ชื่อเรื่องวิทยานิพนธ์	การทำนายการเกิดภาวะเมแทบอลิกซินโดรมในผู้รับบริการ ณ โรงพยาบาลศูนย์การแพทย์มหาวิทยาลัยแม่ฟ้าหลวงด้วยการเรียนรู้ของเครื่อง
ผู้ประพันธ์	ลัดดาวัณย์ เจียมกิจวัฒนา
หลักสูตร	วิทยาศาสตรมหาบัณฑิต (วิทยาการวิเคราะห์ทางสุขภาพและชีวการแพทย์)
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.วิภพ สุทธนะ
อาจารย์ที่ปรึกษาร่วม	ผู้ช่วยศาสตราจารย์ ดร.ภัทรมน วุฒิพิทยามงคล

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินประสิทธิภาพของโมเดลการเรียนรู้ของเครื่องสำหรับการทำนายภาวะเมแทบอลิกซินโดรม (Metabolic Syndrome; MetS) โดยไม่อาศัยการตรวจเลือด ในกลุ่มตัวอย่างผู้รับบริการโรงพยาบาลศูนย์การแพทย์มหาวิทยาลัยแม่ฟ้าหลวง อายุ 18–65 ปี จำนวน 29,499 ราย 192,121 ครั้งของการเข้ารับบริการ ณ โรงพยาบาล ซึ่งครอบคลุมขอบเขตการใช้ข้อมูลประชากรศาสตร์ (อายุ เพศ อาชีพ สถานภาพสมรส) และข้อมูลสรีรวิทยาพื้นฐาน (ดัชนีมวลกาย เส้นรอบเอว ความดันโลหิต)

วิธีการศึกษาเป็นการสำรวจย้อนหลังแบบตัดขวาง (retrospective cross-sectional) โดยรวบรวมข้อมูลจากระบบเวชระเบียนอิเล็กทรอนิกส์ ตั้งแต่วันที่ 1 ธันวาคม พ.ศ. 2561 ถึงวันที่ 31 ธันวาคม พ.ศ. 2565 จากนั้นนำชุดข้อมูลทั้งหมดมาทำการแบ่งออกเป็นสองส่วน ได้แก่ ชุดข้อมูลสำหรับการฝึกสอนร้อยละ 70 และชุดข้อมูลสำหรับการทดสอบร้อยละ 30 ด้วยวิธีการสุ่มตัวอย่างอย่างง่าย ประยุกต์ใช้โมเดลการเรียนรู้ของเครื่องจำนวน 5 อัลกอริธึม ได้แก่ Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest (RF), Logistic Regression และ Extreme Gradient Boosting (XGBoost) โดยใช้ภาษาไพทอนในการดำเนินการ โดยแต่ละโมเดลจะถูกฝึกสอน (training) โดยใช้ข้อมูลชุดฝึกสอน และทดสอบประสิทธิภาพโดยใช้ชุดข้อมูลสำหรับทดสอบ การตรวจสอบความถูกต้องของโมเดลใช้วิธี 10-fold cross-validation และทำการปรับแต่งค่า hyperparameter ของแต่ละโมเดลด้วยวิธี grid search เพื่อหาค่าที่เหมาะสมที่สุด ประเมินประสิทธิภาพของแต่ละโมเดลบนชุดข้อมูลทดสอบ โดยใช้ตัวชี้วัด ได้แก่ ค่าความถูกต้องโดยรวม (Accuracy), ความแม่นยำเชิงบวก (Precision), ความไว (Recall), ค่า F1-score, และ พื้นที่ใต้โค้งลักษณะการทำงานของตัวรับ (ROC-AUC)

ผลการศึกษาพบว่า กลุ่มตัวอย่างที่มีภาวะเมแทบอลิกซินโดรมมีความแตกต่างทางสถิติในหลายตัวแปรสำคัญ เมื่อเทียบกับกลุ่มที่ไม่มีภาวะเมแทบอลิกซินโดรม ได้แก่ น้ำหนัก อายุ ความดันโลหิตตัวบน ความดันโลหิตตัวล่าง เพศ และ สถานภาพการสมรส โดยโมเดล XGBoost มี recall สูงสุดที่ 0.88 สะท้อนถึงศักยภาพในการตรวจจับผู้ที่มีความเสี่ยงต่อภาวะเมแทบอลิกได้ดีที่สุด โมเดล RF มี accuracy สูงสุดที่ 0.89 precision สูงสุดที่ 0.41 f1-score สูงสุดที่ 0.51 ค่า ROC-AUC สูงสุดที่ 0.91 เท่ากับ ANN แต่มี recall ต่ำสุดที่ 0.68 ซึ่งแสดงถึงประสิทธิภาพสมดุลในเกือบทุกตัวชี้วัดเว้นแต่การตรวจจับผู้ที่มีความเสี่ยงต่อภาวะเมแทบอลิก โดยนำโมเดลกลุ่ม XGBoost จากงานวิจัยนี้ มาสร้าง web application ที่สามารถนำมาทดลองใช้คัดกรองประเมินความเสี่ยงภาวะเมแทบอลิกซินโดรมกับข้อมูลสุขภาพในบริบทจริง

คำสำคัญ : ภาวะเมแทบอลิกซินโดรม, การเรียนรู้ของเครื่อง, การทำนายโดยไม่ต้องเจาะเลือด, การคัดกรองเบื้องต้น, XGBoost

Thesis Title	Machine Learning-based Prediction of Metabolic Syndromes Among Clients at Mae Fah Luang University Medical Center Hospital
Author	Laddawan Jiamkitwattana
Degree	Master of Science (Health and Biomedical Analytics)
Advisor	Assistant Professor Wipob Suttana, Ph. D.
Co-Advisor	Assistant Professor Pattaramon Vuttipittayamongkol, Ph. D.

ABSTRACT

This study aimed to develop and evaluate machine-learning models for predicting metabolic syndrome (MetS) without requiring blood tests, using routinely collected data from 29,499 adult clients (192,121 visits) aged 18–65 at the Mae Fah Luang University Medical Center Hospital. Predictor variables included demographic factors (age, sex, occupation, marriage status) and basic physiological measures (body mass index, waist circumference, and blood pressure).

A retrospective cross-sectional analysis of electronic health records was conducted spanning 1 December, 2018 to 31 December, 2022. Data were randomly split into 70% training and 30% testing sets via simple random sampling. Five algorithms, including Artificial Neural Network (ANN), Support Vector Machine, Random Forest (RF), Logistic Regression, and Extreme Gradient Boosting (XGBoost) were implemented in Python. Model validation used 10-fold cross-validation with grid-search hyperparameter tuning. Performance was assessed on the test set using accuracy, precision, recall, F1-score, and ROC-AUC.

The findings indicated that participants with metabolic syndrome exhibited statistically significant differences across several key variables—body weight, age, systolic blood pressure, diastolic blood pressure, sex, and marital status—compared with those without metabolic syndrome. XGBoost achieved the highest recall (0.88), indicating the best ability to detect individuals at risk for MetS. RF yielded the highest accuracy (0.89), precision (0.41), and F1-score (0.51), with ROC-AUC 0.91 (tied with

ANN), but showed the lowest recall (0.68). Based on its sensitivity, the XGBoost model was implemented as a prototype web application for pilot deployment to support real-world MetS risk screening at MFU MCH.

Keywords: Metabolic Syndrome, Machine Learning, Non-invasive Prediction, XGBoost, Primary Care Screening