Evaluating a machine learning models for predicting full recovery in stroke: A case study at Hatyai Hospital

Nisjara Kunaton

Hatyai Hospital, Songkhla

Abstract

Stroke is a leading cause of death and disability worldwide, making the prediction of patient recovery crucial for treatment planning and rehabilitation. This study investigates the application of Machine Learning (ML) techniques to predict stroke patient recovery, using data from 6,081 cases at Hatyai Hospital. In the design and execution of the experiment, four ML models were utilized: Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting. The data preparation process involved label encoding, handling missing values, and balancing the dataset. Model performance was evaluated using K-Fold Cross-Validation and Hyperparameter Tuning. Results show that Random Forest performed the best, achieving an accuracy of 88.61%, with both Precision and Recall at 0.91, and an F1-Score of 0.91. Gradient Boosting followed closely with an accuracy of 88.50%. Logistic Regression and Decision Tree showed lower performance, with accuracies of 87.08% and 83.83%, respectively. The study demonstrates that ML techniques, particularly Random Forest and Gradient Boosting, offer high accuracy in predicting stroke recovery, providing valuable insights for efficient treatment planning.

Keywords: *Machine Learning, Barthel Index, Stroke*

Received: 10 June 2025, Revised: 25 July 2025, Accepted: 1 September 2025

Correspondence: Nisjara Kunaton, Hatyai Hospital, 182 Ratthakan, Tambon Hat Yai, Hat Yai District, Songkhla 90110, Thailand, Tel.: 074 273 100, email: nis.kunfm@gmail.com

การประเมินโมเดลแบบจำลองในการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง

นิศจรา กุณาฑลย์

โรงพยาบาลหาดใหญ่ สงขลา

บทคัดย่อ

โรคหลอดเลือดสมองเป็นสาเหตุหลักของการเสียชีวิตและความทุพพลภาพทั่วโลก การพยากรณ์การฟื้นตัวของผู้ป่วยจึงมีความสำคัญอย่างยิ่งในการวางแผนการรักษาและฟื้นฟู งานวิจัยนี้มุ่งเน้นศึกษาการใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning: ML) ในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง โดยใช้ข้อมูลผู้ป่วย 6,081 ราย จากโรงพยาบาลหาดใหญ่ ในกระบวนการออกแบบและดำเนินการทดลอง ผู้วิจัยได้เลือกใช้แบบจำลอง M 4 แบบ ได้แก่ Logistic Regression, Decision Tree, Random Forest, และ Gradient Boosting การเตรียมข้อมูลได้รวมถึงการเข้ารหัสป้ายกำกับ การจัดการค่าที่หายไป และการปรับสมดุลของชุดข้อมูล เทคนิคการประเมินประสิทธิภาพของโมเดลถูกดำเนินการด้วย K-Fold Cross-Validation และการปรับแต่งพารามิเตอร์ (Hyperparameter Tuning) สำหรับแต่ละโมเดล ผลการทดลองแสดงให้เห็นว่า Random Forest มีประสิทธิภาพสูงสุด โดยมีค่า Accuracy 88.61%, Precision และ Recall ที่ 0.91, และ F1-Score ที่ 0.91 ขณะที่ Gradient Boosting แสดงผลใกล้เคียงกันด้วย Accuracy 88.50% Logistic Regression และ Decision Tree มีประสิทธิภาพต่ำกว่า ด้วยค่า Accuracy ที่ 87.08% และ 83.83% ตามลำดับ งานวิจัยนี้ชี้ให้เห็นว่าเทคนิค ML โดยเฉพาะ Random Forest และ Gradient Boosting มีความสามารถในการช่วยพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมองอย่างแม่นยำ ซึ่งสามารถนำไปใช้ในการวางแผนการรักษาได้อย่างมีประสิทธิภาพ

คำสำคัญ: *การเรียนรู้ของเครื่อง, ดัชนีบาร์เทล, โรคหลอดเลือดสมอง*

วันที่รับต้นฉบับ: 10 มิถุนายน 2568, วันที่แก้ไข: 25 กรกฎาคม 2568, วันที่ตอบรับ: 1 กันยายน 2568

บทนำ

โรคหลอดเลือดสมองยังเป็นสาเหตุการตายอันดับที่ 2 และเป็นสาเหตุของความทุพพลภาพอันดับที่ 3 ของโลก [1] การดูแลผู้ป่วยโรคหลอดเลือดสมองเป็นภาระหนักที่ผู้ดูแลต้องรับมือ โดยเฉพาะในกรณีที่ผู้ป่วยมีความพึ่งพิงสูงในการทำกิจวัตรประจำวัน ระดับความพึ่งพิงของผู้ป่วยซึ่งวัดโดย Barthel Index (BI) มีผลกระทบอย่างมีนัยสำคัญต่อความตึงเครียดของผู้ดูแล [2] การประเมินผลลัพธ์ของผู้ป่วยโรคหลอดเลือดสมองมีความสำคัญอย่างยิ่งต่อการวางแผนการรักษาและการพยากรณ์โรค โดยเครื่องมือหลักที่ใช้ในการประเมินผลลัพธ์ได้แก่ Modified Rankin Scale (mRS), Barthel Index (BI) และ National Institutes of Health Stroke Scale (NIHSS) ซึ่งแต่ละเครื่องมือมีบทบาทเฉพาะในการประเมินการฟื้นตัวและความเป็นอิสระของผู้ป่วยหลังเกิดโรคหลอดเลือดสมอง [3] ปัจจัยทางคลินิก เช่น ความรุนแรงของโรค อายุ โรคร่วม ความสามารถในการทำกิจวัตรประจำวัน และการตอบสนองต่อการรักษา ล้วนมีผลต่อการประเมินด้วย NIHSS, mRS และ Barthel Index ซึ่งสามารถใช้ในการพยากรณ์ผลลัพธ์ของผู้ป่วยโรคหลอดเลือดสมองและวางแผนการรักษาที่เหมาะสมได้ [4]

การนำ Machine Learning (ML) มาใช้ในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง มีศักยภาพในการช่วยวิเคราะห์ปัจจัยของผู้ป่วย โดยอาศัยการวิเคราะห์ข้อมูลจำนวนมากและการเรียนรู้จากข้อมูลเหล่านั้น ทำให้สามารถพยากรณ์การฟื้นตัวของผู้ป่วยได้อย่างแม่นยำและรวดเร็วมากขึ้น ส่งผลให้การวางแผนการคัดกรอง การรักษาและปรับเปลี่ยนพฤติกรรมเสี่ยงเป็นไปอย่างมีประสิทธิภาพ ในยุคที่เทคโนโลยีมีความก้าวหน้าเช่นปัจจุบัน ผู้จัดทำจึงสนใจศึกษาการนำ Machine Learning (ML) มาใช้เพื่อพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง โดยการพัฒนาและประเมินผลโมเดล Machine Learning ในการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง

ผู้นิพนธ์ประสานงาน: นิศจรา กุณาฑลย์ โรงพยาบาลหาดใหญ่ 182 ถ.รัถการ อ.หาดใหญ่ จ.สงขลา 90110, โทร.: 074 273 100, email: nis.kunfm@gmail.com

ทบทวนวรรณกรรม

การทบทวนวรรณกรรมในบทความวิจัยนี้มุ่งเน้นไปที่การสำรวจและสรุปงานวิจัยที่เกี่ยวข้องกับการพัฒนาและการใช้ Machine Learning (ML) ในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง และเครื่องมือที่ใช้ในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง

1. การใช้ Machine Learning ในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง

แบบจำลอง Machine Learning (ML) ได้แสดงให้เห็นถึงศักยภาพในการพยากรณ์คะแนน Barthel Index (BI) สำหรับการฟื้นฟูของผู้ป่วยโรคหลอดเลือดสมอง โดยอาศัยข้อมูลทางคลินิกที่มีโครงสร้างเพื่อเพิ่มความแม่นยำในการพยากรณ์ งานวิจัยหลายชิ้นได้เน้นให้เห็นถึงประสิทธิภาพของเทคนิค ML ในการทำนายผลลัพธ์ของผู้ป่วยในด้านความสามารถในการฟื้นฟูร่างกาย [5]

2. ประสิทธิภาพของแบบจำลองในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง

จากการทบทวนวรรณกรรมอย่างเป็นระบบ งานวิจัยพบว่าแบบจำลอง ML มีความสามารถในการทำนายผลลัพธ์ได้อย่างมีประสิทธิภาพ โดยมีขนาดตัวอย่างเฉลี่ยของผู้ป่วยอยู่ที่ 475 รายในงานวิจัยต่าง ๆ ซึ่งชี้ให้เห็นว่า ML สามารถทำนายผลการฟื้นตัวได้อย่างแม่นยำในกลุ่มผู้ป่วยขนาดใหญ่

นอกจากนี้ ยังมีการเปรียบเทียบแบบจำลองการถดถอยโลจิสติกกับ Neural Networks ในการทำนายความเป็นอิสระในการทำกิจวัตรประจำวัน พบว่าไม่มีความแตกต่างอย่างมีนัยสำคัญในด้านความแม่นยำของการพยากรณ์ ซึ่งชี้ให้เห็นว่า วิธีการทางสถิติดั้งเดิมยังคงมีความน่าเชื่อถือในการพยากรณ์ [6]

3. Machine Learning Model แต่ละแบบ

Decision Tree และ Random Forest เป็นเทคนิคการเรียนรู้ของเครื่องที่ได้รับความนิยมอย่างมาก ทั้งสองมีข้อดีและข้อจำกัดที่แตกต่างกัน Decision Tree นั้นง่าย รวดเร็ว และสามารถตีความได้ง่าย ทำให้เหมาะสมกับงานจำแนกประเภทที่ไม่ซับซ้อน อย่างไรก็ตาม Decision Tree มักมีปัญหาเรื่องการเกิด overfitting ซึ่งอาจทำให้ไม่สามารถทำงานได้ดีเมื่อเจอกับข้อมูลที่ไม่เคยเห็นมาก่อน (Ho, 1995) ในทางกลับกัน Random Forest ได้พัฒนาและแก้ไขปัญหาดังกล่าวด้วยการนำต้นไม้ออกมารวมกันหลายต้น ทำให้เพิ่มความแม่นยำและความทนทานต่อการทำนายที่หลากหลาย [7] [8] [9] [10]

Gradient Boosting และ Random Forest เป็นเทคนิค ML โดยใช้ต้นไม้ (decision trees) โดย Gradient Boosting ให้ความแม่นยำสูงและเน้นที่ข้อมูลที่ยากต่อการพยากรณ์ แต่มีความซับซ้อนในการปรับแต่งและไวต่อการเกิด overfitting ส่วน Random Forest เป็นเทคนิคที่ง่ายต่อการใช้งาน ไม่ซับซ้อนและทนทานต่อข้อมูล noise แต่ไม่แม่นยำเท่า Gradient Boosting [11]

การเปรียบเทียบประสิทธิภาพของ Logistic Regression, Decision Tree, Random Forest และ Gradient Boosting ในการทำนายค่าดัชนี Barthel Index (BI) สำหรับผู้ป่วยโรคหลอดเลือดสมองแสดงให้เห็นถึงคุณลักษณะการทำงานที่แตกต่างกันของโมเดลการเรียนรู้ของเครื่องแต่ละประเภท โดย Logistic Regression มักแสดงผลความแม่นยำที่สูงและการระบุผลลัพธ์เชิงบวกที่สมบูรณ์แบบ (Recall) ซึ่งทำให้มีประสิทธิภาพในการระบุผลลัพธ์เชิงบวก [12] ขณะที่ Random Forest และ Gradient Boosting มักมีประสิทธิภาพสูงกว่าโมเดลอื่น ๆ โดยสามารถทำได้ถึงความแม่นยำ 97.36%, 97.73% ตามลำดับ [13] ส่วน Decision Tree แม้จะมีโครงสร้างที่ง่ายกว่า แต่มีแนวโน้มที่จะมีความแม่นยำและความแม่นตรง (Precision) ต่ำกว่า ซึ่งบ่งบอกถึงความท้าทายในการจัดการกับผลบวกที่ผิดพลาด (False Positives) [12] ส่วน Gradient Boosting เมื่อใช้งานร่วมกับเทคนิค stacking สามารถสร้างผลลัพธ์ที่ดียิ่งขึ้น โดยมีความแม่นยำสูงสุดถึง 98.85% [13] ดังนั้น การเลือกโมเดลอาจขึ้นอยู่กับวัตถุประสงค์เฉพาะ เช่น การให้ความสำคัญกับ Recall หรือความแม่นยำโดยรวม

ตัวแปรที่สำคัญในการทำนายได้แก่ ค่าเริ่มต้นของ Bathel index (BI) [14] แม้ว่า Logistic Regression จะโดดเด่นในด้านการระบุผลลัพธ์เชิงบวก (Recall) แต่ Random Forest และ Gradient Boosting ก็มีความแม่นยำสะท้อนให้เห็นว่าการเลือกโมเดลควรพิจารณาจากความต้องการทางคลินิกและลักษณะของข้อมูล

4. ข้อกังวลและข้อจำกัด

ในกระบวนการสร้างแบบจำลองการเรียนรู้ของเครื่อง (ML) ข้อมูลที่ไม่สมดุลเป็นปัญหาสำคัญที่ส่งผลกระทบต่อความแม่นยำและประสิทธิภาพของโมเดล โดยเฉพาะอย่างยิ่งเมื่อใช้แบบจำลอง Decision Tree และ Random Forest เนื่องจากทั้งสองอัลกอริทึมนี้มีปฏิสัมพันธ์กับข้อมูลในรูปแบบที่แตกต่างกัน จึงเกิดข้อกังวลและข้อจำกัดบางประการที่ผู้พัฒนาโมเดลต้องพิจารณา

แบบจำลองที่ใช้ Decision Tree โดยเฉพาะในกรณีที่มีหลายคลาส มักจะประสบปัญหาเมื่อข้อมูลมีการกระจายตัวไม่เท่ากันระหว่างคลาสที่เป็นส่วนใหญ่และคลาสที่เป็นส่วนน้อย เนื่องจาก Decision Tree มีแนวโน้มที่จะให้ความสำคัญกับคลาสส่วนใหญ่ ทำให้การจำแนกคลาสส่วนน้อยทำได้ไม่ดี [15] เทคนิคอย่างการทำ random over-sampling ซึ่งช่วยเพิ่มจำนวนข้อมูลในคลาสส่วนน้อย จะช่วยให้การทำงานของ Decision Tree ดีขึ้น [16] ส่วนแบบจำลองที่ใช้ Random Forest แสดง

ความทนทานมากกว่าต่อความไม่สมดุลของข้อมูล โดยยังสามารถรักษาประสิทธิภาพที่ดีได้แม้ไม่มีการปรับตัวอย่างข้อมูลมากนัก [17] อย่างไรก็ตาม การปรับปรุงด้วยอัลกอริทึม Random Forest แบบ balanced ซึ่งมีการปรับวิธีการสุ่มตัวอย่างให้สมดุล สามารถเพิ่มความแม่นยำในการจำแนกคลาสส่วนน้อยได้ดียิ่งขึ้น [18] [19] แม้ว่า Random Forest จะทำงานได้ดีกว่าภายใต้ข้อมูลที่ไม่สมดุล แต่ก็ยังสามารถเพิ่มประสิทธิภาพได้ด้วยกลยุทธ์การสุ่มตัวอย่างที่เหมาะสม สรุปแล้ว โมเดลทั้งสองต้องการการพิจารณาถึงการกระจายตัวของข้อมูลอย่างละเอียดเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด ซึ่งข้อมูลที่ไม่สมดุลส่งผลกระทบอย่างมากต่อประสิทธิภาพของโมเดลที่ใช้ในการพยากรณ์ดัชนี Barthel ในผู้ป่วยโรคหลอดเลือดสมอง งานวิจัยหลายชิ้นชี้ให้เห็นถึงความท้าทายที่เกิดขึ้นจากชุดข้อมูลที่มีความไม่สมดุล โดยเฉพาะในด้านการแพทย์ ซึ่งกลุ่มตัวอย่างส่วนใหญ่มักประกอบไปด้วยผู้ป่วยที่มีสภาพร่างกายปกติ ทำให้ความแม่นยำในการพยากรณ์กลุ่มผู้ป่วยที่มีอาการหนักหรือน้อยกว่าไม่สูงมากนัก [20]

แบบจำลอง ML กำลังได้รับความนิยมในการพยากรณ์ผลลัพธ์ของผู้ป่วยโรคหลอดเลือดสมอง แต่ยังมีข้อกังวลเกี่ยวกับมาตรฐานการรายงานและความสามารถในการทำซ้ำของผลการวิจัย ซึ่งชี้ให้เห็นถึงความจำเป็นในการพัฒนาแนวทางวิธีการที่ดีขึ้นในการนำ ML ไปใช้ในการดูแลผู้ป่วยทางคลินิก [5] การทบทวนวรรณกรรมแสดงให้เห็นว่า การใช้แบบจำลอง Machine Learning ในการพยากรณ์ Barthel Index (BI) มีศักยภาพสูงในการช่วยเพิ่มความแม่นยำในการประเมินการฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง แม้ว่า ML จะมีประสิทธิภาพสูงในงานวิจัยหลายชิ้น แต่ก็ยังมีความท้าทายที่ต้องได้รับการแก้ไข โดยเฉพาะในเรื่องของมาตรฐานการรายงานและการทำซ้ำผลการศึกษาเพื่อให้สามารถนำมาใช้ในทางคลินิกได้อย่างเต็มที่ [21] [22] [23]

5. การประเมินผลลัพธ์ของผู้ป่วยโรคหลอดเลือดสมอง: บทบาทของ mRS, Barthel Index (BI) และ NIH Stroke Scale (NIHSS)

การประเมินผลลัพธ์ของผู้ป่วยโรคหลอดเลือดสมองมีความสำคัญอย่างยิ่งต่อการวางแผนการรักษาและการพยากรณ์โรค โดยเครื่องมือหลักที่ใช้ในการประเมินผลลัพธ์ได้แก่ Modified Rankin Scale (mRS), Barthel Index (BI) และ National Institutes of Health Stroke Scale (NIHSS)

6. ภาพรวมของเครื่องมือการประเมิน

Modified Rankin Scale (mRS): mRS เป็นเครื่องมือที่ใช้วัดระดับความพิการโดยรวมของผู้ป่วย ซึ่งสัมพันธ์กับผลลัพธ์ในระยะยาวของผู้ป่วยโรคหลอดเลือดสมอง งานวิจัยโดย Taleb และคณะ (2024) พบว่า การประเมินในช่วงแรก (วันที่ 2 และ 4 หลังจากเกิดอาการ) มีความสอดคล้องกับผลลัพธ์ในระยะ 90 วันของผู้ป่วย อย่างไรก็ตาม ค่าที่ประเมินอาจมีความคลาดเคลื่อนในระยะยาวได้ [24]

Barthel Index (BI): Barthel Index เป็นเครื่องมือที่ใช้ในการประเมินความสามารถของผู้ป่วยในการทำกิจวัตรประจำวัน เช่น การกิน การแต่งตัว และการเคลื่อนไหว งานวิจัยโดย Fu และคณะ (2024) แสดงให้เห็นว่า BI มีความสามารถสูงในการพยากรณ์ผลลัพธ์ของผู้ป่วยที่ได้รับการรักษาด้วย thrombolytic therapy และในบริบทนี้ BI ให้ผลการพยากรณ์ที่ดีกว่า NIHSS การมีคะแนน BI ที่ต่ำมักจะบ่งบอกถึงการฟื้นตัวที่ไม่ดี และการพยากรณ์โรคที่แย่ลง [25]

NIH Stroke Scale (NIHSS): NIHSS เป็นเครื่องมือที่ใช้วัดความบกพร่องทางระบบประสาทของผู้ป่วย ซึ่งใช้กันอย่างแพร่หลายในช่วงแรกหลังจากที่ผู้ป่วยเกิดอาการโรคหลอดเลือดสมอง อย่างไรก็ตาม งานวิจัยของ Fu และคณะ (2024) แสดงให้เห็นว่า NIHSS มีความสามารถในการพยากรณ์โรคต่ำกว่า BI โดยเฉพาะในผู้ป่วยที่ได้รับการรักษาด้วย thrombolysis [25]

แต่ละเครื่องมือมีจุดเด่นที่แตกต่างกัน mRS เน้นการวัดระดับความพิการโดยรวมและสัมพันธ์กับผลลัพธ์ระยะยาว BI วัดความสามารถในการทำกิจวัตรประจำวันและให้การพยากรณ์ที่แม่นยำโดยเฉพาะในกรณีของการรักษาด้วย thrombolytic therapy ขณะที่ NIHSS มีบทบาทในการประเมินความบกพร่องทางระบบประสาท แต่มีความแม่นยำต่ำกว่าในการพยากรณ์ผลลัพธ์ในระยะยาว [26] โรงพยาบาลส่วนใหญ่จึงมักจะใช้ Barthel Index มาประเมินพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมองตั้งแต่นอนโรงพยาบาลจนถึงการติดตามต่อเนื่องในชุมชน

วิธีการศึกษา

Study design, settings, and data source

การศึกษาเชิงเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมองของโรงพยาบาลหาดใหญ่ข้อมูลถูกเก็บรวบรวมย้อนหลังจากฐานข้อมูลผู้ป่วยโรคหลอดเลือดสมองของศูนย์ข้อมูลข่าวสารระบบประสาท สถาบันประสาทวิทยา ระหว่างปี 2561-2566 จากข้อมูลทั้งหมด 6,081 รายการ ข้อมูลส่วนบุคคลที่สามารถระบุตัวตนไม่ได้เป็นส่วนหนึ่งของชุดข้อมูล โดยมุ่งเน้นที่ข้อมูลปัจจัยทางคลินิกและกระบวนการการดูแลรักษาที่ได้รับ ลักษณะของข้อมูล ได้แก่ Categorical Variables เช่น เพศ, สถานะการสูบบุหรี่ Numerical Variables เช่น อายุ, BMI และ Ordinal Variables เช่น Stroke type การประมวลผลและการวิเคราะห์ข้อมูลดำเนินการโดยใช้ภาษาโปรแกรม Python และไลบรารีทางวิทยาศาสตร์ที่เกี่ยวข้อง ได้แก่ Pandas, NumPy และ scikit-learn

Data analysis approach

การศึกษานี้ดำเนินการตามทดลองดังที่แสดงใน รูปที่ 1 เพื่อสร้างแบบจำลองการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง เราได้ประเมินแบบจำลองการเรียนรู้ของเครื่อง (ML) จำนวน 3 แบบ ประกอบด้วย Decision Tree, Random Forest, Gradient Boosting ทั้งนี้ การประเมินประสิทธิภาพของแต่ละโมเดลถูกดำเนินการโดยใช้เทคนิค K-Fold Cross-Validation เพื่อให้ได้ผลลัพธ์ที่แม่นยำและเชื่อถือได้

Data preparation

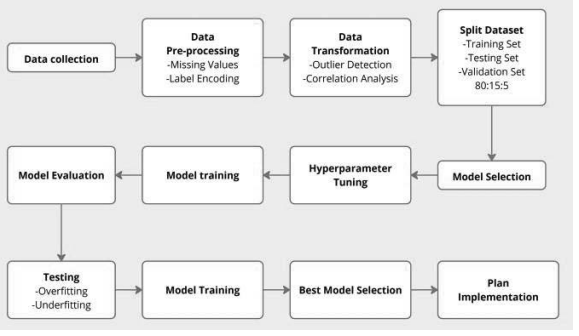
การเตรียมข้อมูลเป็นสิ่งจำเป็นในกระบวนการเรียนรู้ของ ML รวมถึงการเข้ารหัสป้ายกำกับ (label encoding) ข้อมูลบางส่วนถูกตัดออกเพื่อลดความซับซ้อนและเพิ่มความแม่นยำในการจำแนก โดยข้อมูลที่ถูกตัดออก ได้แก่ ผลการประเมิน NIHSS, mRS และ Barthel index ครั้งที่ 2 ก่อน discharge เนื่องจากเป็นผลลัพธ์ในการฟื้นตัว รวมทั้งรายที่ไม่มีข้อมูลการประเมินผลลัพธ์ หลังจากการทำความสะอาดข้อมูล ได้แก่ จัดการกับค่าที่ขาดหาย (Handling Missing Data), จัดการกับค่าผิดปกติ (Outlier Detection and Treatment), แปลงข้อมูลประเภทหมวดหมู่ (Encoding Categorical Variables) ทำให้เหลือข้อมูลที่พร้อมใช้ในการพัฒนาโมเดลจำนวน 6,081 รายการ จากข้อมูลทั้งหมดที่รวบรวมได้ 6,117 รายการ แบ่งตามระดับผลการประเมิน Barthel index ครั้งที่ 2 โดย 100 คะแนน คือ การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง และคะแนน 0-99 คือ การฟื้นตัวอย่างสมบูรณ์ไม่สมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง หลังจากเก็บรวมรวมจึงทำการตรวจสอบข้อมูลค่าผิดปกติ (Outlier Detection) รูปที่ 2 โดยใช้เกณฑ์ในการตัดค่าผิดปกติออกจากข้อมูล เนื่องจากค่าที่ผิดปกติอาจเกิดจากการป้อนข้อมูลที่ไม่ถูกต้องหรือการวัดค่าที่ไม่สมเหตุสมผล ซึ่งอาจทำให้โมเดล Machine Learning เรียนรู้ผิดพลาดและทำให้ผลการจำแนกมีความแม่นยำน้อยลง ในขั้นตอนนี้ ค่าที่สูงกว่าหรืออยู่นอกขอบเขตที่กำหนดไว้จะถูกลบออก ได้แก่ ระดับน้ำตาลในเลือด (BS), ระดับไขมันในเลือด (LDL) อาจเกิดขึ้นได้จากหลายสาเหตุ เช่น ข้อผิดพลาดในการวัด, ความผิดพลาดในการป้อนข้อมูล หลังจากการทำความสะอาดข้อมูล ทำให้เหลือข้อมูลที่พร้อมใช้ในการพัฒนาโมเดลจำนวน 5,058 รายการ แบ่งเป็น 2 ระดับ ตามการฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง ได้ชุดข้อมูล รูปที่ 3 และดำเนินการจัดการความไม่สมดุลของข้อมูล (Class Imbalance) ในชุดข้อมูลเทคนิค Oversampling ด้วย SMOTE (Synthetic Minority Oversampling Technique)

Feature

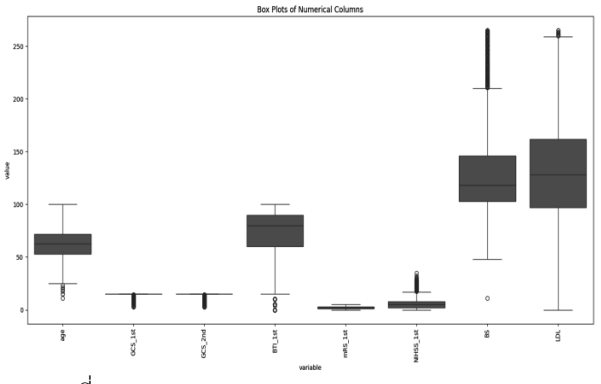
ตัวแปรทางคลินิกที่ใช้ในการทดลองนี้ ประกอบด้วยข้อมูลประชากรของผู้ป่วย (อายุและเพศ), โรคประจำตัว (เบาหวาน, ความดันโลหิตสูง, ไขมันในเลือดสูง, โรคหัวใจสั่นพลิ้ว, โรคหลอดเลือดสมอง) การสูบบุหรี่ดื่มแอลกอฮอล์, การวินิจฉัยโรค, GCS score ครั้งที่ 1 และครั้ง 2, Barthel index ครั้งที่ 1, mRS score ครั้งที่ 1 และ 2, NIHSS ครั้งที่ 1 และ 2, BS, LDL, การได้ยา RTPA, การรักษาด้วย Thrombectomy, การ on ventilator เป็นต้น

Train and Test Data

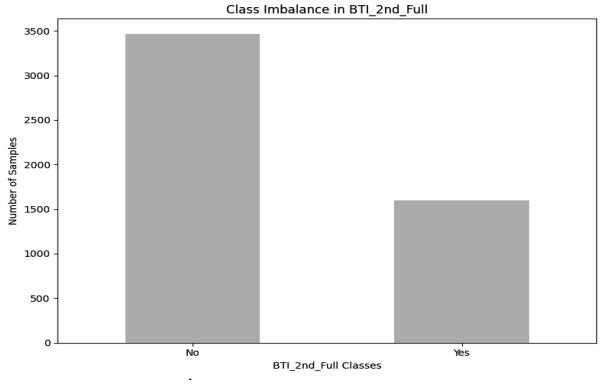
ข้อมูลถูกแบ่งออกเป็น 3 ชุดข้อมูลฝึกอบรม (Training Set) ชุดข้อมูลทดสอบ (Testing Set) และชุดตรวจสอบ (validation set) โดยใช้สัดส่วน 80:15:5 โดยโมเดลที่ถูกทดสอบประกอบด้วย Logistic Regression, Decision Tree, Random Forest, และ Gradient Boosting การประเมินประสิทธิภาพของแต่ละโมเดลถูกดำเนินการโดยใช้เทคนิค K-Fold Cross-Validation แบ่งข้อมูลทดสอบ 5 รอบ ให้แน่ใจว่าโมเดลไม่เกิดการ overfitting เพื่อให้ได้ผลลัพธ์ที่แม่นยำและเชื่อถือได้ จากการเปรียบเทียบพบว่า Random Forest มีประสิทธิภาพโดยรวมดีที่สุด โดยมีค่า Accuracy, Precision, Recall, F1-Score และ Cross-Validation Accuracy สูงสุดใกล้เคียงกับ Gradient Boosting ส่วน Decision Tree มีประสิทธิภาพต่ำสุด ดังตารางที่ 1 และได้มีการปรับแต่งพารามิเตอร์เพิ่มเติม (Hyperparameter Tuning) เช่น จำนวนต้นไม้ในป่า (n estimators) และความลึกสูงสุดของต้นไม้ (max depth) เพื่อเพิ่มประสิทธิภาพของโมเดลในการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง



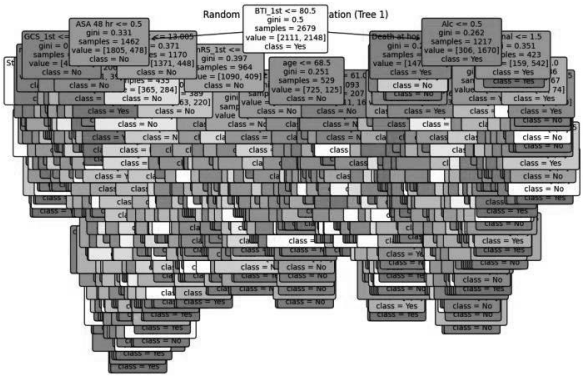
รูปที่ 1 ขั้นตอนที่ใช้ในการออกแบบและดำเนินการทดลอง



รูปที่ 2 Boxplot of Features for Outlier Detection



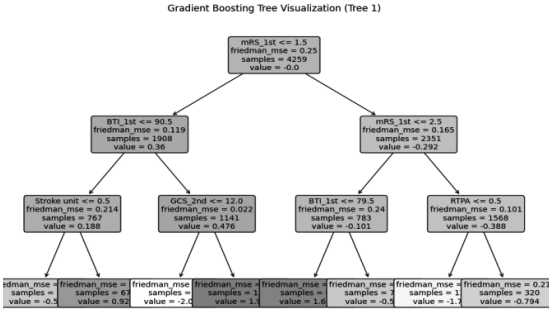
รูปที่ 3 Class Imbalance in Outcomes



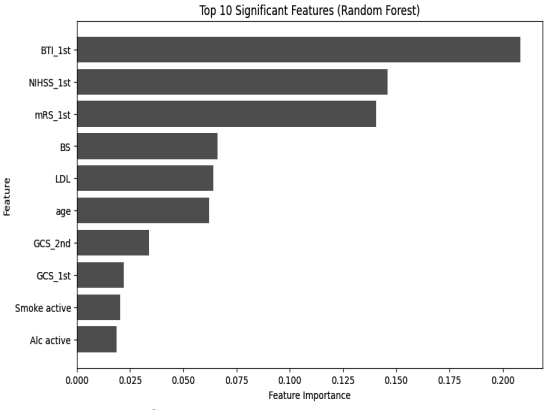
รูปที่ 4 Random Forest Tree Visualization

ตารางที่ 1 ตารางการเปรียบเทียบโมเดลที่ถูกทดสอบ

Class	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.8708	0.8708	0.8708	0.8708
Cross-Validation Accuracy: 0.8669 (+/- 0.0107)				
Decision Tree	0.8383	0.8383	0.8383	0.8383
Cross-Validation Accuracy: 0.8345 (+/- 0.0232)				
Random Forest	0.8861	0.8863	0.8861	0.8861
Cross-Validation Accuracy: 0.8880 (+/- 0.0155)				
Gradient Boosting	0.8850	0.8854	0.8850	0.8850
Cross-Validation Accuracy: 0.8795 (+/- 0.0075)				



รูปที่ 5 Gradient Boosting Tree Visualization



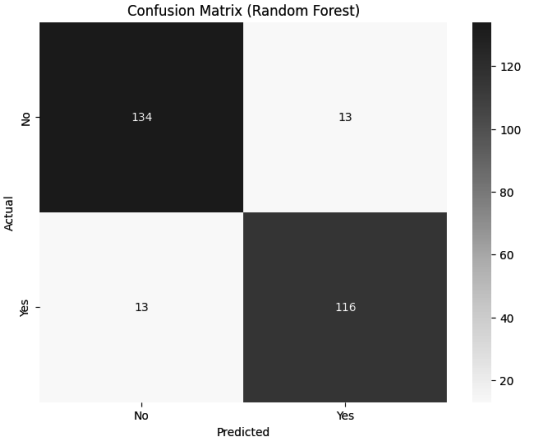
รูปที่ 6 Feature ที่สำคัญใน Random Forest

ผลการศึกษาและการอภิปรายผล

1. ผลการศึกษา

1.1 ประสิทธิภาพของโมเดลที่ได้รับการปรับแต่ง (Tuned Model Performance)

โมเดล Random Forest ที่ได้รับการปรับแต่งแสดงความแม่นยำ (Accuracy) 88.61% บนชุดข้อมูลทดสอบซึ่งแสดงให้เห็นว่าโมเดลสามารถพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมองได้ ผลการทดสอบโมเดลแสดงให้เห็นว่า โมเดลมีค่า Precision, Recall, และ F1-Score ค่อนข้างสูง ตารางที่ 1



รูปที่ 7 Confusion matrix แสดงผลจากประสิทธิภาพของโมเดล

1.2 การวิเคราะห์ Confusion Matrix

โมเดล Random Forest มีประสิทธิภาพที่ดีทั้งในการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง (Yes) และการพยากรณ์การฟื้นตัวอย่างไม่สมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง (No) โดยมีค่า Precision, Recall และ F1-Score อยู่ในช่วง 0.90 ถึง 0.91 และมี ค่า Accuracy ของโมเดลสูงถึง 91% เนื่องจากค่าต่าง ๆ ของคลาส No และ Yes ใกล้เคียงกัน แสดงให้เห็นว่าโมเดลสามารถจัดการกับข้อมูลที่มีความไม่สมดุลได้ดี แต่จำนวนตัวอย่างใน Support ต่ำอาจเกิดจากการแบ่งข้อมูลออกเป็นชุดเล็ก ๆ เพื่อใช้ในการทดสอบ, การทำ Resampling หรือการกรองข้อมูลที่ไม่สมบูรณ์ ซึ่งเป็นสาเหตุหลักที่ทำให้ข้อมูลที่ถูกทดสอบมีจำนวนลดลง **ตารางที่ 2**

ตารางที่ 2 Confusion Matrix ของ Random Forest

Class	Precision	Recall	F1-Score	Support
No	0.91	0.91	0.91	147
Yes	0.90	0.90	0.90	129
accuracy			0.91	276
macro avg	0.91	0.91	0.91	276
weighted avg	0.91	0.91	0.91	276

1.3 การเปรียบเทียบโมเดล (Model Comparison)

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

2. อภิปรายผล

ผลการศึกษานี้แสดงให้เห็นว่าโมเดล Random Forest ที่พัฒนาขึ้นมีความแม่นยำในการพยากรณ์การฟื้นตัวอย่างไม่สมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง ซึ่งสอดคล้องกับงานวิจัยอื่น ๆ ที่แสดงให้เห็นว่าโมเดล Random Forest และ Gradient Boosting มีประสิทธิภาพในการจัดการข้อมูลที่มีความซับซ้อนและไม่สมดุลได้ดีกว่า โมเดลอื่น ๆ เช่น Logistic Regression และ Decision Tree เช่นเดียวกับที่พบในงานวิจัยของ (Wang et al., 2020) ที่แสดงว่า Random Forest มีความสามารถสูงสามารถทำนายผลการฟื้นตัวได้อย่างแม่นยำในกลุ่มผู้ป่วยขนาดใหญ่ [5] [13]

การศึกษานี้ยังชี้ให้เห็นถึงข้อดีของการใช้ Random Forest เมื่อเทียบกับโมเดล Decision Tree ซึ่งได้รับการสนับสนุนจากงานวิจัยโดย (Marron et al., 2014) ที่ระบุว่า Random Forest มักมีประสิทธิภาพสูงกว่าในด้านความแม่นยำและการจัดการกับข้อมูลที่มีหลายมิติ [10] อย่างไรก็ตาม โมเดล Random Forest ที่พัฒนาขึ้นมีความแม่นยำในการพยากรณ์การฟื้นตัวอย่างไม่สมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมองมากกว่าการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง แสดงให้เห็นว่ามีความท้าทายในการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมองในก่อนออกจากโรงพยาบาล ซึ่งอาจสอดคล้องกับข้อค้นพบในงานวิจัยของ (Ungkawa & Rafi, 2024) ที่ระบุว่าความไม่สมดุลของข้อมูลในระดับต่าง ๆ อาจส่งผลต่อประสิทธิภาพของโมเดล [20] จากการวิเคราะห์อาจเนื่องมาจากจำนวน feature ที่มาก ทำให้โมเดลพยากรณ์การฟื้นตัวได้ไม่ค่อยแม่นยำ และจากการพิจารณาพบว่าคุณลักษณะที่จะการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง ขึ้นอยู่กับคุณลักษณะทางการประเมินทางคลินิก เช่น Barthel index ครั้งที่ 1, NIHSS ครั้งที่ 1, mRS ครั้งที่ 1, BS, LDL, อายุ, GCS score เป็นต้น ความสำคัญของผลลัพธ์ที่ได้รับในงานวิจัยนี้อยู่ที่การพิสูจน์ว่าโมเดล Machine Learning สามารถใช้ในการพยากรณ์ผลลัพธ์ของผู้ป่วยโรคหลอดเลือดสมองและวางแผนการรักษาหรือปรับเปลี่ยนพฤติกรรมที่เหมาะสมได้

การทดสอบและเปรียบเทียบโมเดลหลายรูปแบบช่วยยืนยันว่า Random Forest เป็นตัวเลือกที่เหมาะสมที่สุดในการใช้งานจริง การตรวจสอบ Overfitting/Underfitting เป็นส่วนสำคัญของการพัฒนาโมเดล Machine Learning เพื่อให้แน่ใจว่าโมเดลที่พัฒนาไม่เรียนรู้ข้อมูลมากเกินไป (Overfitting) หรือเรียนรู้น้อยเกินไป (Underfitting) การตรวจสอบนี้ถูกดำเนินการโดยการเปรียบเทียบผลการทำงานของโมเดลบนชุดข้อมูลฝึกอบรมและชุดข้อมูลทดสอบ และผลการทดสอบโมเดล Random Forest ที่ถูกปรับแต่ง (Tuned Random Forest) บนชุดข้อมูลทดสอบแสดงให้เห็นว่าโมเดลมีความแม่นยำ (Accuracy) ถึง 88.61% ซึ่งบ่งชี้ว่าโมเดลสามารถพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมองได้อย่างแม่นยำในกรณีส่วนใหญ่

สรุปผลการศึกษา

จากการศึกษาเปรียบเทียบแบบจำลองการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมองโดยใช้ข้อมูลจากโรงพยาบาลหาดใหญ่ พบว่าแบบจำลอง Random Forest มีประสิทธิภาพสูงสุดและใกล้เคียงกับ Gradient Boosting เมื่อเปรียบเทียบกับแบบจำลองอื่น เช่น Logistic Regression และ Decision

Tree โดย Random Forest และ Gradient Boosting ให้ความแม่นยำ (Accuracy) สูงถึง 88.61%, 88.50% ตามลำดับ โดยผลการศึกษานี้พบว่าการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมองด้วยโมเดล Random Forest ขึ้นอยู่กับคุณลักษณะทางคลินิกสำคัญ ได้แก่ Barthel index ครั้งที่ 1, NIHSS ครั้งที่ 1, mRS ครั้งที่ 1, BS, LDL, อายุ, GCS score ข้อมูลเหล่านี้เป็นปัจจัยสำคัญที่สามารถช่วยให้โมเดล Machine Learning สามารถเรียนรู้ได้ดีและมีประสิทธิภาพในการทำนายการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง

แม้ว่าโมเดล Random Forest มีประสิทธิภาพที่ดีทั้งในการพยากรณ์การฟื้นตัวอย่างสมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง (Yes) และการพยากรณ์การฟื้นตัวอย่างไม่สมบูรณ์ของผู้ป่วยโรคหลอดเลือดสมอง (No) แต่จำนวนตัวอย่างใน Support ต่ำอาจเกิดจากการแบ่งข้อมูลออกเป็นชุดเล็ก ๆ เพื่อใช้ในการทดสอบ, การทำ Resampling หรือการกรองข้อมูลที่ไม่สมบูรณ์ ซึ่งเป็นสาเหตุหลักที่ทำให้ข้อมูลที่ถูกทดสอบมีจำนวนลดลงในภาพรวม แบบจำลอง Random Forest ได้รับการยอมรับว่าเป็นโมเดลที่เหมาะสมในการพยากรณ์การฟื้นตัวของผู้ป่วยโรคหลอดเลือดสมอง โดยผลการวิจัยนี้สามารถนำไปประยุกต์ใช้ในการวางแผนการรักษาหรือปรับเปลี่ยนพฤติกรรมที่เหมาะสมได้รวมทั้งฟื้นฟูผู้ป่วยได้อย่างมีประสิทธิภาพ

REFERENCES

[1] V. L. Feigin et al., “Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019,” The Lancet Neurology, vol. 20, no. 10, pp. 795–820, Oct. 2021, doi: 10.1016/S1474-4422(21)00252-0.

[2] A. Sardar, K. Shahzad, A. R. Arshad, K. Shabbir, and S. Raza, “Correlation of caregivers’ strain with patients’ disability in stroke,” vol. 34, pp. 326–30, Mar. 2022, doi: 10.55519/JAMC-02-9488.

[3] Q. Wu et al., “Comparison of Three Instruments for Activity Disability in Acute Ischemic Stroke Survivors,” Can. J. Neurol. Sci., vol. 48, no. 1, pp. 94–104, Jan. 2021, doi: 10.1017/cjn.2020.149.

[4] M. Murie-Fernแndez and M. M. Marzo, “Predictors of Neurological and Functional Recovery in Patients with Moderate to Severe Ischemic Stroke: The EPICA Study,” Stroke Research and Treatment, vol. 2020, no. 1, p. 1419720, 2020, doi: 10.1155/2020/1419720.

[5] W. Wang et al., “A systematic review of machine learning models for predicting outcomes of stroke with structured data,” PLoS ONE, vol. 15, no. 6, p. e0234722, Jun. 2020, doi: 10.1371/journal.pone.0234722.

[6] Q. Zhang, Z. Zhang, X. Huang, C. Zhou, and J. Xu, “Application of Logistic Regression and Decision Tree Models in the Prediction of Activities of Daily Living in Patients with Stroke,” Neural Plasticity, vol. 2022, no. 1, p. 9662630, 2022, doi: 10.1155/2022/9662630.

[7] A. Criminisi, J. Shotton, and E. Konukoglu, “Decision Forests for Classication, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning”.

[8] A. Cuzzocrea, S. Francis, and M. Gaber, An Information-Theoretic Approach for Setting the Optimal Number of Decision Trees in Random Forests. 2013, p. 1019.

[9] Tin Kam Ho, “Random decision forests,” in Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, Que., Canada: IEEE Comput. Soc. Press, 1995, pp. 278–282. doi: 10.1109/ICDAR.1995.598994.

[10] D. Marron, A. Bifet, and G. De Francisci Morales, “Random Forests of Very Fast Decision Trees on GPU for Mining Evolving Big Data Streams,” in ECAI 2014, IOS Press, 2014, pp. 615–620. doi: 10.3233/978-1-61499-419-0-615.

[11] A. V. Konstantinov and L. V. Utkin, “Gradient boosting machine with partially randomized decision trees,” Jun. 19, 2020, arXiv: arXiv:2006.11014. Accessed: Oct. 19, 2024. [Online]. Available: http://arxiv.org/abs/2006.11014

[12] G. C. Okoye and E. U. Umeh, “Predicting Functional Outcome After Ischemic Stroke Using Logistic Regression and Machine Learning Models,” Earthline Journal of Mathematical Sciences, vol. 14, no. 1, pp. 133–150, 2024.

[13] D. Sengupta, S. Mondal, Y. R. Singh, and A. Pandey, “Performance Analysis of Machine Learning Algorithms for Prediction of Cerebral Attack (Stroke),” in Frontiers of ICT in Healthcare, vol. 519, J. K. Mandal and D. De,

Eds., in Lecture Notes in Networks and Systems, vol. 519. , Singapore: Springer Nature Singapore, 2023, pp. 215–228. doi: 10.1007/978-981-19-5191-6_18.

[14] S.-C. Chang et al., "The comparison and interpretation of machine-learning models in post-stroke functional outcome prediction," Diagnostics, vol. 11, no. 10, p. 1784, 2021.

[15] "Building decision trees for the multi-class imbalance problem," in SciSpace - Paper, Springer, Berlin, Heidelberg, May 2012, pp. 122–134. doi: 10.1007/978-3-642-30217-6_11.

[16] "A Study with Class Imbalance and Random Sampling for a Decision Tree Learning System," in SciSpace - Paper, Springer, Boston, MA, Sep. 2008, pp. 131–140. doi: 10.1007/978-0-387-09695-7_13.

[17] D. J. Dittman, T. M. Khoshgoftaar, and A. Napolitano, "The Effect of Data Sampling When Using Random Forest on Imbalanced Bioinformatics Data," in 2015 IEEE International Conference on Information Reuse and Integration, Aug. 2015, pp. 457–463. doi: 10.1109/IRI.2015.76.

[18] "An Improved Random Forest Algorithm for Class-Imbalanced Data Classification and its Application in PAD Risk Factors Analysis," The Open Electrical & Electronic Engineering Journal, vol. 7, no. 1, pp. 62–70, Jun. 2013, doi: 10.2174/1874129001307010062.

[19] M. Amrehn, F. Mualla, E. Angelopoulou, S. Steidl, and A. Maier, "The Random Forest Classifier in WEKA: Discussion and New Developments for Imbalanced Data," Jan. 04, 2019, arXiv: arXiv: 1812.08102. Accessed: Oct. 02, 2024. [Online]. Available: http://arxiv.org/abs/1812.08102

[20] "Data Balancing Techniques Using the PCA-KMeans and ADASYN for Possible Stroke Disease Cases | Jurnal Online Informatika." Accessed: Oct. 02, 2024. [Online]. Available: https://join.if.uinsgd.ac.id/index.php/join/article/view/1293

[21] S. Campagnini, C. Arienti, M. Patrini, P. Liuzzi, A. Mannini, and M. C. Carrozza, "Machine learning methods for functional recovery prediction and prognosis in post-stroke rehabilitation: a systematic review," J NeuroEngineering Rehabil, vol. 19, no. 1, p. 54, Jun. 2022, doi: 10.1186/s12984-022-01032-4.

[22] S.-C. Chang et al., "The Comparison and Interpretation of Machine-Learning Models in Post-Stroke Functional Outcome Prediction," Diagnostics, vol. 11, no. 10, Art. no. 10, Oct. 2021, doi: 10.3390/diagnostics11101784.

[23] "Inpatient stroke rehabilitation: prediction of clinical outcomes using a machine-learning approach | Journal of NeuroEngineering and Rehabilitation." Accessed: Oct. 02, 2024. [Online]. Available: https://link.springer.com/article/10.1186/s12984-020-00704-3

[24] "Abstract TP77: Stroke Rank Order Correlation but Moderate Absolute Value Drift Between Early Day 2/4 and Day 90 Stroke Outcome Scales: mRS, BI, and NIHSS," Stroke, Feb. 2024, doi: 10.1161/str.55.suppl_1.tp77.

[25] M. Fu et al., "Barthel Index, SPAN-100, and NIHSS Studies on the Predictive Value of Prognosis in Patients With Thrombolysis," The Neurologist, vol. 29, no. 3, p. 158, May 2024, doi: 10.1097NRL.0000000000000554.

[26] K. Ghandehari, "Challenging comparison of stroke scales," J Res Med Sci, vol. 18, no. 10, pp.906–910, Oct. 2013.