การทำนายการเข้ารับบริการซ้ำภายใน 72 ชั่วโมงของผู้ป่วยที่แผนกฉุกเฉิน โดยใช้การเรียนรู้ของเครื่อง

พิชญ์พงศ์ พูลผล[1*], ศิริสรรพ เหล่าหะเกียรติ[2]

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อประยุกต์ใช้การเรียนรู้ของเครื่องมาช่วยทำนายการกลับมารับบริการซ้ำที่แผนกฉุกเฉินภายในระยะเวลา 72 ชั่วโมง เพื่อสนับสนุนแพทย์ประจำแผนกฉุกเฉินในการประเมินผู้ป่วยหลังได้รับอนุญาตให้สามารถกลับบ้านได้ โดยใช้ข้อมูลชุด MIMIC-IV-ED ซึ่งเป็นข้อมูลจากฐานข้อมูลเวชระเบียนของแผนกฉุกเฉินที่ Beth Israel Deaconess Medical Center เมืองบอสตัน ประเทศสหรัฐอเมริการะหว่างปี ค.ศ. 2011 – ค.ศ. 2019 จากเว็บไซต์ Physionet ซึ่งมีจำนวนการเข้ารับบริการที่แผนกฉุกเฉินทั้งหมด 220,378 ครั้ง โดยมีการกลับมารับบริการซ้ำที่แผนกฉุกเฉินภายใน 72 ชั่วโมง 10,090 ครั้ง (4.58%) ใช้ตัวแปรทั้งหมด 22 ตัวแปร โดยฝึกฝนด้วยข้อมูลทั้งแบบที่มีความไม่สมดุลของข้อมูล และข้อมูลที่ใช้วิธีการจัดการความไม่สมดุลของข้อมูล 4 วิธี คือ Random oversampling, SMOTE, Random undersampling และ Class weight เพื่อนำมาเปรียบเทียบประสิทธิภาพของแบบจำลอง 4 ประเภทหลัก ได้แก่ Logistic regression, KNN classifier, Random forest classifier และ XGBoost classifer พบว่า ข้อมูลที่ไม่มีการจัดการความไม่สมดุลส่งผลให้แบบจำลองมีค่า Accuracy สูงถึง 0.95 แต่ค่า Recall เท่ากับ 0 ซึ่งมี AUC เท่ากับ 0.5 ไม่แตกต่างจากการทำนายแบบสุ่ม แต่เมื่อมีการจัดการความไม่สมดุลของข้อมูล แต่ละวิธีของการจัดการความไม่สมดุลของข้อมูลได้ผลดีกว่าแต่ไม่แตกต่างกัน โดยแบบจำลองที่ใช้วิธี Logistic regression ได้ค่า AUC สูงที่สุด คือ เท่ากับ 0.61 ทั้งนี้จากการสร้างแบบจำลองพบว่า ตัวแปรที่มีความสำคัญได้แก่ การวินิจฉัย เพศ อายุ เชื้อชาติ ระดับความปวด ระยะเวลาการอยู่ในแผนกฉุกเฉิน อัตราการเต้นของหัวใจ

คำสำคัญ : แผนกฉุกเฉิน, การกลับเข้ามารับบริการซ้ำ, 72 ชั่วฌมง

---

[1] หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาการข้อมูล คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

[2] คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ กรุงเทพฯ 10110

* Corresponding author: E-mail address: peachapong.p@gmail.com

Prediction of 72-hour Re-visit to the Emergency Department Using Machine Learning

Peachapong Poolpol[1*], Sirisup Laohakiat[2]

Abstract

This study aimed to develop machine learning to predict re-visit to Emergency Department (ED) within 72 hours of patients who are discharged from the ED, supporting emergency physicians in evaluating patient risks for re-visit after discharge. Utilizing the MIMIC-IV-ED dataset from Beth Israel Deaconess Medical Center in Boston, Massachusetts, USA, spanning 2011 to 2019. This dataset, available on Physionet, contains 220,378 emergency department visits with 10,090 (4.58%) resulting in a revisit within 72 hours with a total of 22 variables. The study investigates techniques for handling data imbalance, including Random Oversampling, SMOTE, Random Undersampling, and Class Weight, to train various models: Logistic Regression, KNN Classifier, Random Forest Classifier, and XGBoost Classifier. The findings indicate that addressing data imbalance significantly affects model learning, with accuracies reaching up to 0.95 but recall at zero, resulting in an AUC of 0.5, which is equivalent to random prediction. However, the performance of the imbalance management methods did not show significant differences, with Logistic Regression models achieving the highest AUC at 0.61. Key predictive variables identified include diagnosis, gender, age, race, pain scores, length of stay in the ED, and heart rate.

Keywords : Emergency Department, Re-visit, 72-hour

---

[1] Data Science, Faculty of Science, Srinakharinwirot University, Bangkok, 10110, Thailand

[2] Faculty of Science, Srinakharinwirot University, Bangkok, 10110, Thailand

* Corresponding author: E-mail address: peachapong.p@gmail.com

บทนำ

แผนกฉุกเฉินเป็นแผนกที่เปิดให้บริการตลอด 24 ชั่วโมงในทุกวัน เพื่อให้บริการกับผู้ป่วยที่มีความต้องการเร่งด่วนในการรักษา และเป็นแผนกที่มีผู้ป่วยเข้ารับบริการจำนวนมาก ในขณะที่ทรัพยากรมีอย่างจำกัด แต่ทั้งนี้มีผู้ป่วยเพียง 21.4% - 53.0% ที่มารับบริการที่แผนกฉุกเฉินแล้วได้รับการรับตัวไว้รักษาในโรงพยาบาล[1, 2] ในขณะที่ผู้ป่วยที่เหลือ จะได้รับการวินิจฉัยให้การรักษา และอนุญาตให้ผู้ป่วยสามารถกลับบ้านได้ โดยจะต้องมีการประเมินจากแพทย์ที่ประจำแผนกฉุกเฉินแล้วว่าผู้ป่วยมีอาการคงที่ในระดับที่สามารถกลับไปรับการรักษาต่อที่บ้านได้ ทั้งนี้มีผู้ป่วย 9% - 48% ที่เมื่อได้รับการอนุญาตให้สามารถกลับบ้านได้ แต่กลับมีอาการที่รุนแรงขึ้น และต้องกลับเข้ามารับการบริการซ้ำที่แผนกฉุกเฉิน[3] โดยเฉพาะภายในระยะเวลา 72 ชั่วโมงหลังจากผู้ป่วยได้รับอนุญาตให้กลับบ้าน ที่อาจจะบ่งบอกว่าการประเมินของแพทย์ประจำแผนฉุกเฉินนั้นอาจเกิดความคลาดเคลื่อน ผู้ป่วยอาจมีอาการที่ยังไม่คงที่เพียงพอ การวิจัยในครั้งนี้จัดทำขึ้นเพื่อศึกษาการประยุกต์ใช้เทคโนโลยีเรียนรู้ของเครื่อง (machine learning) เพื่อทำนายการกลับมารับบริการที่แผนกฉุกเฉินภายในระยะเวลา 72 ชั่วโมง และค้นหาปัจจัยสำคัญที่เกี่ยวข้อง เพื่อเพิ่มประสิทธิภาพและประสิทธิผลของบริการด้านสุขภาพ

การศึกษาของ Hsu และคณะ[4] ทำการศึกษาเพื่อทำนายการกลับมารับรับบริการซ้ำที่แผนกฉุกเฉินโดยไม่ได้นัดหมายภายใน 72 ชั่วโมงในผู้ป่วยที่มีอาการปวดท้องโดยใช้การเรียนรู้ของ โดยมีการเข้ารับบริการที่แผนกฉุกเฉินด้วยอาการปวดท้อง 25,151 ครั้ง และมีการกลับมารับบริการซ้ำภายใน 72 ชั่วโมง จำนวน 1,471 ครั้ง คิดเป็น 5.85% ซึ่งมีปัจจัยที่นำมาวิเคราะห์ทั้งหมด 617 ปัจจัย เช่น ข้อมูลทั่วไปของผู้ป่วย วิธีการมาโรงพยาบาล ช่วงเวลาที่ถึงโรงพยาบาล ความอาวุโสของแพทย์ ระยะเวลาการนอนในแผนกฉุกเฉิน การวินิจฉัย ความถี่ในการมาที่แผนกฉุกเฉินในปีก่อน ความถี่ในการถูกรับรักษาตัวในโรงพยาบาลในปีก่อนหรือการวินิจฉัยในสองปีก่อน โดยอัตราส่วนระหว่างผู้ป่วยที่มารับการบริการซ้ำภายใน 72 ชั่วโมง และผู้ป่วยที่ไม่ได้มารับบริการซ้ำภายใน 72 ชั่วโมง คือ 19:1 จึงใช้ one-sided selection algorithm ช่วยในการปรับอัตราส่วนเป็น 15:1 และนำข้อมูลที่อยู่ในบริเวณ Decision boundaries ออกโดยวิธี Tomek link algorithm (k=1) โดยมีแบบจำลองที่นำมาใช้ คือ logistic regression (LR), random forest (RF), extreme gradient boosting (XGB), and voting classifiers (VC) โดยที่แบบจำลอง voting classifiers จะใช้การทำนายสุดท้ายจากแบบจำลองสามชนิดก่อนหน้า และทำการฝึก ด้วย 10-fold cross-validation และจัดการความไม่เท่ากันของข้อมูลด้วย bootstrapping method เฉพาะข้อมูลที่นำมาฝึก จากการทำ voting weight ของ VC พบว่าน้ำหนักของ 1(LR):2(RF):1(XGB) ให้น้ำหนักที่ดีที่สุด ผลประสิทธิภาพของแบบจำลองพบว่า ในทุกแบบจำลองให้ผลของค่า AUC มากกว่า 0.69 โดยที่ XGB และ voting classifiers มีค่า AUC ดีที่สุดคือ 0.74 แต่แบบจำลอง VC ให้ค่า F1 score ที่สูงที่สุดคือ 0.25 ซึ่งมากกว่าแบบจำลอง XGB ที่มีค่าเท่ากับ 0.07

การศึกษาของ Gao และคณะ[5] ทำการศึกษาเพื่อทำการทำนายการกลับมาห้องฉุกเฉินซ้ำใน 30 วัน ของปีงบประมาณ 2014 โดยมีผู้ป่วย 22,734 คนในการวิเคราะห์ และมี 4,937 คนที่กลับมาซ้ำภายใน 30 วัน คิดเป็น 22% โดยใช้ปัจจัยที่นำมาวิเคราะห์ได้แก่ ข้อมูลทั่วไปของผู้ป่วยเช่น อายุ เพศ สถานภาพสมรส เชื้อชาติ ระดับความพิการ ระยะเวลาที่อยู่ในหน่วยบริการทหาร สถานะทางเศรษฐกิจและสังคม ค่าใช้จ่ายจากการเข้าพักในโรงพยาบาลในปีก่อนหน้า และโรคประจำตัวหรือปัจจัยเสี่ยงของผู้ป่วย โดยใช้แบบจำลอง logistic regression ในการวิเคราะห์และวัดผลด้วยค่า C statistic ซึ่งคล้ายกับค่า AUC พบว่าแบบจำลองที่ 1 โดยใช้ข้อมูลเฉพาะ ข้อมูลทั่วไปของผู้ป่วย และสถานะทางเศรษฐกิจและสังคม มีค่า C-statistic 0.556 แบบจำลองที่ 2 ใช้ข้อมูลจากแบบจำลองที่ 1 และค่าใช้จ่ายจากการเข้าพักในโรงพยาบาลในปีก่อนหน้า มีค่า C-statistic เท่ากับ 0.748 และแบบจำลองที่ 3 ใช้ข้อมูลจากแบบจำลองที่ 2 และโรคประจำตัวหรือปัจจัยเสี่ยงของผู้ป่วย มีค่า C-statistic เท่ากับ 0.763

วิธีดำเนินการ

ขั้นตอนที่ 1 : แนะนำชุดข้อมูลที่ใช้ในการศึกษานี้

ผู้วิจัยนำข้อมูลผู้ป่วยที่เข้ารับบริการในแผนกฉุกเฉิน ชื่อ MIMIC-IV-ED ซึ่งเป็นข้อมูลจากฐานข้อมูลเวชระเบียนของแผนกฉุกเฉินที่ Beth Israel Deaconess Medical Center เมืองบอสตัน รัฐแมสซาชูเซตส์ ประเทศสหรัฐอเมริการะหว่างปี ค.ศ. 2011 – ค.ศ. 2019 จากเว็บไซต์ https://physionet.org/content/mimic-iv-ed/2.2/ โดยการได้มาซึ่งข้อมูลต้องเป็นผู้ใช้งานที่ผ่านการอบรมการวิจัยในคนของ CITI Data or Specimens Only Research และต้องเซ็นยินยอมใช้ข้อมูล sign the data use agreement for the project ก่อนจะสามารถเข้าถึงข้อมูลได้[6, 7] ใช้เกณฑ์นำเข้า ได้แก่ อายุตั้งแต่ 18 ปีขึ้นไป และเป็นผู้ที่มีประวัติการได้รับอนุญาตให้กลับบ้าน เมื่อมารับบริการที่แผนกฉุกเฉิน และเกณฑ์การคัดออก ได้แก่ เป็นผู้ที่ไม่ได้รับการรักษา จนถึงขั้นตอนที่แพทย์วินิจฉัยและตัดสินใจในการจำหน่ายผู้ป่วย เช่น ไม่รอตรวจ หลบหนีระหว่างการรักษา ปฏิเสธการรักษา หรือเป็นผู้ที่มีประวัติได้รับอนุญาตให้รับการรักษาในโรงพยาบาล หรือส่งต่อไปยังสถานพยาบาลอื่น หรือเป็นผู้ที่เสียชีวิตภายในแผนกฉุกเฉิน

ขั้นตอนที่ 2 : การนำเข้าข้อมูล ตรวจสอบข้อมูล และพิจารณาข้อมูล

ในการวิจัยนี้จะมีการนำเข้าไลบรารี่สำหรับการใช้งานภาษาไพทอน (python) โดยไลบรารี่หลักที่ใช้ในการจัดการข้อมูล วิเคราะห์ข้อมูล และสร้างแบบจำลองได้แก่ numpy, pandas, datetime, matplotlib, sklearn, imblearn, xgboost จากนั้นจะนำเข้าข้อมูลในรูปแบบไฟล์ Comma Separated Value (*.csv) ด้วย pandas

ขั้นตอนที่ 3 : การวิเคราะห์ข้อความเพื่อสร้างกลุ่มสินค้า

การจัดเตรียมข้อมูล ผู้วิจัยได้ตัดข้อมูลบางแถวออกไม่นำข้อมูลต่อไปนี้มาร่วมวิเคราะห์ ได้แก่ ใบคำสั่งซื้อมีการยกเลิกหรือแก้ไข, ข้อมูลที่ไม่มีรหัสลูกค้า และทำการกำหนดช่วงวันที่ของข้อมูลที่ใช้ในการวิเคราะห์ข้อมูลเท่ากับ 1 ปี ต่อมาวิเคราะห์ข้อความจากชื่อสินค้า ใช้ข้อมูลคอลัมน์ชื่อสินค้า เลือกเฉพาะสินค้าที่ไม่ซ้ำโดยมีรหัสสินค้าเป็นตัวบ่งบอก เนื่องจากชื่อสินค้าอยู่ในรูปแบบภาษาปกติของมนุษย์จึงจำเป็นต้องแปลงชื่อสินให้อยู่ในรูปแบบที่คอมพิวเตอร์สามารถเข้าใจ โดยผ่านกระบวนการวิเคราะห์ข้อความเริ่มจากนำข้อมูลจากคอลัมน์สินค้ามาตัดคำชื่อสินค้าภาษาอังกฤษและกำจัดคำหยุด โดยอาศัยพจนานุกรมคำศัพท์มาช่วยในกระบวนการตัดคำ สำหรับงานวิจัยนี้ผู้วิจัยใช้เครื่องมือประมวลภาษาธรรมชาติ โมดูล NLTK (Natural Language Toolkit) จากนั้นหาคำหลักที่พบบ่อยที่สุด สำหรับนำไปใช้กำหนดประเภทให้กับชื่อสินค้า ผู้วิจัยใช้เทคนิค K-Means Clustering ในการแบ่งกลุ่ม เพื่อกำหนดประเภทให้กับชื่อสินค้า ใช้ Silhouette Score เป็นเทคนิคที่ใช้วัดตัวอย่างนั้นๆมีความเหมือนกับกลุ่มที่อยู่มากเพียงใด เมื่อเทียบกับกลุ่มอื่นๆ และทำการรวมข้อมูลประเภทสินค้าโดยจัดกลุ่มตามแอตทริบิวต์รหัสลูกค้า

ขั้นตอนที่ 4 : การเตรียมข้อมูล

ผู้วิจัยทำการเตรียมข้อมูล ได้แก่ การนำเข้าข้อมูลจากตารางในฐานข้อมูล MIMIC-IV-ED และ MIMIC-IV จากนั้นทำการสำรวจข้อมูล ทำความสะอาดข้อมูล คัดออกข้อมูลที่ซ้ำ มีการสร้างตัวแปรใหม่ ได้แก่ อายุ ระยะเวลาที่ผู้ป่วยอยู่ในแผนกฉุกเฉิน รวมทั้งตัดข้อมูลตามเกณฑ์คัดเข้าและคัดออก โดยมีจำนวนการเข้ารับบริการที่แผนกฉุกเฉิน 220,378 ครั้ง เป็นการกลับมารับบริการ

ซ้ำภายใน 72 ชั่วโมง จำนวน 10,090 ครั้ง คิดเป็น 4.58% การแปลงข้อมูลแบบกลุ่ม ได้แก่ เพศ เชื้อชาติ และ การเดินทางมาที่แผนกฉุกเฉิน สำหรับข้อมูลเพศใช้การแปลงค่า เพศหญิง เป็น 0 และเพศชายเป็น 1 ในขณะที่เชื้อชาติ และการเดินทางมาที่แผนกฉุกเฉิน ซึ่งมีมากกว่า 1 กลุ่ม จึงใช้ OneHotEncoder ในการแปลงข้อมูลดังกล่าวให้เป็นตัวเลข เนื่องจากข้อมูลการวินิจฉัยโรค มีทั้งหมด 9 คอลัมน์ และข้อมูลภายในเป็นข้อมูลที่แตกต่างกันค่อนข้างมาก จึงต้องใช้ TargetEncoder มาช่วยในการแปลงข้อมูลถ่วงน้ำหนักเป็นตัวเลข โดยอ้างอิงจากผลของ Label จึงต้องป้องกันไม่ให้เกิด Data leakage โดยแบ่งข้อมูลเป็น Training set และ Test set ก่อนจากนั้นจึงทำ Target Encoder เฉพาะข้อมูล Training set เท่านั้น

ขั้นตอนที่ 5 : การสเกลข้อมูล และการจัดการความไม่สมดุลของข้อมูล

สำหรับข้อมูลที่เป็นตัวเลขทั้งหมดในงานวิจัยนี้จะใช้การสเกลข้อมูลด้วยวิธี StandardScaler สำหรับข้อมูล Training set เพื่อใช้ในการพิจารณา feature importance เนื่องจากหากข้อมูลตัวเลขมี scale ที่ไม่เท่ากัน จะทำให้ค่าความสำคัญของตัวแปรถูกคิดถ่วงไปที่ข้อมูลที่มีค่ามาก และมีการจัดการความไม่สมดุลของข้อมูล โดยในงานวิจัยฉบับนี้จะใช้วิธีจัดการ คือ Random oversampling, SMOTE, Random undersampling และ Class weight ในทุกแบบจำลอง เปรียบเทียบกับการไม่จัดการความไม่สมดุลของข้อมูล

ขั้นตอนที่ 6 : การสร้างแบบจำลอง

ในการศึกษานี้ ผู้วิจัยสร้างแบบจำลองสำหรับการกลับมารับบริการซ้ำที่แผนกฉุกเฉินภายใน 72 ชั่วโมง โดยเลือกใช้แบบจำลอง Logistic regression, KNN classification, Random Forest classification, XGBoost classification

ผลการวิจัยและอภิปรายผลการวิจัย

ในงานวิจัยนี้มีการสร้างแบบจำลองทั้งสิ้น 29 แบบ โดยแบ่งเป็น 4 แบบจำลองหลัก จากการนำข้อมูลเข้าฝึก และนำมาใช้กับกลุ่มทดสอบได้ประสิทธิภาพของแบบจำลองต่าง ๆ ดังตารางที่ 1

ตารางที่ 1 ผลประสิทธิภาพของแบบจำลองต่าง ๆ

แบบจำลอง	accuracy	roc_auc score	recall score
1. Logistic regression			
1.1. Logistic regression ที่ไม่มีการจัดการความไม่สมดุลของข้อมูล และกำหนดค่า C เป็นค่าเริ่มต้น (C = 1)	0.95	0.5	0
1.2. Logistic regression ที่ไม่มีการจัดการความไม่สมดุลของข้อมูล และกำหนดค่า C เป็น 0.01	0.95	0.5	0
1.3. Logistic regression ที่ไม่มีการจัดการความไม่สมดุลของข้อมูล และกำหนดค่า C เป็น 100	0.95	0.5	0
1.4. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Oversampling และกำหนดค่า C เป็นค่าเริ่มต้น (C = 1)	0.66	0.61	0.55
1.5. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Oversampling และกำหนดค่า C เป็น 0.01	0.66	0.61	0.55

แบบจำลอง	accuracy	roc_auc score	recall score
1.6. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Oversampling และกำหนดค่า C เป็น 100	0.66	0.61	0.55
1.7. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี SMOTE และกำหนดค่า C เป็นค่าเริ่มต้น (C = 1)	0.66	0.61	0.55
1.8. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี SMOTE และกำหนดค่า C เป็น 0.01	0.66	0.61	0.55
1.9. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี SMOTE และกำหนดค่า C เป็น 100	0.66	0.61	0.55
1.10. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Undersampling และกำหนดค่า C เป็นค่าเริ่มต้น (C = 1)	0.66	0.61	0.55
1.11. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Undersampling และกำหนดค่า C เป็น 0.01	0.66	0.61	0.55
1.12. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Undersampling และกำหนดค่า C เป็น 100	0.68	0.61	0.53
1.13. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยการกำหนด class weight ให้ y สมดุล และ C เป็นค่าเริ่มต้น (C = 1)	0.66	0.61	0.56
1.14. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยการกำหนด class weight ให้ y สมดุล และ C เป็น 0.01	0.66	0.61	0.56
1.15. Logistic regression ที่จัดการความไม่สมดุลของข้อมูลด้วยการกำหนด class weight ให้ y สมดุล และ C เป็น 100	0.66	0.61	0.56
2. KNN Classification			
2.1. KNN Classification ที่ไม่มีการจัดการความไม่สมดุลของข้อมูล	0.95	0.5	0.01
2.2. KNN Classification ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Oversampling	0.85	0.53	0.18
2.3. KNN Classification ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี SMOTE	0.76	0.54	0.31
2.4. KNN Classification ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Undersampling	0.6	0.56	0.53
3. Random forest classifier			
3.1. Random forest classifier ที่ไม่มีการจัดการความไม่สมดุลของข้อมูล และกำหนด max_depth = 2	0.95	0.5	0
3.2. Random forest classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Oversampling และกำหนด max_depth = 2	0.72	0.59	0.45
3.3. Random forest classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี SMOTE และกำหนด max_depth = 2	0.95	0.5	0
3.4. Random forest classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Undersampling และกำหนด max_depth = 2	0.69	0.6	0.49
3.5. Random forest classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยการกำหนด class weight ให้ y สมดุล และ max_depth = 2	0.71	0.6	0.48
4. XGBoost classifier			
4.1. XGBoost classifier ที่ไม่มีการจัดการความไม่สมดุลของข้อมูล	0.95	0.5	0
4.2. XGBoost classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Oversampling	0.92	0.54	0.11
4.3. XGBoost classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี SMOTE	0.94	0.51	0.03

แบบจำลอง	accuracy	roc_auc score	recall score
4.4. XGBoost classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยวิธี Random Undersampling	0.23	0.54	0.88
4.5. XGBoost classifier ที่จัดการความไม่สมดุลของข้อมูลด้วยการกำหนด scale_pos_weight ให้ y สมดุล	0.95	0.5	0

เมื่อพิจารณาประสิทธิภาพของแบบจำลองในการทำนายการเข้ารับบริการซ้ำภายใน 72 ชั่วโมงของผู้ป่วยที่แผนกฉุกเฉิน Accuracy สูงที่สุดคือ เท่ากับ 0.95 รองลงมาคือเท่ากับ 0.94 และ 0.92 และน้อยที่สุด เท่ากับ 0.23 Recall สูงที่สุด เท่ากับ 0.88 รองลงมา มีค่าเท่ากับ 0.56 และต่ำที่สุด เท่ากับ 0 AUC สูงที่สุด เท่ากับ 0.61 รองลงมา เท่ากับ 0.60 และน้อยที่สุด เท่ากับ 0.5 จะเห็นว่าทุกแบบจำลองที่ไม่มีการจัดการความสมดุลของข้อมูล มีค่า Accuracy สูงมาก คือเท่ากับ 0.95 แต่ในขณะที่ค่า Recall ซึ่งคำนวณจาก การทำนาย True Positive มีค่าเป็น 0 นั่นหมายถึงแบบจำลองนั้นทำนายถูกเฉพาะค่า Negative ในขณะที่ค่า Positive ที่มีปริมาณน้อยนั้น ถูกทำนายเป็น Negative ทั้งหมด ดังนั้นข้อมูลที่มีความไม่สมดุลของข้อมูลทำนาย จึงไม่เหมาะสมที่จะพิจารณาประสิทธิภาพของแบบจำลองจากค่า Accuracy เพียงอย่างเดียว เพราะเมื่อพิจารณาค่า Recall แล้วมีค่าต่ำมาก เนื่องจากแบบจำลองไม่สามารถทำนายข้อมูลที่เป็น Positive ได้ถูก แต่กลับทำนายข้อมูลที่เป็น Negative ได้ถูกต้อง ดังนั้นจึงต้องมีการจัดการของความไม่สมดุลของข้อมูล โดยในการวิจัยนี้เปรียบเทียบประสิทธิภาพของแบบจำลอง จากการใช้วิธีการจัดการความไม่สมดุลของข้อมูลที่แตกต่างกัน ได้แก่ Random Oversampling, SMOTE, Random Undersampling และใช้ความสมดุลของ Class weight ในบางแบบจำลอง ซึ่งจะเห็นว่า เมื่อมีการจัดการความไม่สมดุลของข้อมูลในหลายแบบจำลองที่ใช้ parameter เดียวกัน พบว่ามีค่า Accuracy ลดลง และมีค่า Recall สูงขึ้น ซึ่งแสดงให้เห็นว่าแม้เป็นข้อมูลชุดเดียวกัน หากมีการทำให้ข้อมูลทำนายมีปริมาณใกล้เคียงกัน แบบจำลองที่ทำนายได้อาจจะมีประสิทธิภาพในการทำนายไม่ดีมากนัก ดังนั้นในการพิจารณาประสิทธิภาพของแบบจำลองชุดนี้ จึงไม่สามารถพิจารณาเฉพาะค่า Accuracy ได้ แต่ก็ไม่สามารถพิจารณาเฉพาะค่า Recall ได้ เนื่องจากบางแบบจำลองอาจจะมี Recall ที่สูงมาก แต่ Accuracy อาจจะต่ำมากได้ด้วยเช่นกัน ดังนั้นการดูประสิทธิภาพที่เหมาะสมกับข้อมูลชุดนี้มากที่สุด คือ การเลือกแบบจำลองโดยใช้ค่า AUC (Area Under Receiver Operating Characteristic curve) ซึ่งพบว่า แบบจำลองที่ใช้วิธี Logistic Regression ที่มีการจัดการข้อมูลที่ไม่สมดุล มีค่า AUC สูงที่สุด คือเท่ากับ 0.61 เมื่อเทียบกับแบบจำลองอื่น ๆ เช่น KNN Classifier, Random Forest หรือ XGBoost ซึ่งแสดงให้เห็นว่า ข้อมูลบางประเภท อาจจะไม่จำเป็นต้องใช้แบบจำลองที่ซับซ้อนเพื่อสร้างแบบจำลองในการทำนาย สามารถใช้แบบจำลองที่เข้าใจง่าย ก็เพียงพอในการทำนาย

เมื่อพิจารณาปัจจัยที่มีผลต่อการสร้างแบบจำลองในการทำนายการเข้ารับบริการซ้ำภายใน 72 ชั่วโมงของผู้ป่วยที่แผนกฉุกเฉิน จากแบบจำลองทั้งหมด 25 แบบจำลองที่สามารถคำนวณค่าความสำคัญสูงสุดได้ พบว่า ตัวแปรที่มีความสำคัญ ได้แก่ การวินิจฉัยโรคลำดับที่ 1 การวินิจฉัยโรคลำดับที่ 2 เพศ อายุ เชื้อชาติ ระดับความปวด ระยะเวลาการอยู่ในแผนกฉุกเฉิน และอัตราการเต้นของ โดยที่ตัวแปร การวินิจฉัยโรคลำดับที่ 1 มีค่าความสำคัญสูงสุดเป็นอันดับที่ 1 ใน 20 แบบจำลอง

เมื่อเปรียบเทียบกับงานวิจัยที่ใกล้เคียงกัน พบว่างานวิจัยนี้มีค่า AUC ที่ค่อนข้างต่ำเมื่อเปรียบเทียบกับการศึกษาของ Hsu และคณะ[4] ที่ในทุกแบบจำลองให้ผลของค่า AUC มากกว่า 0.69 โดยที่ XGB และ voting classifiers มีค่า AUC ดีที่สุดคือ 0.74

อาจเกิดจากในการศึกษาของ Hsu และคณะ มีการเฉพาะเจาะจงไปในผู้ป่วยที่มีอาการปวดท้อง และคัดออกผู้ป่วยที่มีการนัดหมายกลับมาซ้ำ ในขณะที่งานวิจัยนี้ไม่ได้เฉพาะเจาะจงโรคหรืออาการ ทำให้การทำนายมีประสิทธิภาพลดลง รวมทั้งในงานวิจัยนี้ไม่มีการคัดออกกลุ่มผู้ป่วยที่มีการนัดหมายกลับมาตรวจซ้ำ ในขณะที่การศึกษาของ Gao และคณะ[5] หากสร้างแบบจำลองที่ใช้ข้อมูลเฉพาะข้อมูลทั่วไปของผู้ป่วย และสถานะทางเศรษฐกิจและสังคม แบบจำลองในงานวิจัยของผู้วิจัยมีค่าสูงกว่า เนื่องจากในงานวิจัยนี้มีค่า C-statistic (เปรียบเทียบได้กับค่า AUC) เท่ากับ 0.556 แต่เมื่อการการเพิ่มปัจจัยอื่น ๆ ได้แก่ ค่าใช้จ่ายจากการเข้าพักในโรงพยาบาลในปีก่อนหน้า โรคประจำตัว พบว่าค่า C-statistic ค่อนข้างสูง คือ 0.763 นั่นอาจจะช่วยบ่งว่า หากงานวิจัยนี้มีการเพิ่มข้อมูลเช่น ข้อมูลโรคประจำตัว หรือการรักษาในปีที่ผ่านมา อาจจะช่วยเพิ่มประสิทธิภาพของแบบจำลองสำหรับงานวิจัยนี้ได้ แต่ทั้งนี้งานวิจัยของ Gao และคณะ เป็นการทำนายการกลับมาซ้ำใน 30 วัน ซึ่งแตกต่างจากงานวิจัยนี้ของผู้วิจัย การที่ระยะเวลาการใช้พิจารณาการกลับมาที่แผนกฉุกเฉินซ้ำนานขึ้น ส่งผลให้โอกาสที่มีข้อมูลที่เป็น Positive สูงขึ้นด้วย เมื่อพิจารณาเปรียบเทียบกับการศึกษาอื่น ที่ทำในลักษณะใกล้เคียงกัน งานวิจัยนี้อาจจะมีประสิทธิภาพของแบบจำลองที่ยังไม่สูงมากนัก โดยที่หากสามารถสร้างแบบจำลองที่มีค่า AUC ตั้งแต่ 0.7 ขึ้นไป จึงจะทำให้แบบจำลองนั้นได้รับการยอมรับที่มากขึ้น

กิตติกรรมประกาศ

การจัดทำวิจัยได้รับการสนับสนุนจากบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ ในการนำเสนอผลงานวิจัย ผู้วิจัยจึงขอขอบคุณมา ณ ที่นี้

เอกสารอ้างอิง

1. Pines JM. The Central Importance of Emergency Department Admission Rate Variation in Value-Based Care. Ann Emerg Med. 2021;78(4):484-6.

2. Smulowitz PB, O'Malley AJ, McWilliams JM, Zaborski L, Landon BE. Variation in Rates of Hospital Admission from the Emergency Department Among Medicare Patients at the Regional, Hospital, and Physician Levels. Annals of Emergency Medicine. 2021;78(4):474-83.

3. Benbassat J, Taragin M. Hospital readmissions as a measure of quality of health care: advantages and limitations. Arch Intern Med. 2000;160(8):1074-81.

4. Hsu CC, Chu CJ, Lin CH, Huang CH, Ng CJ, Lin GY, et al. A Machine Learning Model for Predicting Unscheduled 72 h Return Visits to the Emergency Department by Patients with Abdominal Pain. Diagnostics (Basel). 2021;12(1).

5. Gao K, Pellerin G, Kaminsky L. Predicting 30-day emergency department revisits. Am J Manag Care. 2018;24(11):e358-e64.

6. Johnson A, Bulgarelli L, Pollard T, Celi LA, Mark R, Steven H. MIMIC-IV-ED (version 2.2). PhysioNet. 2023.

7. Goldberger A, Amaral L, Glass L, Hausdorff J, Ivanov P, Mark R, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 2000;101(23):e215–e20.