิพลของตัวแปรด้วย SHAP

ผลการศึกษา: แบบจำลอง random forest มีความสามารถในการจำแนกโรคภาพรวม (AUROC) สูงที่สุดที่ 0.811 อย่างไรก็ตาม เมื่อพิจารณาถึงความไม่สมดุลของชุดข้อมูล (imbalanced data) พบว่าแบบจำลอง logistic regression มีประสิทธิภาพและความน่าเชื่อถือสูงสุด โดยให้ค่า AUPRC ที่ 0.612 และมีความสอดคล้องของการทำนายดีที่สุด (calibration slope 0.986) ในแง่การนำไปใช้ทางคลินิก LR ให้ความไว (sensitivity) สูงถึงร้อยละ 94.1 ซึ่งเหมาะสำหรับการเป็นด่านแรกเพื่อคัดออก (rule-out) ในขณะที่ RF ให้ค่าการทำนายผลลบ (NPV) ที่ร้อยละ 89.3 ซึ่งช่วยสร้างเชื่อมั่นในการวินิจฉัย ส่วนแบบจำลอง XGBoost พบปัญหาการยึดติดข้อมูล (overfitting) จากขนาดตัวอย่างที่จำกัด นอกจากนี้ การวิเคราะห์ด้วย SHAP เผยให้เห็นปัจจัยหลักที่ขับเคลื่อนความเสี่ยง ได้แก่ ความดันโลหิตตัวบน

แบบจำลองการเรียนรู้ของเครื่อง เพื่อทำนายภาวะโรคหลอดเลือด สมองชนิดหลอดเลือดสมองแตก ในผู้ป่วยโรคหลอดเลือดสมอง เฉียบพลัน ณ ห้องฉุกเฉิน โรงพยาบาลสิรินธร จังหวัดขอนแก่น

เจษฎา อัครภาควัย

เจษฎา อัครภาควัย

นายแพทย์ชำนาญการ สาขาเวชศาสตร์ฉุกเฉิน
โรงพยาบาลสิรินธร จังหวัดขอนแก่น

ผู้รับผิดชอบบทความ:

นายแพทย์ เจษฎา อัครภาควัย

นายแพทย์ชำนาญการ สาขาเวชศาสตร์ฉุกเฉิน โรงพยาบาลสิรินธร
เลขที่ 81 หมู่ที่ 10 ตำบลโนนสมบูรณ์ อำเภอบ้านแฮด

จังหวัดขอนแก่น 40110

Email: jessada.u@kkumail.com

ที่สูง, มีภาวะแขนขาอ่อนแรง, สูญอายุ, มีประวัติการดื่มแอลกอฮอล์ และระดับความรู้สึกร่างกาย (GCS) ที่ลดลง ซึ่งสะท้อนความสมเหตุสมผลทางคลินิกอย่างชัดเจน

สรุปผลการศึกษา: การประยุกต์ใช้แบบจำลองทั้งสามรูปแบบร่วมกันแบบผสมผสาน โดยใช้ข้อมูลเพียงข้อมูลพื้นฐานทางคลินิก สามารถเป็นเครื่องมือช่วยสนับสนุนการตัดสินใจ ณ จุดคัดกรอง โดยแบบจำลองนี้ไม่ได้ทดแทนการตรวจ CT scan แต่ช่วยให้แพทย์วางแผนการรักษาให้เหมาะสมกับบริบทมากยิ่งขึ้น

คำสำคัญ: โรคหลอดเลือดสมองแตก, การเรียนรู้ของเครื่อง, ห่วงฉุกเฉิน, ปัญหาประสิทธิภาพทางการแพทย์, โรงพยาบาลชุมชน

บทนำ

โรคหลอดเลือดในสมอง (stroke) เป็นสาเหตุสำคัญของความพิการและการเสียชีวิตเป็นอันดับ 3 ทั่วโลก¹ โดยโรคหลอดเลือดสมองชนิดหลอดเลือดแตก (hemorrhagic stroke) คิดเป็นร้อยละ 34.6 ของผู้ป่วยโรคหลอดเลือดในสมองทั่วโลก^{1,2} และประมาณร้อยละ 20 ในประเทศไทย³

มาตรฐานที่ใช้ในการวินิจฉัยแยกโรคระหว่างโรคหลอดเลือดสมองชนิดขาดเลือดและชนิดหลอดเลือดแตกคือการตรวจเอกซเรย์คอมพิวเตอร์สมอง (CT brain)⁴ ซึ่งทั้งสองภาวะมีแนวทางการรักษาที่แตกต่างกัน⁵ ในบริบทของโรงพยาบาลระดับตติยภูมิหรือโรงพยาบาลชุมชนขนาดใหญ่ เช่น โรงพยาบาลสิรินธร จังหวัดขอนแก่น แม้จะมีศักยภาพในการตรวจ CT scan แต่ยังคงเผชิญข้อจำกัดด้านการขาดแคลนประสาทศัลยแพทย์ (neurosurgeon)⁶⁻⁸ ดังนั้น การคัดกรองประเมินความเสี่ยงล่วงหน้า ณ ห่วงฉุกเฉิน เพื่อทำนายว่าผู้ป่วยที่สงสัยโรคหลอดเลือดสมองเฉียบพลันมีแนวโน้มเป็นชนิดโรคหลอดเลือดแตกหรือไม่ จึงมีความสำคัญต่อการเตรียมความพร้อมในการกู้ชีพเบื้องต้น การประสานงานปรึกษาแพทย์เฉพาะทาง และการส่งต่อผู้ป่วยได้อย่างทันเวลาที่⁵

ปัจจุบันการใช้ปัญญาประดิษฐ์และการเรียนรู้ของเครื่อง (machine learning: ML) ได้เข้ามามีบทบาทในการช่วยวินิจฉัยและพยากรณ์โรคทางการแพทย์

การศึกษานี้จึงมีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลองทางคณิตศาสตร์ด้วยเทคนิค ML สำหรับทำนายความเสี่ยงโอกาสของการเกิดโรคหลอดเลือดสมองชนิดหลอดเลือดแตก โดยใช้เพียงข้อมูลพื้นฐานทางคลินิกแรกรับของผู้ป่วยที่มาใช้บริการ ณ ห่วงฉุกเฉิน โรงพยาบาลสิรินธร จังหวัดขอนแก่น

วัตถุประสงค์

เพื่อศึกษาลักษณะทางคลินิกของกลุ่มผู้ป่วยโรคหลอดเลือดสมอง และพัฒนาแบบจำลอง ML สำหรับทำนายความเสี่ยงของการเกิดโรคหลอดเลือดสมองชนิดหลอดเลือดแตก ในผู้ป่วยโรคหลอดเลือดสมองที่มาใช้บริการที่ห่วงฉุกเฉินโรงพยาบาลสิรินธร จังหวัดขอนแก่น

วิธีการศึกษา

รูปแบบการศึกษาและประชากรเป้าหมาย

การศึกษานี้เป็นการศึกษาวิเคราะห์แบบย้อนหลัง (retrospective cohort study) ณ โรงพยาบาลเดียว (single-center) โดยใช้ข้อมูลจากเวชระเบียนผู้ป่วยอิเล็กทรอนิกส์ (electronic medical records: EMR) ของโรงพยาบาลสิรินธร จังหวัดขอนแก่น ซึ่งข้อมูลที่สามารถระบุตัวผู้ป่วย (เช่น hospital number: HN, visit number: VN, ชื่อ-สกุล, ที่อยู่) จะถูกลบออกก่อนนำไปศึกษา

การศึกษานี้ผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ โรงพยาบาลสิรินธร จังหวัดขอนแก่น เป็นที่เรียบร้อยแล้ว (รหัสโครงการพิจารณา REC68010/เลขที่หนังสือรับรอง ECCERT68010) และใช้แนวทางการรายงานผลการศึกษาตามมาตรฐาน TRIPOD+AI (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis with Artificial Intelligence)¹⁰

เกณฑ์การรับกลุ่มตัวอย่างที่ศึกษา

ผู้ป่วยอายุ 18 ปีขึ้นไป ที่มารับบริการ ณ ห่วงฉุกเฉิน หรือได้รับการส่งต่อจากโรงพยาบาลเครือข่าย ระหว่างวันที่ 1 กันยายน 2567 ถึง 31 สิงหาคม 2568 โดยได้รับการวินิจฉัยตามรหัส ICD-10 กลุ่ม I60-I69 และได้รับการตรวจเอกซเรย์คอมพิวเตอร์สมอง

เกณฑ์การตัดกลุ่มตัวอย่างที่ศึกษา

ผู้ป่วยที่มีระยะเวลาเกิดอาการ (onset) มากกว่า 72 ชั่วโมง, มีประวัติอุบัติเหตุในการเข้ารับการรักษาคั้งนี้, ไม่ได้รับการตรวจ CT scan หรือข้อมูลเวชระเบียนไม่สมบูรณ์จนไม่สามารถนำมาวิเคราะห์ได้

การจัดการข้อมูลและตัวแปร

ตัวแปรที่ใช้ได้แก่ 1.ข้อมูลประชากรเบื้องต้น (อายุ เพศ ดัชนีมวลกาย) 2.ช่วงเวลาที่มารับบริการ 3. ระดับความฉุกเฉิน (emergency severity index level: ESI level) 4. อาการนำ (present illness) ได้แก่ อ่อนแรง พูดไม่ชัด ปากเบี้ยว ชี้นิ้วเวียนชา ไม่พูด ตามัว เดินเซ ปวดศีรษะ 5.ประวัติโรคประจำตัว ได้แก่ โรคความดันโลหิตสูง โรคเบาหวาน โรคไขมันในเลือดผิดปกติ โรคหลอดเลือดสมอง และ โรคหัวใจ 6.ประวัติการสูบบุหรี่ และดื่มเครื่องดื่มแอลกอฮอล์ 7.ความดันโลหิตตัวบนและตัวล่าง (systolic blood pressure: SBP/ diastolic blood pressure: DBP) 8.ระดับน้ำตาลในเลือดปลายนิ้ว (point of care testing blood sugar) 9.ระดับความรู้สึกตัว (Glasgow Coma Scale: GCS) 10.ระดับกำลังกล้ามเนื้อ (motor power grading) 11. ความรุนแรงของอาการทางระบบประสาท National Institutes of Health Stroke Scale (NIHSS) score

Outcome

ผลลัพธ์หลัก คือ โอกาสที่จะเป็นโรคหลอดเลือดสมองชนิดหลอดเลือดแตก วินิจฉัยโดยใช้แบบจำลองที่ใช้ข้อมูลที่มีอยู่ที่ห้องฉุกเฉิน ก่อนจะทราบ ผล CT scan brain (gold standard)

การพัฒนาแบบจำลอง

การคำนวณกลุ่มตัวอย่าง ใช้การคำนวณโดยใช้แนวทางของ Riley et al.¹¹ hundreds of prediction models are being developed to predict health outcomes in individuals. For continuous outcomes, typically a linear regression model is developed to predict an individual's outcome value conditional on values of multiple predictors (covariates ด้วยโปรแกรม R (package *pmsampsize* module) โดยกำหนดความชุกของโรคหลอดเลือดสมองชนิดหลอดเลือด

แตกในประเทศไทยที่ร้อยละ 23¹², Cox-Snell R-squared = 0.2, apparent & adjusted R-squared < 0.05 , shrinkage factor ≥ 0.9 และ กำหนดจำนวนตัวแปรอิสระที่คาดว่าจะนำเข้าสู่แบบจำลอง (candidate predictor parameters) 7 ตัวแปร พบว่าขนาดตัวอย่างขั้นต่ำที่ต้องการคือ **279 ราย** โดยต้องมีจำนวนเหตุการณ์ (events) อย่างน้อย 65 ราย ในการศึกษาที่มีกลุ่มตัวอย่างที่เข้าเกณฑ์ **406 ราย** ซึ่งมีจำนวนผู้ป่วยโรคหลอดเลือดสมองชนิดหลอดเลือดแตก 71 ราย จึงถือว่ามีขนาดตัวอย่างเพียงพอตามเกณฑ์ที่กำหนด ข้อมูลตัวอย่างจะถูกนำมาแบ่งกลุ่มผ่านวิธีการ stratified and randomized โดยแบ่งเป็นชุดข้อมูลที่ใช้ฝึกแบบจำลอง (70%) และทดสอบแบบจำลอง (30%) เพื่อรักษาสัดส่วนของผลลัพธ์ให้สมดุล

การจัดการข้อมูลสูญหาย ข้อมูลตัวแปรต่อเนื่อง (continuous predictors) ใช้วิธีแทนค่าด้วยค่ามัธยฐาน (median) ส่วนข้อมูลตัวแปรจัดกลุ่ม (categorical predictors) ใช้วิธีแทนที่ด้วยอัลกอริทึม k-Nearest Neighbors (k-NN, K=5) โดยคำนวณพารามิเตอร์จากชุดข้อมูลฝึกสอนเท่านั้นเพื่อป้องกันข้อมูลรั่วไหล (data leakage)

การคัดเลือกตัวแปร (feature selection and engineering) เนื่องจากชุดข้อมูลตั้งต้นมีตัวแปรจำนวนมากเพื่อให้แบบจำลองสามารถนำไปปฏิบัติได้จริงทางคลินิก (clinical constraint) ผู้วิจัยจึงดำเนินการจัดการข้อมูล 2 ขั้นตอน โดยขั้นตอนแรก ใช้การรวมกลุ่มตัวแปร (feature engineering) เพื่อลดความซ้ำซ้อน เป็นตัวแปรใหม่ 2 ตัวแปร ได้แก่ 1."ปัจจัยเสี่ยงภาวะหลอดเลือดอุดตัน" (atherosclerotic risk) โดยกำหนดให้มีค่าเมื่อพบประวัติอย่างน้อย 1 ข้อ: การสูบบุหรี่, โรคความดันโลหิตสูง, โรคเบาหวาน, หรือโรคไขมันในเลือดผิดปกติ) 2."ปัจจัยเสี่ยงเลือดออก" (bleeding risk) โดยกำหนดให้มีค่าเมื่อพบประวัติโรคหลอดเลือดสมองหรือโรคหัวใจ ที่ต้องรักษาด้วยการทานยาต้านเกร็ดเลือดหรือยาละลายลิ่มเลือด) ขั้นตอนต่อมาการคัดเลือกตัวแปรด้วยเทคนิค LASSO (least absolute shrinkage and selection operator) เพื่อคัดเลือกตัวแปรอิสระที่มี

ความสำคัญได้แก่ อายุ, ความดันโลหิตตัวบน (SBP), และ GCS score ให้คงอยู่ในสมการ และคัดเลือกว่าตัวแปรเพิ่มเติมจากกลุ่มตัวแปรที่เหลือจนครบ 7 ตัวแปร ได้แก่ อายุ, ความดันโลหิตตัวบน (SBP) ระดับความรู้สึกตัว ปัจจัยเสี่ยงภาวะหลอดเลือดอุดตัน ระดับกำลังกล้ามเนื้อ ประวัติการดื่มเครื่องดื่มแอลกอฮอล์ และปัจจัยเสี่ยงเลือดออก

ผู้วิจัยพัฒนาแบบจำลอง 3 อัลกอริทึม ได้แก่ logistic regression, random forest และ XGBoost (eXtreme gradient boosting) โดยมีการปรับแต่งพารามิเตอร์ (hyperparameter tuning) ผ่านกระบวนการ 5-fold cross-validation บนชุดข้อมูลฝึกสอน

การประเมินผลแบบจำลอง

ประเมินประสิทธิภาพผ่านพื้นที่ใต้โค้ง (AUROC: area under the receiver operating characteristic curve), พื้นที่ใต้โค้งความแม่นยำและการเรียกคืน (AUPRC: area under the precision-recall curve)¹³, ความสอดคล้องของการทำนาย (calibration slope/intercept, brier score), ความไว (sensitivity), ความจำเพาะ (specificity), PPV และ NPV โดยเลือกจุดตัดที่ดีที่สุดด้วย Youden's Index การอธิบายความสมเหตุสมผลของแบบจำลองใช้ค่า SHAP (SHapley Additive exPlanations) การวิเคราะห์ทางสถิติและการพัฒนาแบบจำลองทั้งหมดดำเนินการโดยใช้โปรแกรม RStudio version 2026.01.0 Build 392

ผลการศึกษา

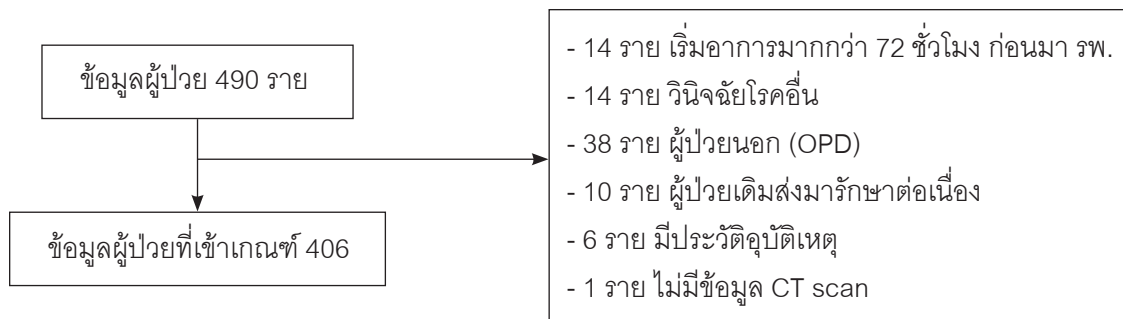
ข้อมูลพื้นฐานทางคลินิก

จากข้อมูลตัวอย่างทั้งหมด 490 ราย มีข้อมูลจำนวน 84 รายถูกคัดออกเนื่องจากจากข้อมูลไม่เข้าเกณฑ์การศึกษาวิจัย มีผู้ป่วยโรคหลอดเลือดสมองเข้าเกณฑ์การศึกษา 406 ราย (รูปที่ 1) แบ่งเป็นกลุ่มชนิดหลอดเลือดแตก 71 ราย (ร้อยละ 17.5) และกลุ่มชนิดขาดเลือด 335 ราย (ร้อยละ 82.5)

เมื่อเปรียบเทียบลักษณะทางคลินิก (ตารางที่ 1) ผู้ป่วยกลุ่มโรคหลอดเลือดสมองชนิดหลอดเลือดแตกมีอายุน้อยกว่ากลุ่มโรคหลอดเลือดสมองชนิดขาดเลือดอย่างมีนัยสำคัญ (มัธยฐานอายุ 58.0 ปี (IQR 48.0-72.0) เทียบกับ 65.0 ปี (IQR 57.0-74.0); $P = 0.009$) ในขณะที่สัดส่วนของเพศไม่พบความแตกต่างอย่างมีนัยสำคัญระหว่างสองกลุ่ม ($P = 0.19$)

ด้านข้อมูลทางคลินิกแรกเริ่ม ผู้ป่วยโรคหลอดเลือดสมองชนิดหลอดเลือดแตกพบความดันโลหิตสูงกว่าอย่างชัดเจน ทั้งความดันโลหิตตัวบน (systolic blood pressure) มัธยฐานอยู่ที่ 164 mmHg (IQR 150.0 - 189.0) เทียบกับ 154 mmHg (IQR 135.0 - 173.0 ; $P=0.001$) และความดันโลหิตตัวล่าง (diastolic blood pressure) มัธยฐาน 98 mmHg (IQR 76.0 - 96.0) เทียบกับ 86 mmHg (IQR 83.0 - 113.0) ; $P<0.001$) นอกจากนี้ ผู้ป่วยกลุ่มนี้ยังมีระดับความรู้สึกตัว Glasgow Coma Scale (GCS) ที่ต่ำกว่า (14 เทียบกับ 15 ; $P<0.001$) และมาโรงพยาบาลด้วยอาการความรู้สึกตัวเปลี่ยนแปลง (altered level of consciousness) ในสัดส่วนที่สูงกว่า (ร้อยละ 33.8 เทียบกับ ร้อยละ 11.6; $P<0.001$) และตรวจพบกำลังของกล้ามเนื้อที่ระดับต่ำมากกว่า ($P<0.001$) ซึ่งสอดคล้องกับการถูกจัดระดับความฉุกเฉินเป็น ESI level 1 ที่พบในกลุ่มโรคหลอดเลือดสมองชนิดหลอดเลือดแตกมากกว่า (ร้อยละ 32.4 เทียบกับ 5.7; $P<0.001$) อย่างไรก็ตาม เมื่อประเมินความรุนแรงของอาการทางระบบประสาทด้วยแบบประเมิน NIHSS ไม่พบความแตกต่างทางสถิติระหว่างผู้ป่วยทั้งสองกลุ่ม (median score 0 vs. 0; $P=0.31$)

สำหรับปัจจัยเสี่ยง พบว่าประวัติการดื่มแอลกอฮอล์ (35.2% vs. 13.0%; $P<0.001$) และการสูบบุหรี่ (23.9% vs. 13.1%; $P=0.028$) พบในกลุ่มชนิดหลอดเลือดสมองแตกมากกว่ากลุ่มชนิดขาดเลือดอย่างมีนัยสำคัญ ในขณะที่โรคประจำตัวเดิม ได้แก่ ความดันโลหิตสูง เบาหวาน และไขมันในเลือดสูง ไม่พบความแตกต่างทางสถิติระหว่างทั้งสองกลุ่ม



รูปที่ 1: แผนภาพกระบวนการคัดเลือกกลุ่มตัวอย่าง (Patient Flow Diagram)

ตารางที่ 1 ข้อมูลผู้ป่วยที่เข้าเกณฑ์การศึกษา

	Overall (N=406)	Non - Hemorrhagic Stroke (N=335)	Hemorrhagic Stroke (N=71)	p-value
อายุ*	65.0 (56.0 - 74.0)	65.0 (57.0 - 74.0)	58.0 (48.0 - 72.0)	0.009
เพศชาย - ราย(%)	239 (58.9)	192 (57.3)	47 (66.2)	0.19
ความดันโลหิตตัวบน*	156.0 (138.0 - 175.0)	154.0 (135.0 - 173.0)	164.0 (150.0 - 189.0)	0.001
ความดันโลหิตตัวล่าง*	87.0 (77.0 - 99.0)	86.0 (76.0 - 96.0)	98.0 (83.0 - 113.0)	<0.001
BMI*	23.1 (20.3 - 26.0)	23.4 (20.6 - 26.2)	22.1 (19.7 - 25.7)	0.12
GCS score*	15.0 (15 - 15)	15.0 (15 - 15)	14.0 (10 - 15)	<0.001
POCT blood sugar*	125.0 (100.0 - 172.0)	125.0 (99.0 - 172.5)	125.0 (106.0 - 169.0)	0.38
ช่วงเวลาให้บริการ (%)				0.047
8.00 - 16.00	208 (51.2)	176 (52.5)	32 (45.1)	
16.00 - 24.00	149 (36.7)	125 (37.3)	24 (33.8)	
24.00 - 8.00	49 (12.1)	34 (10.1)	15 (21.1)	
การมารับบริการ (%)				0.24
มาเอง	189 (46.6)	152 (45.4)	37 (52.1)	
ส่งต่อจากโรงพยาบาล	203 (50.0)	173 (51.6)	30 (42.3)	
กู้ชีพนำส่ง	14 (3.4)	10 (3.0)	4 (5.6)	
ระดับความฉุกเฉิน ESI				<0.001
Level1 (%)	42 (10.3)	19 (5.7)	23 (32.4)	
Level2 (%)	268 (66.0)	223 (66.6)	45 (63.4)	
Level3 (%)	96 (23.6)	93 (27.8)	3 (4.2)	
อาการที่พบ (%)				
อ่อนแรง	266 (65.5)	216 (64.5)	50 (70.4)	0.41
พูดไม่ชัด ลิ้นแข็ง	231 (56.9)	197 (58.8)	34 (47.9)	0.11
ใบหน้าเบี้ยว	151 (37.2)	129 (38.5)	22 (31.0)	0.28
ซีมี	63 (15.5)	39 (11.6)	24 (33.8)	<0.001
วิงเวียน	45 (11.1)	42 (12.5)	3 (4.2)	0.058
ชา	49 (12.1)	46 (13.7)	3 (4.2)	0.026

	Overall (N=406)	Non - Hemorrhagic Stroke (N=335)	Hemorrhagic Stroke (N=71)	p-value
ไม่พูด	31 (7.6)	23 (6.9)	8 (11.3)	0.22
ตาพร่ามัว	2 (0.5)	1 (0.3)	1 (1.4)	0.32
เดินเซ	17 (4.2)	17 (5.1)	0 (0.0)	0.052
ปวดศีรษะ	13 (3.2)	9 (2.7)	4 (5.6)	0.26
เวลาที่มีอาการ (%)				<0.001
น้อยกว่า 4.5 ชั่วโมง	215 (53.0)	165 (49.3)	50 (70.4)	
4.5 - 24 ชั่วโมง	151 (37.2)	133 (39.7)	18 (25.4)	
24 - 48 ชั่วโมง	27 (6.7)	27 (8.1)	0 (0.0)	
48 - 72 ชั่วโมง	13 (3.2)	10 (3.0)	3 (4.2)	
ปัจจัยเสี่ยง (%)				
โรคความดันโลหิตสูง	215 (53.0)	178 (53.1)	37 (52.1)	0.90
โรคเบาหวาน	132 (32.5)	114 (34.0)	18 (25.4)	0.17
โรคไขมันในเลือดสูง	47 (11.6)	43 (12.8)	4 (5.6)	0.10
โรคหลอดเลือดสมอง	57 (14.0)	51 (15.2)	6 (8.5)	0.19
โรคหัวใจ	35 (8.6)	31 (9.3)	4 (5.6)	0.48
สูบบุหรี่	61 (15.0)	44 (13.1)	17 (23.9)	0.028
ดื่มเครื่องดื่มแอลกอฮอล์	67 (16.5)	42 (12.5)	25 (35.2)	<0.001
กำลังกล้ามเนื้อที่ต่ำที่สุด (%)				<0.001
0	40 (9.9)	28 (8.4)	12 (16.9)	
1	39 (9.6)	16 (4.8)	23 (32.4)	
2	28 (6.9)	22 (6.6)	6 (8.5)	
3	74 (18.2)	64 (19.1)	10 (14.1)	
4	121 (29.8)	110 (32.8)	11 (15.5)	
5	104 (25.6)	95 (28.4)	9 (12.7)	
NIHSS Score*	0 (0.0 - 0)	0 (0 - 0)	0 (0 - 0)	0.31

* ข้อมูลต่อเนื่อง (continuous data) ค่าในวงเล็บแสดงถึงค่าร้อยละสำหรับข้อมูลเชิงกลุ่ม (categorical data) และค่าพิสัยระหว่างควอไทล์ (IQR)

† ทดสอบด้วยสถิติ Wilcoxon rank sum test สำหรับตัวแปรต่อเนื่องที่มีการแจกแจงแบบเบ้ (skewed) และ Fisher's exact test สำหรับตัวแปรเชิงกลุ่ม

‡ BMI = body mass index (ดัชนีมวลกาย), GCS score = Glasgow Coma Scale score (ระดับความรู้สึกตัว), POCT blood sugar = point of care testing blood sugar (ระดับน้ำตาลในเลือดปลายนิ้ว), ESI = emergency severity index (ระดับความฉุกเฉิน), NIHSS = National Institutes of Health Stroke Scale score (คะแนนความรุนแรงของอาการทางระบบประสาท)

§

¶

||

ตารางที่ 2 ประสิทธิภาพของแบบจำลอง

	Logistic Regression	Random Forest	XGBoost
Discrimination			
AUROC (95% CI)	0.736 (0.567-0.905)	0.811 (0.709-0.912)	0.781 (0.669-0.893)
AUPRC (95% CI)	0.612 (0.352-0.810)	0.549 (0.307-0.755)	0.403 (0.206-0.687)
Calibration			
Calibration Slope	0.986	1.078	0.691
Calibration Intercept	0.179	0.025	0.012
Brier Score	0.100	0.111	0.125
Diagnostic Performance*			
Selected Threshold	0.029	0.229	0.137
Sensitivity (95% CI)	0.941 (0.713–0.999)	0.529 (0.278–0.770)	0.588 (0.329–0.816)
Specificity (95% CI)	0.036 (0.008–0.102)	0.807 (0.706–0.886)	0.711 (0.601–0.805)
PPV (95% CI)	0.167 (0.098–0.256)	0.360 (0.180–0.575)	0.294 (0.151–0.475)
NPV (95% CI)	0.750 (0.194–0.994)	0.893 (0.801–0.953)	0.894 (0.794–0.956)

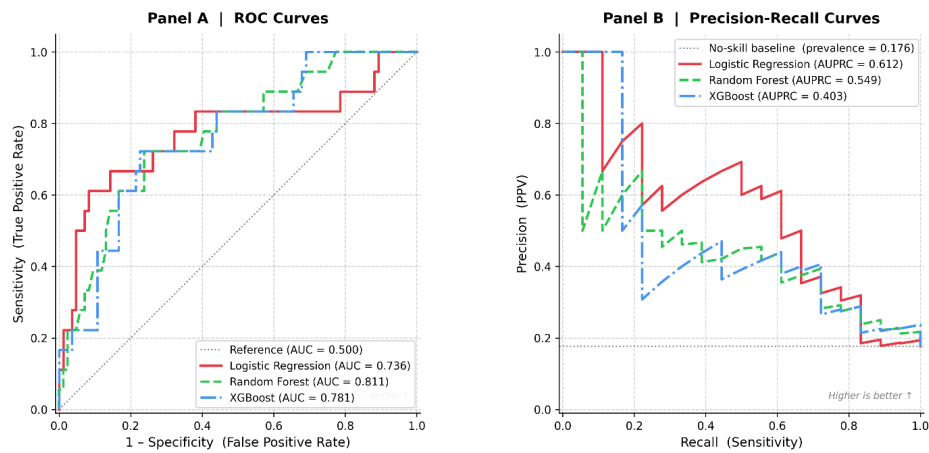
* AUROC = area under the receiver operating characteristic, AUPRC = area under the precision-recall curve, CI = Confidence interval, PPV = positive predictive value, NPV = negative predictive value

†
‡
§
¶
||

AUROC denotes area under the receiver-operating-characteristic curve; AUPRC, area under the precision-recall curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value. * Diagnostic performance metrics were calculated based on the optimal thresholds determined by Youden’s index or specific clinical constraints on the training set.

ประสิทธิภาพของแบบจำลอง
ความสามารถในการจำแนกโรค (discrimination performance): เมื่อทดสอบบนชุดข้อมูลทดสอบ (ตารางที่ 2) พบว่าแบบจำลอง random forest (RF) มีความสามารถในการจำแนกผู้ป่วยที่มีภาวะหลอดเลือดสมองชนิดหลอดเลือดแตกสูงสุด โดยมีพื้นที่ใต้โค้งอาร์โอซี (AUROC: 0.811 [95%CI 0.709 - 0.912]) ดังที่แสดงการเปรียบเทียบใน รูปที่ 2 (ROC curve) รองลงมาคือ XGBoost และ logistic regression (LR) มีค่า AUROC

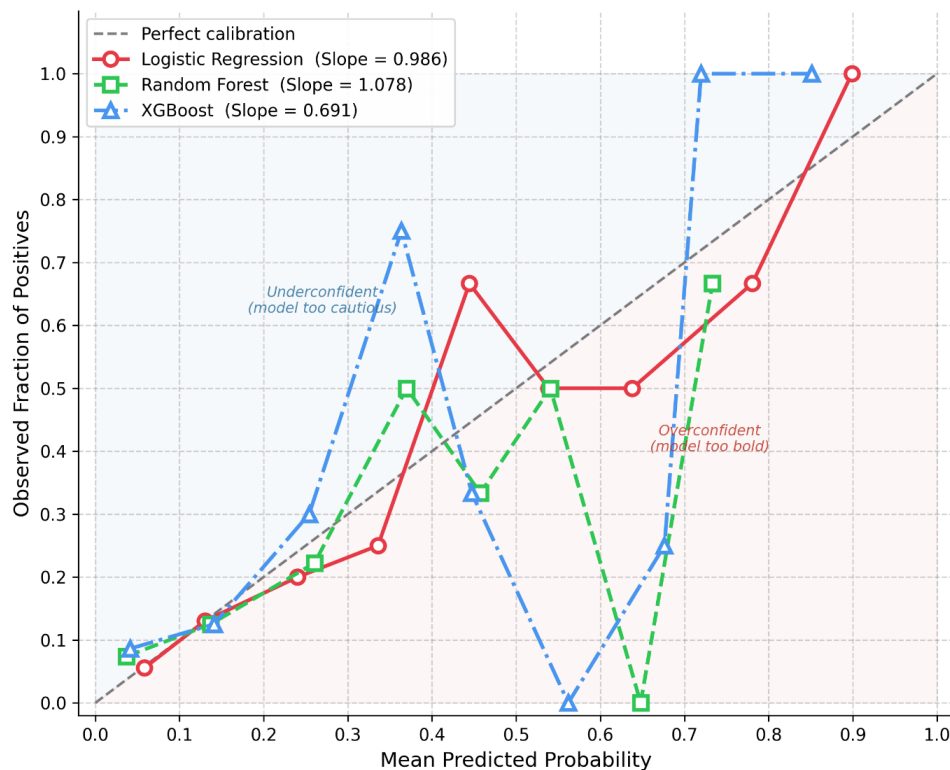
รองลงมาที่ (0.781 [95%CI 0.669-0.893]) และ (0.736 [95%CI 0.567 – 0.905]) ตามลำดับ อย่างไรก็ตาม เมื่อพิจารณาค่า AUPRC ซึ่งสะท้อนประสิทธิภาพในชุดข้อมูลที่มีความไม่สมดุล (imbalanced data) พบว่า Logistic Regression มีค่าสูงสุดที่ (AUPRC: 0.612 [95%CI 0.352 - 0.810]) สอดคล้องกับลักษณะเส้นโค้งใน รูปที่ 2 รองลงมาคือ RF (0.549 [95%CI 0.307 - 0.755]) และ XGBoost ที่ (0.403 [95% CI 0.206 - 0.687])



รูปที่ 3 การเปรียบเทียบประสิทธิภาพในการจำแนกโรค (discrimination performance)

ความสอดคล้องของการทำนาย (calibration)
ระหว่างค่าความน่าจะเป็นที่แบบจำลองทำนายกับอัตรา
การเกิดโรคจริง ดังปรากฏใน รูปที่ 3 (Calibration Plot)
พบว่า logistic regression มีความน่าเชื่อถือในการทำ
นายดีที่สุด โดยมีค่า calibration slope ที่ 0.986 (ค่า
ใกล้เคียงอุดมคติที่ 1.0) และมีค่า brier score ต่ำที่สุดที่
0.100 แบบจำลอง random forest มีค่า calibration

slope ที่ 1.078 และ brier score ที่ 0.111 ในขณะที่แบบ
จำลอง XGBoost มีแนวโน้มที่จะยึดติดกับข้อมูลชุด
ฝึกสอนมากเกินไป (overfitting) และมีความมั่นใจในการทำ
นายความเสี่ยงสูงเกินจริง (overconfidence) โดยมี
ค่า calibration slope เพียง 0.691 และ brier score ที่
0.125

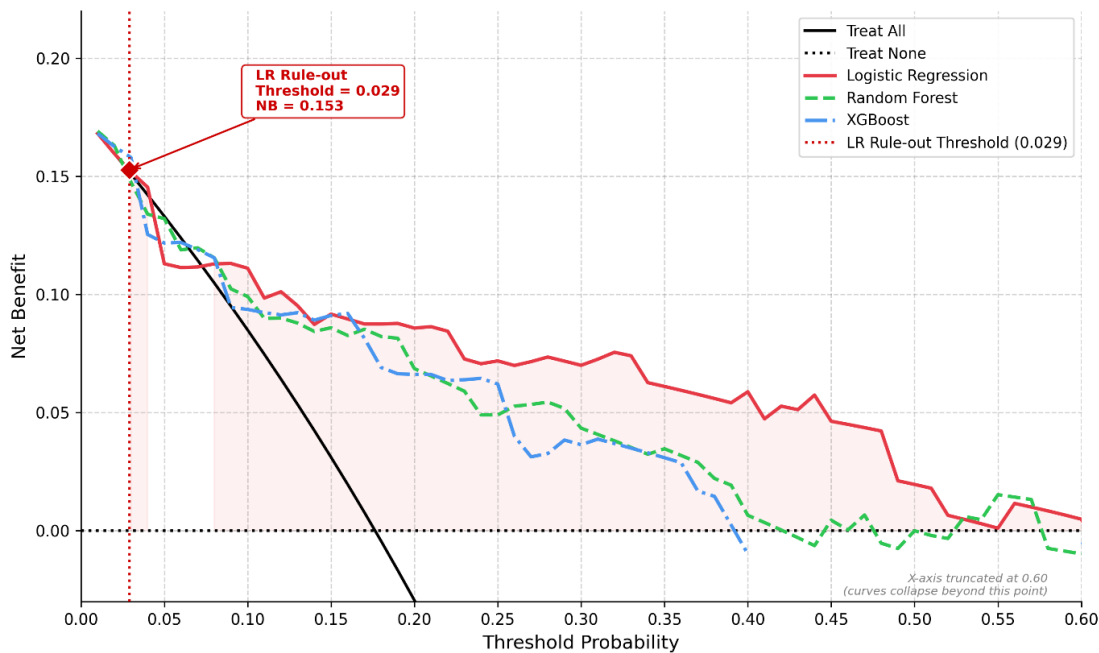


รูปที่ 3 ความสอดคล้องของการทำนาย (Calibration Plot / Reliability Diagram)

ประสิทธิภาพทางคลินิก (Clinical and Diagnostic Utility)

เมื่อวิเคราะห์จุดตัดที่เหมาะสม (threshold) ดังแสดงในรูปที่ 4 (Decision Curve Analysis) และ ตารางที่ 2 แบบจำลอง logistic regression (ที่จุดตัด 0.029) ให้ค่าความไว (sensitivity) สูงสุดที่ 0.941 (95% CI, 0.713 - 0.999) ซึ่งเหมาะสำหรับการคัดกรองเพื่อคัดออก (rule-out) แต่มีค่าความจำเพาะ (specificity) ต่ำเพียง 0.036 (95% CI, 0.008 ถึง 0.102) มีค่าการทำนายผล

บวก (PPV) ที่ 0.167 (95% CI, 0.098 - 0.256) สำหรับแบบจำลอง random forest (ที่จุดตัด 0.229) ให้ความสมมูลทางคลินิกที่ดีกว่า โดยมีความจำเพาะที่ 0.807 (95% CI, 0.706 ถึง 0.886) ความไวที่ 0.529 (95% CI, 0.278 ถึง 0.770) และค่าการทำนายผลลบ (NPV) ที่ 0.893 (95% CI, 0.801 ถึง 0.953) ส่วนแบบจำลอง XGBoost (ที่จุดตัด 0.137) มีความไวที่ 0.588 (95% CI, 0.329 ถึง 0.816) และความจำเพาะที่ 0.711 (95% CI, 0.601 ถึง 0.805)



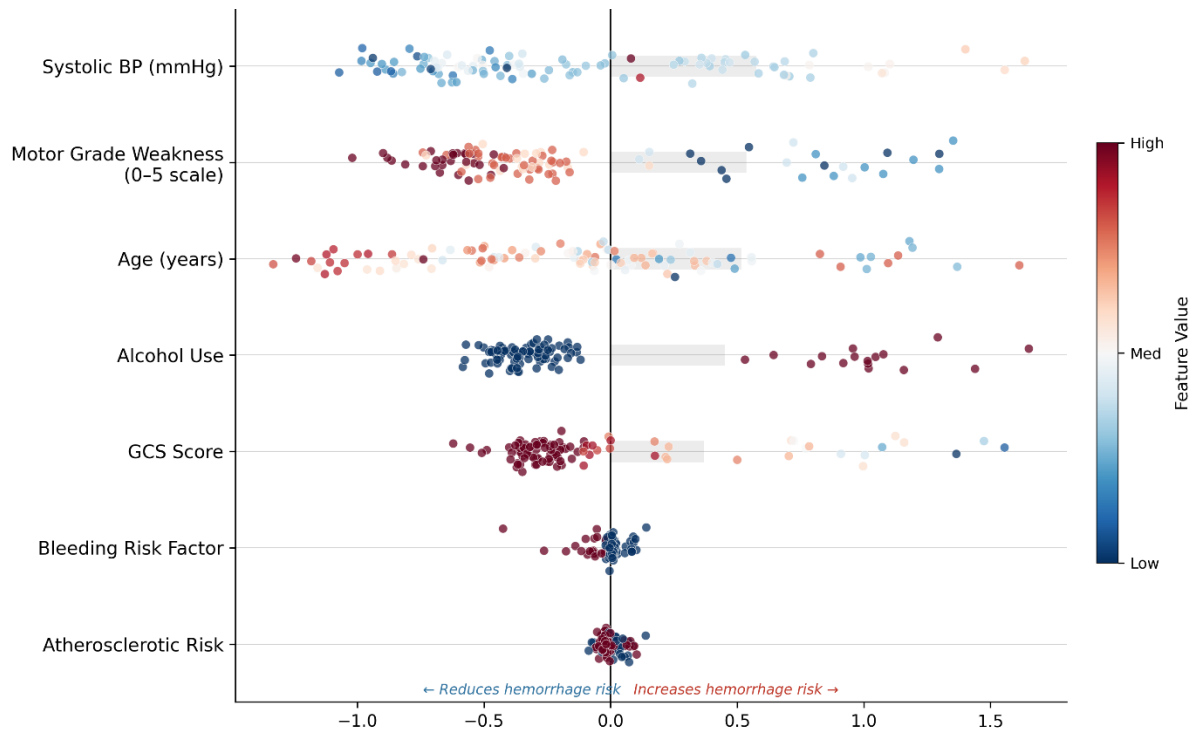
รูปที่ 4 การวิเคราะห์เส้นโค้งการตัดสินใจ

(Decision Curve Analysis; DCA) , LR = Logistic Regression , NB = Net Benefit

ความสามารถในการอธิบายแบบจำลอง (Model Explainability)

การประเมินอิทธิพลของตัวแปรผ่าน SHAP (SHapley Additive exPlanations) (รูปที่ 5) แสดงให้เห็นว่าตัวแปรที่ส่งผลต่อการเพิ่มความเสียหายของภาวะหลอดเลือดสมอง

ชนิดหลอดเลือดแตก 5 อันดับแรก ได้แก่ (1) ค่าความดันโลหิตตัวบนที่สูง, (2) การตรวจพบอาการแขนขาอ่อนแรง, (3) ช่วงอายุผู้ป่วยที่น้อยกว่าชนิดหลอดเลือดตีบ, (4) มีประวัติการดื่มแอลกอฮอล์ และ (5) ระดับความรู้สึกตัวที่ต่ำลง ล้วนสอดคล้องกับลักษณะของผู้ป่วยในตารางที่ 1



รูปที่ 5 SHAP (SHapley Additive exPlanations) Summary Plot

อภิปรายผล

การศึกษานี้พัฒนาแบบจำลองการเรียนรู้ของเครื่อง (machine learning) เพื่อทำนายความเสี่ยงของโรคหลอดเลือดสมองชนิดหลอดเลือดแตก โดยใช้เพียงข้อมูลพื้นฐานทางคลินิก ณ ห้องฉุกเฉิน ผลการศึกษาพบว่าแบบจำลอง RF ให้ความสามารถในการจำแนกโรค (AUROC) สูงที่สุดที่ 0.811 อย่างไรก็ตาม ในชุดข้อมูลการศึกษานี้มีความไม่สมดุลของข้อมูล (ความชุกของชนิดหลอดเลือดแตกเพียงร้อยละ 17.5) แบบจำลอง LR กลับมีความเหมาะสมที่สุดในด้านความสอดคล้องของการทำนาย และให้ค่า AUPRC สูงสุด (calibration slope 0.986, AUPRC 0.612) ซึ่งสอดคล้องกับการศึกษาของ Bhardwaj และคณะ ที่ระบุว่าชุดข้อมูลผู้ป่วยโรคหลอดเลือดสมองมักมีความไม่สมดุลสูง โดยมีสัดส่วนของชนิดหลอดเลือดชนิดขาดเลือด มักสูงถึงร้อยละ 84 และการศึกษาของ Silveira และคณะ ซึ่งมีสัดส่วนผู้ป่วยชนิดหลอดเลือดแตกเพียงร้อยละ 3 การใช้ค่าความแม่นยำ (accuracy) หรือ AUROC อาจทำให้ประสิทธิภาพของเครื่องมือที่สูงเกินจริง (overestimation)^{14,15}

สำหรับแบบจำลองที่มีความซับซ้อนอย่าง XGBoost พบว่าในการศึกษานี้เกิดปัญหายืดติดกับชุดข้อมูลมากเกินไป (overfitting) ซึ่งน่าจะเป็นผลมาจากข้อจำกัดด้านจำนวนตัวอย่างของเหตุการณ์ (events) ที่มีจำนวนน้อย อย่างไรก็ตาม เมื่อพิจารณาในบริบทของการใช้ชุดข้อมูลขนาดใหญ่ เช่น การศึกษาแบบ prospective multicenter ของ Hayashi และคณะ ที่ประเมินผู้ป่วยระยะก่อนถึงโรงพยาบาล (Prehospital) จำนวน 1,446 ราย พบว่า XGBoost เป็นแบบจำลองที่มีประสิทธิภาพในการทำนายโรคหลอดเลือดสมองชนิดหลอดเลือดแตก (AUROC 0.866)¹⁶ ดังนั้นสำหรับการศึกษาจากศูนย์ข้อมูลเดียว (single-center) การใช้แบบจำลองที่มีความซับซ้อนน้อยกว่าอย่าง LR หรือ RF เป็นทางเลือกที่เหมาะสมกว่า

ดังนั้น การนำแบบจำลองทั้งสามรูปแบบ ได้แก่ LR, RF และ XGBoost มาประยุกต์ใช้งานร่วมกัน (ensemble approach) ถือเป็นทางเลือกที่เหมาะสมในทางปฏิบัติ ซึ่งสอดคล้องกับข้อเสนอแนะของ Silveira และคณะ ที่สนับสนุนการรวมหลายอัลกอริทึมเข้าด้วยกันเพื่อดึงจุดเด่นมาชดเชยข้อจำกัดซึ่งกันและกัน¹⁵ เมื่อพิจารณาจากการศึกษานี้ แบบจำลอง LR มีความไว (sensitivity)

สูงถึงร้อยละ 94.1 จึงเหมาะสมที่จะเป็น “การคัดกรองเพื่อคัดออก” (rule-out) ป้องกันไม่ให้ผู้ป่วยตกหล่นควบคู่ไปกับการใช้ RF ที่มีค่าการทำนายผลลบ (NPV) สูงถึงร้อยละ 89.3 เพื่อช่วยสนับสนุนการบริหารจัดการคิวตรวจ CT scan หรือเตรียมการปรึกษาประสาทศัลยแพทย์ล่วงหน้าสำหรับผู้ป่วยกลุ่มความเสี่ยงสูง ส่วน XGBoost แม้ยังมีข้อจำกัดในปัจจุบัน แต่จะเป็นรากฐานสำคัญในการพัฒนาระบบที่รองรับข้อมูลขนาดใหญ่ขึ้นในอนาคต

นอกจากนี้ จุดเด่นที่สำคัญของแบบจำลองในการศึกษานี้คือความสมเหตุสมผลทางคลินิก (clinical plausibility) การวิเคราะห์ด้วย SHAP เผยให้เห็นว่าระดับความดันโลหิตตัวบน (SBP) ที่สูงขึ้น, ภาวะแขนขาอ่อนแรง, อายุที่เพิ่มขึ้น, และระดับความรู้สึกตัว (GCS) ที่ลดลง เป็นปัจจัยพยากรณ์หลักในการตัดสินใจของแบบจำลอง ซึ่งสอดคล้องโดยตรงกับการวิเคราะห์ SHAP ในการศึกษาของ Hayashi และคณะ ที่พบว่า ระดับ SBP, ภาวะแขนขาอ่อนแรง, และระดับความรู้สึกตัว เป็นตัวแปรที่มีอิทธิพลสูงสุดในระยะก่อนถึงโรงพยาบาล¹⁶ นอกจากนี้ งานวิจัยของ Liu และคณะ รวมถึง Silveira และคณะ ยังเน้นย้ำถึงบทบาทของระดับความดันโลหิตตัวบน (SBP) และอายุ ว่าเป็นปัจจัยพยากรณ์หลักทางคลินิกที่ส่งผลต่อการแยกชนิดของโรคและการพยากรณ์ผลลัพธ์การรักษา^{15,17} และในขณะที่แบบจำลองของ Silveira ใช้ข้อมูลภาพถ่ายรังสีเป็นตัวแปรหลักที่มีอิทธิพลสูงสุด¹⁵ แบบจำลองในการศึกษาของผู้วิจัยพิสูจน์ให้เห็นว่า แม้จะใช้เพียงข้อมูลทางคลินิกพื้นฐานก็สามารถใช้ในการทำนายโรคได้

การศึกษานี้มีข้อจำกัดหลายประการ ประการแรก รูปแบบการศึกษาแบบย้อนหลัง (retrospective) ทำให้ไม่สามารถควบคุมคุณภาพการบันทึกข้อมูลได้ มีโอกาสเกิดข้อมูลไม่ครบถ้วนหรือสูญหาย ประการที่สอง จำนวนเหตุการณ์ (event rate) ของกลุ่มโรคหลอดเลือดสมองชนิดหลอดเลือดแตกที่ค่อนข้างน้อย เป็นปัจจัยสำคัญที่ทำให้แบบจำลองที่มีความซับซ้อนสูงอย่าง XGBoost เกิดภาวะ overfitting แสดงให้เห็นถึงความจำเป็นในการใช้ข้อมูลที่มีขนาดใหญ่ขึ้น ประการที่สาม การศึกษานี้

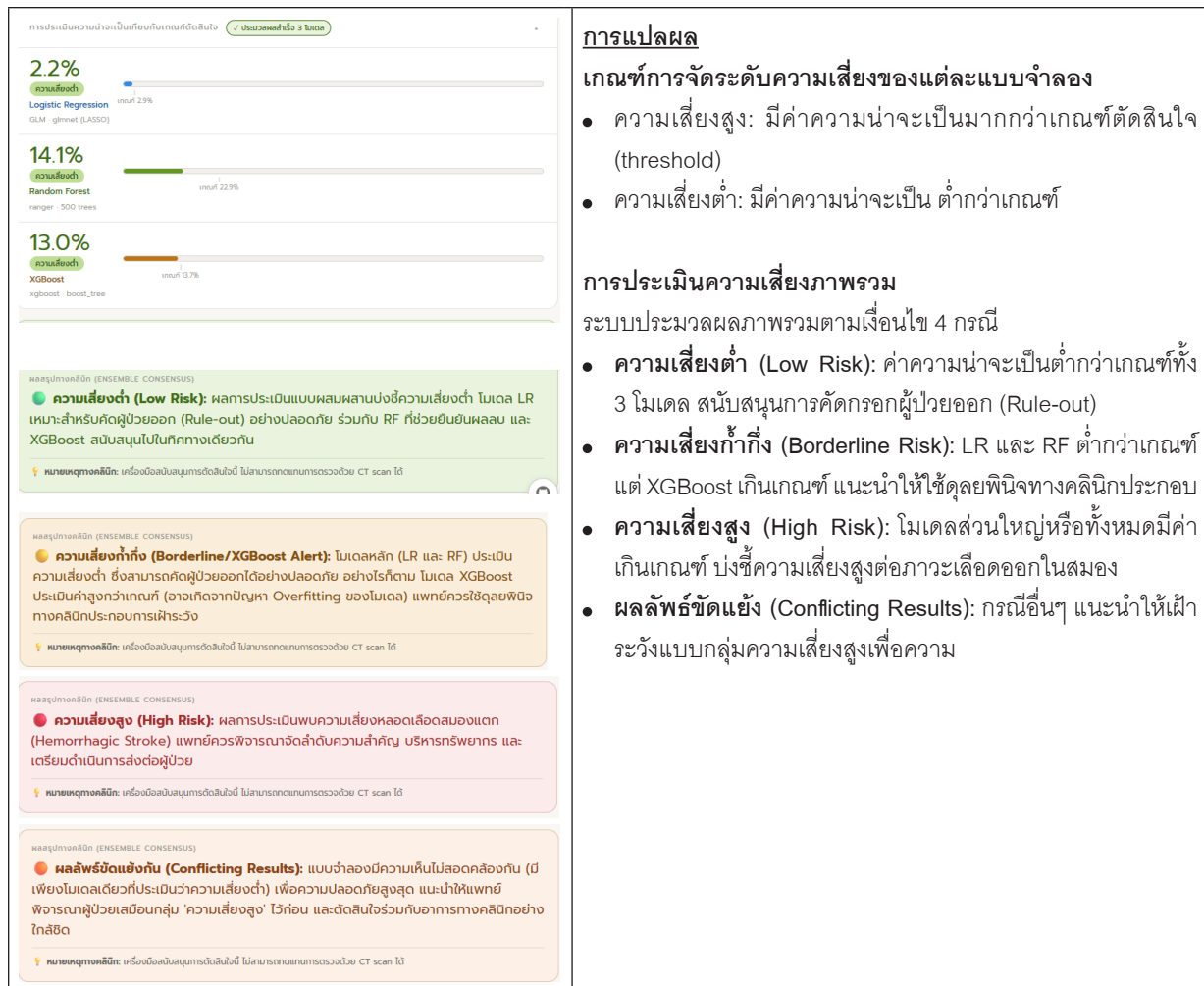
ดำเนินการในศูนย์ข้อมูลเดียว (single-center) ซึ่งกลุ่มประชากรและรูปแบบการส่งต่อผู้ป่วยอาจแตกต่างจากบริบทอื่น ประการสุดท้าย แบบจำลองนี้ถูกประเมินเฉพาะในกลุ่มผู้ป่วยที่ได้รับการยืนยันแล้วว่าเป็นโรคหลอดเลือดสมอง การใช้งานจริงในประชากรห้องฉุกเฉินทั่วไปซึ่งรวมถึงกลุ่มผู้ป่วยที่มีอาการคล้ายโรคหลอดเลือดสมอง (stroke mimics) อาจส่งผลให้ความจำเพาะ (specificity) ลดลง

สรุปผลการศึกษา

การประยุกต์ใช้แบบจำลองทั้ง 3 รูปแบบร่วมกันโดยอาศัยเพียงข้อมูลทางคลินิกพื้นฐานของผู้ป่วย มีศักยภาพสูงในการเป็นเครื่องมือสนับสนุนการตัดสินใจของแพทย์ ณ ห้องฉุกเฉิน โดยแบบจำลอง logistic regression ให้ค่าความไว (sensitivity) สูงจึงเหมาะสำหรับการใช้เพื่อคัดกรองผู้ป่วยที่ไม่สงสัยโรคหลอดเลือดสมองชนิดหลอดเลือดแตกออก (rule-out) ในขณะที่แบบจำลอง RF ให้ค่าการทำนายผลลบ (negative predictive value: NPV) ที่สูง ซึ่งช่วยยืนยันในการคัดกรองผู้ป่วยออก อย่างไรก็ตาม สำหรับแบบจำลอง XGBoost ยังคงพบปัญหาการเรียนรู้ข้อมูลมากเกินไป (overfitting) ซึ่งเป็นผลมาจากข้อจำกัดด้านขนาดของกลุ่มตัวอย่าง สุดท้ายการใช้งานแบบจำลองดังกล่าวไม่ได้มีวัตถุประสงค์เพื่อนำมาทดแทนการตรวจวินิจฉัยด้วยเครื่องเอกซเรย์คอมพิวเตอร์ (CT scan) แต่เป็นเครื่องมือเสริมที่ช่วยให้แพทย์ในโรงพยาบาลระดับทุติยภูมิสามารถจัดลำดับความสำคัญเร่งด่วน ให้การรักษาและดำเนินการส่งต่อผู้ป่วยได้อย่างปลอดภัยและมีประสิทธิภาพมากยิ่งขึ้น

ผู้วิจัยได้ทำตัวอย่างไว้เพื่อใช้ในการศึกษาที่ <https://tinyurl.com/MLHStroke>

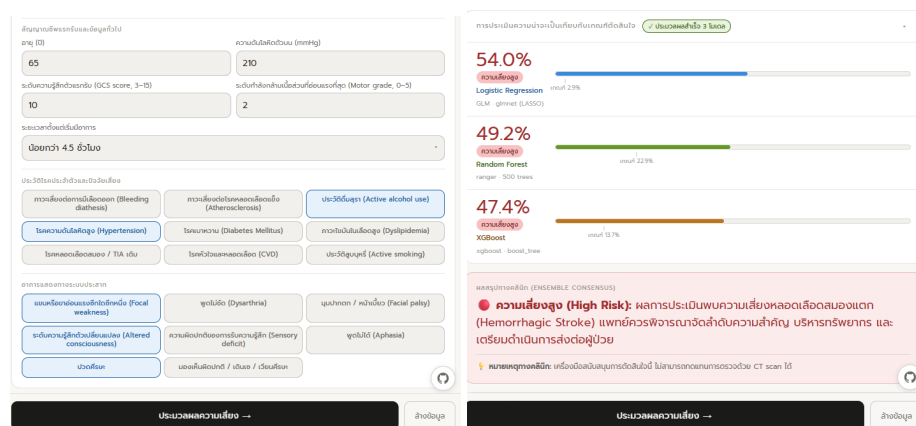




ตัวอย่างผู้ป่วย

ผู้ป่วยที่มีความเสี่ยงสูง (High Risk for Hemorrhagic Stroke)

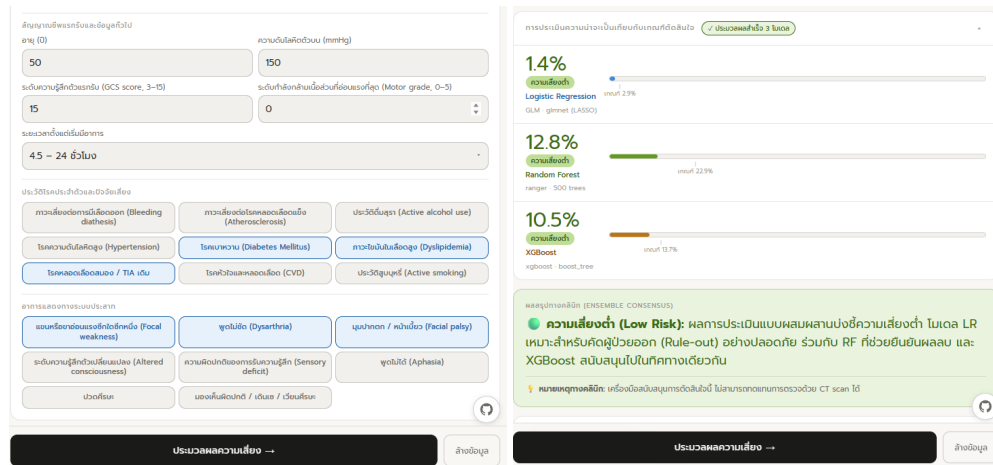
ผู้ป่วยชาย อายุ 65 ปี ถูกนำส่งห้องฉุกเฉินด้วยอาการปวดศีรษะรุนแรงเฉียบพลัน ตามด้วยอาการซึมลงและแขนขาซีกขวาล้มลงแรง ญาติแจ้งว่าเริ่มมีอาการเมื่อ 2 ชั่วโมงที่แล้ว ผู้ป่วยมีประวัติดื่มสุราเป็นประจำและเป็นโรคความดันโลหิตสูงแต่รับประทานยาไม่สม่ำเสมอ



รูปที่ 1 แสดงตัวอย่างการแสดงผลผู้ป่วยที่มีความเสี่ยงสูง

ผู้ป่วยที่มีความเสี่ยงต่ำ (Likely Ischemic Stroke / Low Hemorrhage Risk)

ผู้ป่วยหญิง อายุ 50 ปี มีอาการปากเบี้ยว พูดไม่ชัด และแขนขาซ้ายอ่อนแรง ญาติพบความผิดปกติเมื่อตื่นนอนตอนเช้า (ประเมินเวลา Onset ประมาณ 8 ชั่วโมงที่แล้ว) ผู้ป่วยยังรู้สึกตัวดี ทำตามสั่งได้ มีประวัติโรคเบาหวาน ไขมันในเลือดสูง และเคยมีประวัติ TIA (สมองขาดเลือดชั่วคราว) มาก่อน



รูปที่ 2 แสดงตัวอย่างการแสดงผลผู้ป่วยที่มีความเสี่ยงต่ำ

References

- Feigin VL, Brainin M, Norrving B, Martins SO, Pandian J, Lindsay P, et al. World Stroke Organization: Global Stroke Fact Sheet 2025. *Int J Stroke* 2025 ;20:132–44. doi:10.1177/17474930241308142 PubMed PMID: 39635884; PubMed Central PMCID: PMC11786524.
- van Asch CJ, Luitse MJ, Rinkel GJ, van der Tweel I, Algra A, Klijn CJ. Incidence, case fatality, and functional outcome of intracerebral haemorrhage over time, according to age, sex, and ethnic origin: a systematic review and meta-analysis. *Lancet Neurol* 2010 ;9:167–76. doi:10.1016/S1474-4422(09)70340-0 PubMed PMID: 20056489.
- Kumluang S, Wu O, Langhorne P, Geue C. Stroke resource utilisation and all-cause mortality in Thailand 2017–2020: A retrospective, cross-sectional study. *BMJ Open* 2023 ;13:e072259. doi:10.1136/bmjopen-2023-072259 PubMed PMID: 37280024; PubMed Central PMCID: PMC10255610.
- Hemphill JC, Greenberg SM, Anderson CS, Becker K, Bendok BR, Cushman M, et al. Guidelines for the management of spontaneous intracerebral hemorrhage. *Stroke* 2015 ;46:2032–60. doi:10.1161/STR.0000000000000069
- วีรสาร ก. คู่มือในการดูแลรักษาผู้ป่วยโรคหลอดเลือดสมองแตก ปี 2561. 1st ed. สถาบันประสาทวิทยา; 2562.
- ศานติธรรม ว, กลุ่มงานวิจัยและพัฒนา สส. วิจัยปริทัศน์ฉบับที่ 38 (ก.ย. 2566) [ปัญหาการขาดแคลนแพทย์ของประเทศไทย] [Internet]. 2023 Dec 1 [cited 2025 Jun 5]. Available from: <https://hdl.handle.net/20.500.14156/2001570>
- สถาบันวิจัยและประเมินเทคโนโลยีทางการแพทย์ กรมการแพทย์. Google Docs [Internet]. 2565 [cited 2025 Jun 5]. บัญชีเครื่องมือแพทย์ที่จำเป็นสำหรับประเทศไทย (Essential Medical Devices list) สาขาโรคหลอดเลือดสมอง. Available from: https://drive.google.com/file/d/1GZp4_omjvgqcuY-viZSfDObc_z4Z-6Dr_B/view?usp=sharing&usp=embed_facebook
- สถาบันวิจัยและประเมินเทคโนโลยีทางการแพทย์ กรมการแพทย์ กระทรวงสาธารณสุข. Google Docs [Internet]. 2555 [cited 2025 Jun 5]. จำนวน การกระจาย และลักษณะเครื่องตรวจวินิจฉัยทางรังสีอาศัยเทคโนโลยีขั้นสูงในประเทศไทย (Study of numbers, distribution and characteristics of high technology radiodiagnostic imaging in Thailand). Available from: https://drive.google.com/file/d/1jQ553n5c50rE6JWhdyNb3ZhHlIC5mVCZ/view?usp=drive_link&usp=embed_facebook
- Sirsat MS, Fermé E, Câmara J. Machine learning for brain stroke: A review. *J Stroke Cerebrovasc Dis Off J Natl*

- Stroke Assoc 2020 ;29:105162. doi:10.1016/j.jstrokecerebrovasdis.2020.105162 PubMed PMID: 32912543.
10. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Calster BV, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024 ;385:e078378. doi:10.1136/bmj-2023-078378 PubMed PMID: 38626948.
 11. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell Jr FE, Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: Part I – Continuous outcomes. *Stat Med* 2019;38:1262–75. doi:10.1002/sim.7993
 12. HDC Service [Internet]. [cited 2024 Jul 10]. Available from: <https://hdc.moph.go.th/center/public/standard-report-detail/db98b653287b753aab3b86222589f12e>
 13. ACM Other conferences [Internet]. [cited 2025 Dec 31]. The relationship between precision-recall and ROC curves | Proceedings of the 23rd international conference on Machine learning. Located at: world. Available from: <https://dl.acm.org/doi/10.1145/1143844.1143874> doi:10.1145/1143844.1143874
 14. Bhardwaj A, Srivastava MP, Wilson PV, Mehndiratta A, Vishnu VY, Garg R. Machine learning based reanalysis of clinical scores for distinguishing between ischemic and hemorrhagic stroke in low resource setting. *J Stroke Cerebrovasc Dis* 2022 ;31. doi:10.1016/j.jstrokecerebrovasdis.2022.106638 PubMed PMID: 35926404.
 15. Silveira SPP, Lima GDR, Alberto GB, Marcolino LCM, Xavier LB, Pereira CU, et al. Machine learning models to differentiate hemorrhagic from ischemic stroke: insights from the International Stroke Trial. *JBNC - J Bras Neurocri* 2026;37:169–83. doi:10.22290/jbnc.2026.02031047
 16. Hayashi Y, Shimada T, Hattori N, Shimazui T, Yoshida Y, Miura RE, et al. A prehospital diagnostic algorithm for strokes using machine learning: a prospective observational study. *Sci Rep* 2021 ;11:20519. doi:10.1038/s41598-021-99828-2 PubMed PMID: 34654860; PubMed Central PMCID: PMC8521587.
 17. Liu J, Niu S, Zhang R, Zhang B. Predictive modeling and feature analysis for clinical prognosis in hemorrhagic stroke patients using machine learning. In: Proceedings of the 2024 4th International Conference on Bioinformatics and Intelligent Computing [Internet]. New York, NY, USA: Association for Computing Machinery; 2024 [cited 2026 May 12]. p. 236–9. (BIC '24). Available from: <https://dl.acm.org/doi/10.1145/3665689.3665730> doi:10.1145/3665689.3665730