

การพยากรณ์การสำรองเงินสดสำหรับสหกรณ์ออมทรัพย์ด้วยข้อมูล การถอนเงินช่วงเช้าโดยใช้การเรียนรู้ของเครื่อง

Forecasting Cash Withdrawals at Service Counters in Savings Cooperative Using Morning Transaction Data and Machine Learning

พิธิวัต ถาวรกุลชัย และ เสฏฐวิทย์ เกิดผล*

ภาควิชาวิทยาการคอมพิวเตอร์
มหาวิทยาลัยเกษตรศาสตร์
* ผู้รับผิดชอบบทความ
sethavidh.g@ku.th

Received: 6 Mar 2025

Revised: 27 Apr 2025

Accepted: 30 May 2025

บทคัดย่อ

การบริหารเงินสดสำรองที่มีประสิทธิภาพเป็นปัจจัยสำคัญต่อการดำเนินงานของสหกรณ์ออมทรัพย์ โดยเฉพาะอย่างยิ่งสำหรับเคาน์เตอร์บริการที่ต้องเตรียมเงินสดให้เพียงพอต่อความต้องการของสมาชิก งานวิจัยนี้นำเสนอการพัฒนาแบบจำลองการเรียนรู้ของเครื่องเพื่อทำนายปริมาณการถอนเงินของเคาน์เตอร์บริการ โดยใช้เทคโนโลยีการเรียนรู้ของเครื่อง โดยจำแนกเป็นสองช่วง ได้แก่ ไม่เกิน 3 ล้านบาท และมากกว่า 3 ล้านบาท เพื่อเพิ่มประสิทธิภาพในการบริหารเงินสดและลดต้นทุนการสำรองเงินสด ข้อมูลที่ใช้มาจากธุรกรรมการถอนเงินสดหน้าเคาน์เตอร์ย้อนหลัง 2 ปี ของสหกรณ์ออมทรัพย์แห่งหนึ่งในประเทศไทย โดยเลือกใช้ยอดการถอนเงินในช่วงเช้าและยอดการถอนเงินของวันก่อนหน้าเป็นตัวแปรอิสระในการสร้างแบบจำลอง การศึกษาได้เปรียบเทียบประสิทธิภาพของแบบจำลอง Gradient Boosting, Logistic Regression และ MLP พร้อมทั้งศึกษาผลกระทบของการเลือกช่วงเวลาที่ใช้เป็นตัวแปรอิสระ ผลการวิจัยชี้ให้เห็นว่าแบบจำลอง MLP ให้ผลลัพธ์ที่ดีที่สุดเมื่อใช้ข้อมูลการถอนเงินในช่วงเวลา 08:00–12:00 โดยมีความแม่นยำสูงสุดถึง 85%

คำสำคัญ: สำรองเงินสด การเรียนรู้ของเครื่อง การพยากรณ์ สหกรณ์ออมทรัพย์ แบบจำลอง Gradient Boosting แบบจำลอง Logistic Regression แบบจำลอง MLP

Abstract

Efficient cash reserve management is a critical factor in the operations of savings cooperatives, particularly for service counters that must ensure sufficient cash availability for members' transactions. This study proposes the development of machine learning models to forecast daily cash withdrawal for service counters. The approach utilizes machine learning techniques to classify withdrawal amounts into two categories: Not exceeding 3 million THB and more than 3 million THB, aiming to enhance cash management efficiency and reduce cash holding costs. The dataset comprises two years of historical counter withdrawal transactions from a savings cooperative in Thailand. The model employs morning withdrawal amounts and previous-day withdrawals as independent variables for prediction. This study evaluates the performance of Gradient Boosting, Logistic Regression, and Multi-Layer Perceptron (MLP) models and examines the impact of different time intervals on prediction accuracy. The results indicate that the MLP model achieved the highest accuracy of 84.80% when utilizing withdrawal data from 08:00 to 12:00, demonstrating its effectiveness in optimizing cash reserve forecasting.

Keywords: Cash reserve, Machine learning, Forecasting, Savings cooperative, Gradient Boosting, Logistic Regression, Multi-Layer Perceptron

1. บทนำ

สหกรณ์ออมทรัพย์ เป็นองค์กรทางการเงินที่เป็นลักษณะของสมาคมหรือสหภาพ มีวัตถุประสงค์เพื่อส่งเสริมให้สมาชิกรู้จักการออมทรัพย์และให้กู้ยืมเงินเมื่อเกิดความจำเป็น และได้รับการจดทะเบียนตามพระราชบัญญัติสหกรณ์ พ.ศ. 2542 [1] สหกรณ์ออมทรัพย์จึงเป็นสถาบันการเงินที่มีความสำคัญต่อการพัฒนาเศรษฐกิจและสังคมของประเทศ [2] เนื่องจากสหกรณ์ออมทรัพย์เป็นสถาบันทางการเงินที่มีบทบาทในการให้บริการแก่สมาชิกอย่างหลากหลาย บริการที่สำคัญอย่างหนึ่งคือ การรับฝากเงิน และการถอนเงินหน้าเคาน์เตอร์บริการ ซึ่งการฝากถอนเงินผ่านช่องทางเคาน์เตอร์บริการนั้น สหกรณ์จะต้องมีการสำรองเงินสดที่เพียงพอสำหรับรองรับการทำธุรกรรมของสมาชิกในแต่ละวัน การจัดการและบริหารเงินสดสำรองจึงเป็นสิ่งที่สำคัญ

เนื่องจากการให้บริการรับฝากเงินของสหกรณ์ออมทรัพย์นั้น เงินที่รับฝากไว้นั้นสหกรณ์จะต้องจ่ายดอกเบี้ยให้แก่สมาชิก ซึ่งหากสหกรณ์มีการสำรองเงินสดประจำวันไว้มากเกินไป จะทำให้สหกรณ์เกิดภาระในการจ่ายดอกเบี้ยของเงินส่วนนั้น และขาดโอกาสในการนำเงินส่วนที่สำรองไว้นั้นไปลงทุนเพิ่มเพื่อให้เกิดกำไร หรือในมุมกลับกันหากสำรองเงินสดประจำวันไว้น้อยเกินไป อาจไม่เพียงพอต่อการให้บริการสมาชิก ทำให้สมาชิกขาดความเชื่อมั่นหรือไม่พึงพอใจกับการใช้บริการ อย่างไรก็ตาม การคาดการณ์การสำรองเงินสดสำหรับการให้บริการหน้าเคาน์เตอร์เป็นสิ่งที่ยากลำบาก เนื่องจากมีปัจจัยหลายอย่างที่มีผลต่อพฤติกรรมการณ์การถอนเงินของสมาชิก เช่น วันที่จ่ายเงินเดือน ฤดูกาล และเหตุการณ์พิเศษต่าง ๆ เป็นต้น ทำให้การคาดการณ์ต้องคำนึงถึงปัจจัยที่หลากหลายและซับซ้อน

ในปัจจุบัน ขั้นตอนการสำรองเงินสดของสหกรณ์ออมทรัพย์ตัวอย่างในการศึกษานี้ มีสำนักงานให้บริการอยู่หนึ่งแห่ง โดยขั้นตอนการสำรองเงินสดนั้น สหกรณ์จะสำรองเงินสดพื้นฐานไว้ที่สำนักงานจำนวน 3 ล้านบาท ในทุก ๆ วันที่เปิดให้บริการ สหกรณ์จะทำการเบิกเงินสดเพิ่มเติมเมื่อจำเป็นจากธนาคารวัน

ละ 5 ล้านบาท เพื่อรองรับการทำธุรกรรมเงินสดหน้าเคาน์เตอร์ของสมาชิกในระหว่างวัน เมื่อถึงเวลาปิดให้บริการในช่วงเย็น สหกรณ์จะส่งเงินคืนส่วนเกินกลับเข้าธนาคาร โดยคำนวณเงินคืนจากยอดคงเหลือของวันนั้นลบด้วย 3 ล้านบาท ซึ่งเป็นการคืนเงินส่วนเกินที่เกินจากเงินสดสำรองพื้นฐาน นอกจากนี้ สหกรณ์ยังมีค่าบริการขนส่งเงินสดระหว่างธนาคารและสำนักงานวันละ 2 รอบ คือช่วงเช้าในการเบิกเงินและช่วงเย็นในการส่งคืน ทำให้เกิดต้นทุนค่าใช้จ่ายสำหรับการขนส่งเหล่านี้เป็นประจำทุกวัน รวมถึงต้นทุนจากการสำรองเงินสด ซึ่งมีภาระดอกเบี้ยสำหรับเงินสดที่ถูกสำรองไว้ อย่างไรก็ตาม วิธีการสำรองเงินในลักษณะดังกล่าวขาดความยืดหยุ่นและไม่สอดคล้องกับพฤติกรรมการณ์การทำธุรกรรมที่มีความผันผวนในแต่ละวัน ส่งผลให้บางวันมีการสำรองเงินสดเกินความจำเป็นหรือไม่เพียงพอต่อความต้องการ

เรียนรู้ของเครื่อง เป็นสาขาหนึ่งของปัญญาประดิษฐ์ ที่ช่วยให้ระบบเรียนรู้จากข้อมูลโดยไม่ต้องใช้โปรแกรมที่เขียนคำสั่งไว้ล่วงหน้า ใช้อัลกอริทึมในการปรับปรุง และทำนายผลลัพธ์จากข้อมูลที่ได้รับการฝึกฝน [3] ในงานวิจัยนี้ ใช้เพื่อพยากรณ์ปริมาณการถอนเงินรายวัน โดยจำแนกปริมาณการถอนเงินออกเป็นสองกลุ่ม ได้แก่ ไม่เกิน 3 ล้านบาท และมากกว่า 3 ล้านบาท การกำหนดช่วงที่ 3 ล้านบาทเป็นเกณฑ์ในการจำแนกจากนโยบายสำรองเงินสดของสหกรณ์ ที่กำหนดให้มีเงินสดคงเหลือขั้นต่ำ 3 ล้านบาทในแต่ละวัน หากปริมาณการถอนเงินคาดการณ์ว่าไม่เกิน 3 ล้านบาท หมายความว่าเงินสดสำรองที่มีอยู่เพียงพอ ไม่จำเป็นต้องเบิกเงินเพิ่มเติมจากธนาคาร ซึ่งสามารถช่วยลดค่าธรรมเนียมการขนส่งเงินสดและลดภาระดอกเบี้ยของเงินสดสำรอง ขณะที่หากคาดการณ์ว่าปริมาณการถอนจะสูงกว่า 3 ล้านบาท สหกรณ์จำเป็นต้องเบิกเงินสดเพิ่มเติมเพื่อรองรับการให้บริการถอนเงินของสมาชิก ดังนั้น การพยากรณ์ปริมาณการถอนเงินเป็นสองกลุ่ม จึงเพียงพอสำหรับการตัดสินใจเบิกเงินสดของสหกรณ์ภายใต้นโยบายสำรองเงินสดที่สหกรณ์กำหนด

2. การทบทวนผลงานที่เกี่ยวข้อง

Poorzaker Arabani และ Ebrahimpour Komleh [4] ทำการพยากรณ์ความต้องการเงินสดของตู้ ATM ในประเทศอิหร่าน โดยใช้แบบจำลอง CNN, SVM และ ANN โดยฝึกฝนแบบใช้และไม่ใช้

Cross-Validation พบว่า CNN ที่ใช้ Cross-Validation มีความแม่นยำสูงที่สุด (RMSE = 4.74%)

Riabykh และคณะ [5] ได้พยากรณ์ความต้องการถอนเงินสดของตู้ ATM ในรัสเซีย โดยใช้แบบจำลอง Holt-Winters SARIMA, Fbprophet, MLP, Random Forest, CNN และ LSTM โดยทดสอบโดยการทำนายล่วงหน้า 1 วัน, 15 วัน และ 30 วัน ผลลัพธ์พบว่า LSTM มีความแม่นยำสูงที่สุดในการทำนาย 1 วัน (WAPE = 0.288) และ 15 วัน (WAPE = 0.309) ขณะที่ Random Forest เหมาะกับการทำนาย 30 วัน (WAPE = 0.47)

Rajwani และคณะ [6] ได้พัฒนาแบบจำลองทำนายการถอนเงินสด โดยใช้แบบจำลอง Linear Regression, Ridge, Lasso, Bayesian Ridge, และ LSTM แล้วเปรียบเทียบประสิทธิภาพของแบบจำลอง โดยใช้ข้อมูลการถอนเงินสดผ่านตู้ ATM ของธนาคารแห่งหนึ่งในปากีสถาน พบว่า Linear Regression ให้ความแม่นยำสูงที่สุดถึง 98%

Rafi และคณะ [7] ได้ใช้ข้อมูลการถอนเงินสดผ่านตู้ ATM จำนวน 2,040 เครื่อง จากสถาบันการเงิน 6 แห่งในปากีสถาน โดยนำเสนอแบบจำลอง Time Series ARIMA Model for ATM (TASM4ATM) และเปรียบเทียบกับ RNN และ Amazon's DeepAR โดยทดสอบการทำนายถัดจากชุดข้อมูลฝึกฝน 30 วัน พบว่า TASM4ATM มีประสิทธิภาพดีที่สุด (MAPE = 3.6%)

Moubariki และคณะ [8] ได้พัฒนาแบบจำลอง Decision Tree, Random Forest และ Neural Networks คาดการณ์การใช้จ่ายภาครัฐ โดยใช้ข้อมูลคำสั่งจ่ายเงินที่อยู่ระหว่างดำเนินการ เพื่อพยากรณ์วันที่จะมีการจ่ายเงินจริง พบว่า Decision Tree มีความแม่นยำสูงที่สุดเมื่อใช้ Grid Search

Cedolin และ Genevois [9] ศึกษาการพยากรณ์การถอนเงินสดจากชุดข้อมูล NN5 ของ ATM ในสหราชอาณาจักร โดยเปรียบเทียบแบบจำลอง Regression, ETS, ARIMA, Prophet, RNN และ LSTM พบว่า การใช้เทคนิค DSMT ช่วยเพิ่มความแม่นยำได้ชัดเจน โดย Regression (DSMT) และ Prophet (DSMT) ให้ค่า SMAPE ต่ำสุดที่ 19.36 และ 19.69

Pratiwi และคณะ [10] ได้นำเสนอการใช้แบบจำลอง Linear Regression, Prophet, ARIMA และ LSTM เพื่อทำนายปริมาณธุรกรรมที่เกิดขึ้นในเครื่อง ATM ของธนาคาร XYZ พบว่า LSTM ให้ผลลัพธ์ดีที่สุด ($R^2 = 0.72$ MSE = 710,590,544.24

MAE = 20,686.91) ในขณะที่ Prophet และ ARIMA ทำงานได้ไม่ดีเนื่องจากข้อมูลมีความแปรปรวนสูง

Asad และคณะ [11] ใช้แบบจำลอง LSTM ทำนายปริมาณเงินสดที่ต้องเติมใน ATM ของธนาคารในปากีสถาน เปรียบเทียบกับแบบจำลอง DNN พบว่าแบบจำลอง LSTM ใช้ข้อมูลย้อนหลัง 7 วัน ซึ่งให้ค่าความผิดพลาด (RMSE = 132.53) ซึ่งแม่นยำกว่าแบบจำลอง DNN

Söylen และคณะ [12] ได้พัฒนาแบบจำลองพยากรณ์กระแสเงินสดโดยใช้ Wavelet Transform ร่วมกับ LSTM, GRU, CNN/LSTM และ XGBoost โดยใช้ข้อมูลบัญชีกระแสเงินสดของบริษัทแห่งหนึ่ง จำนวน 4 บัญชี พบว่า LSTM เหมาะกับบัญชีที่ยอดคงเหลือผันผวน (MAPE = 0.82%) และ XGBoost เหมาะกับบัญชีที่ยอดคงเหลือไม่ผันผวน (MAPE = 4.73%)

จากการทบทวนวรรณกรรมที่เกี่ยวข้อง งานวิจัยเน้นไปที่การแก้ปัญหาการบริหารจัดการเงินสดหรือสภาพคล่องทางการเงิน โดยส่วนใหญ่นำเสนอแนวทางการพยากรณ์โดยใช้ข้อมูลแบบอนุกรมเวลา โดยการใช้อยอดคงเหลือของวันก่อนหน้าเป็นปัจจัยหลักในการพยากรณ์แนวโน้มของวันถัดไป ขณะที่งานวิจัยนี้ใช้ข้อมูลธุรกรรมรายวันของสหกรณ์ ซึ่งยอดถอนเงินจะเริ่มต้นจากศูนย์บาททุกวัน ไม่สามารถอิงแนวโน้มจากอดีตได้โดยตรง งานวิจัยนี้จึงนำเสนอการใช้ข้อมูลการทำธุรกรรมช่วงเช้าเป็นปัจจัยหลักในการพยากรณ์ จุดเด่นของวิธีการนี้คือสามารถสะท้อนพฤติกรรมในวันที่พยากรณ์ได้โดยตรงและสอดคล้องกับข้อจำกัดด้านเวลาของการตัดสินใจสำรองเงินสดภายในวัน

3. ขั้นตอนวิธีดำเนินการวิจัย

การพยากรณ์ในงานวิจัยนี้ถูกออกแบบให้เป็นการจำแนกจำนวนการถอนหน้าเคาน์เตอร์รายวันออกเป็นสองกลุ่ม (Binary Classification) ได้แก่ ยอดถอนเงินไม่เกิน 3 ล้านบาท และยอดถอนเงินมากกว่า 3 ล้านบาท เพื่อสนับสนุนการตัดสินใจเบิกเงินสดเพิ่มจากธนาคาร ว่าควรดำเนินการเบิกเงินสดเพิ่มเติมจากธนาคารหรือไม่ การจำแนกยอดการถอนเงินออกเป็นสองกลุ่ม (Binary Classification) เป็นแนวทางที่เหมาะสมกว่าการพยากรณ์ค่าปริมาณการถอนเงินที่แน่นอน (Regression) สาเหตุสำคัญคือ หากปริมาณการถอนเงินเกิน 3 ล้านบาท ไม่ว่าจำนวน

ที่แท้จริงจะอยู่ที่ 4 ล้าน หรือ 7 ล้าน สหกรณ์จำเป็นต้องดำเนินการเบิกเงินสดเพิ่มเติมจากธนาคารอยู่ดี และการใช้เพียงข้อมูลช่วงเข้าอย่างเดียว ยังไม่เพียงพอต่อการพยากรณ์แบบเชิงปริมาณ (Regression) ให้ได้ค่าตัวเลขที่แม่นยำ เนื่องจากข้อมูลมีความผันผวนสูงและเริ่มต้นจากศูนย์ใหม่ทุกวัน ดังนั้นการจำแนกยอดการถอนเงินออกเป็นสองกลุ่ม (Binary Classification) จึงเพียงพอสำหรับการตัดสินใจเชิงปฏิบัติ

3.1 การรวบรวมและการเตรียมข้อมูล

ข้อมูลที่ใช้ในงานวิจัยนี้เป็นข้อมูลการทำธุรกรรมทางการเงินช่องทางเคาน์เตอร์บริการของสหกรณ์ออมทรัพย์แห่งหนึ่งในประเทศไทย ซึ่งเป็นข้อมูลการทำธุรกรรมย้อนหลังที่อยู่ในช่วงระหว่างวันที่ 1 มกราคม พ.ศ. 2565 ถึง 30 พฤศจิกายน พ.ศ. 2567 โดยข้อมูลดังกล่าวถูกบันทึกในรูปแบบธุรกรรมแยกรายการแสดงดังตารางที่ 1

ตารางที่ 1 รายละเอียดชุดข้อมูลธุรกรรมแยกรายการ

ชื่อคอลัมน์	คำอธิบาย
ACCOUNT_NO	เลขที่บัญชีเงินฝาก (เข้ารหัส)
ITEM_CODE	รหัสประเภทธุรกรรม
ITEM_DESC	ชื่อประเภทธุรกรรม
TRANSACTION_DATE	วันที่และเวลาที่ทำธุรกรรม
WITHDRAW	จำนวนเงินที่ถอน
DEPOSIT	จำนวนเงินที่ฝาก
PRNCBAL	ยอดเงินคงเหลือของบัญชีเงินฝาก

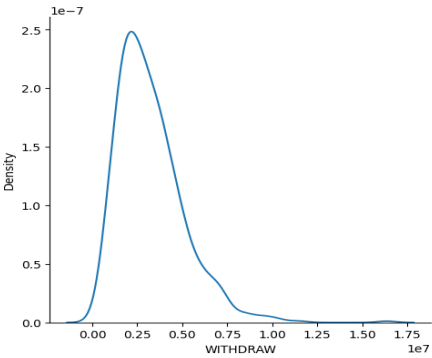
จากการตรวจสอบค่าทางสถิติ พบว่าค่าเฉลี่ยการถอนเงินอยู่ที่ 85,914.46 บาท ซึ่งสูงกว่าค่าเฉลี่ยการฝากที่ 55,166.81 บาท แสดงให้เห็นว่าปริมาณการถอนเงินโดยเฉลี่ยมีค่าสูงกว่าการรับฝากเงิน ขณะที่ค่ามัธยฐานอยู่ที่ 20,000 บาท และ 6,000 บาท ตามลำดับ สะท้อนให้เห็นว่าธุรกรรมการถอนเงินส่วนใหญ่อยู่ในระดับที่สูงกว่าการฝาก นอกจากนี้ การฝากเงินมีค่าสูงสุดที่ 10,000,000 บาท ซึ่งสูงกว่าการถอนเงินที่มีค่าสูงสุด 2,000,000 บาท ส่วนเบี่ยงเบนมาตรฐานของการถอนและฝากเงินอยู่ที่

181,894.10 บาท และ 171,593.70 บาท ตามลำดับ แสดงถึงการกระจายตัวของข้อมูลที่ค่อนข้างสูง

ตารางที่ 2 ค่าสถิติของธุรกรรมแยกรายการ

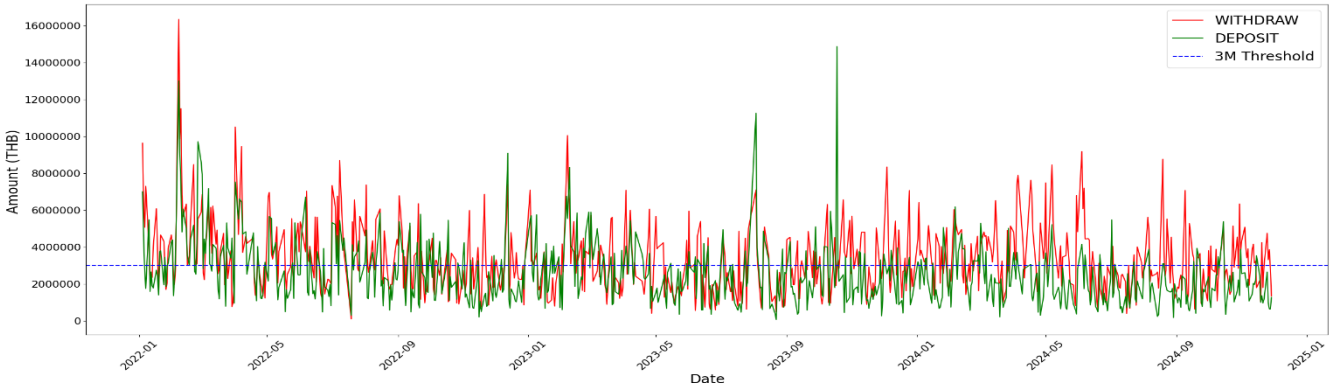
ค่าสถิติ	ยอดถอน(บาท) (N = 26,530)	ยอดฝาก(บาท) (N = 31,790)
Mean	85,914.46	55,166.81
Min	20	1.25
Q1	5,000	1,000
Median	20,000	6,000
Q3	68,968.41	30,000
Max	2,000,000	10,000,000
Std	181,894.10	171,593.70

ข้อมูลธุรกรรมเดิมที่อยู่ในรูปแบบของธุรกรรมแยกรายการ ถูกนำมาปรับรูปแบบข้อมูลให้เป็นรูปแบบรายวัน โดยการสรุปเป็นยอดรวมการถอนเงินแบบรายวัน เพื่อเตรียมเป็นชุดข้อมูลในการฝึกฝนและทดสอบแบบจำลอง



รูปที่ 1 แผนภูมิการแจกแจงข้อมูลปริมาณการถอนเงินสดรายวัน

หลังจากปรับข้อมูลให้อยู่ในรูปแบบรายวัน ผู้วิจัยได้สร้างแผนภูมิการแจกแจงข้อมูล (Distribution Plot) ตามรูปที่ 1 เพื่อตรวจสอบการกระจายตัวของยอดถอนเงิน พบว่ามีลักษณะ เบ้ขวา (Right-Skewed Distribution) ซึ่งสะท้อนให้เห็นว่าส่วนใหญ่ของวันมีการถอนเงินในปริมาณไม่มาก แต่มีบางวันที่มีการถอนเงินจำนวนมากผิดปกติ ลักษณะเช่นนี้แสดงถึงการมีอยู่ของ



รูปที่ 2 ปริมาณการถอนและฝากเงินสตรรายวัน

ค่าผิดปกติ (Outliers) ที่มีมูลค่าสูงมากเพียงไม่กี่รายการ โดยค่ามัธยฐานอยู่ที่ 20,000 บาท ขณะที่ค่าเฉลี่ยสูงกว่ามากที่ 85,914.46 บาท แสดงให้เห็นว่ามีธุรกรรมจำนวนน้อยที่มีมูลค่าสูงมาก

จากรูปที่ 2 กราฟปริมาณการถอนและฝากเงินรายวัน โดยแกน X เป็นวันที่ และแกน Y เป็นยอดถอนเงินสตรรวมในแต่ละวัน และเส้นประสีน้ำเงินแสดงระดับเกณฑ์ที่ 3 ล้านบาท จากกราฟพบว่า ปริมาณการถอนเงินมีแนวโน้มสูงกว่ากับการฝากเงินในหลายช่วง แต่มีบางจุดที่การฝากเงินพุ่งสูงขึ้นอย่างมีนัยสำคัญ ซึ่งสอดคล้องกับค่าสถิติของตารางที่ 2

จากนั้น ผู้วิจัยได้สร้างตัวแปรเป้าหมาย โดยจำแนกปริมาณการถอนเงินออกเป็นสองกลุ่ม แสดงดังตารางที่ 3

ตารางที่ 3 จำนวนข้อมูลตัวแปรเป้าหมาย

กลุ่มเป้าหมาย	จำนวน (วัน)
ยอดถอนเงินไม่เกิน 3 ล้านบาท	372
ยอดถอนเงินมากกว่า 3 ล้านบาท	330
รวม	702

ขั้นตอนถัดมา เป็นการสร้างตัวแปรอิสระจากยอดถอนเงินและจำนวนธุรกรรมในช่วงเวลาต่าง ๆ โดยในระยะเริ่มต้น ผู้วิจัยได้ทดลองใช้ข้อมูลยอดถอนของวันก่อนหน้าในแบบจำลองเชิงถดถอย (Regression Model) แต่ผลลัพธ์มีความแม่นยำต่ำ เนื่องจากยอดถอนเริ่มจากศูนย์ทุกวัน จึงไม่เหมาะกับการวิเคราะห์แบบอนุกรมเวลา ผู้วิจัยจึงเปลี่ยนมาใช้ข้อมูลช่วงเช้า

ของวันเดียวกัน และแบ่งช่วงเวลาออกเป็นหลายช่วง เพื่อตรวจสอบว่าธุรกรรมในช่วงเช้าและวันก่อนหน้า มีความสัมพันธ์กับยอดถอนเงินทั้งวันเพียงใด

ตารางที่ 4 รายละเอียดข้อมูลที่ได้หลังจากการเตรียมข้อมูล

ชื่อคอลัมน์	คำอธิบาย
GROUP_ENCODED	กลุ่มเป้าหมาย (0-3M, >3M)
WITHDRAW_SUM_PREVIOUS	ยอดถอนเงินวันก่อนหน้า
WITHDRAW_COUNT_PREVIOUS	จำนวนครั้งถอนเงินวันก่อนหน้า
WITHDRAW_SUM_08:00_09:00	ยอดถอนเงิน 1 ชม.แรกOfDay
WITHDRAW_COUNT_08:00_09:00	จำนวนครั้งถอน 1 ชม.แรกOfDay
WITHDRAW_SUM_08:00_10:00	ยอดถอนเงิน 2 ชม.แรกOfDay
WITHDRAW_COUNT_08:00_10:00	จำนวนครั้งถอน 2 ชม.แรกOfDay
WITHDRAW_SUM_08:00_11:00	ยอดถอนเงิน 3 ชม.แรกOfDay
WITHDRAW_COUNT_08:00_11:00	จำนวนครั้งถอน 3 ชม.แรกOfDay
WITHDRAW_SUM_08:00_12:00	ยอดถอนเงิน 4 ชม.แรกOfDay
WITHDRAW_COUNT_08:00_12:00	จำนวนครั้งถอน 4 ชม.แรกOfDay
DAY_NUMBER	หมายเลขวันที่
MONTH_NUMBER	หมายเลขเดือน
YEAR_NUMBER	ปี
DAY_OF_WEEK	วันในสัปดาห์
WEEK_OF_MONTH	สัปดาห์ของเดือน

ข้อมูลการถอนเงินถูกจัดกลุ่มเป็น 5 ช่วง ได้แก่ 08:00–09:00, 09:00–10:00, 10:00–11:00, 11:00–12:00 และวันก่อนหน้า เพื่อเปรียบเทียบความแม่นยำในการพยากรณ์ นอกจากนี้ ยังพิจารณาปัจจัยทางปฏิทิน เช่น วันในสัปดาห์ หรือต้นเดือน-ปลายเดือน โดยเพิ่มเป็นตัวแปรอิสระ เพื่อให้แบบจำลองสามารถเรียนรู้

แนวโน้มพฤติกรรมการถอนเงินได้ดียิ่งขึ้น รายละเอียดชุดข้อมูลที่ได้จากการรวบรวมและการเตรียมข้อมูล แสดงดังตารางที่ 4

3.2 การคัดเลือกคุณลักษณะและการวิเคราะห์ความสัมพันธ์
ขั้นตอนนี้เป็น การตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรเป้าหมาย โดยใช้สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) [13] เนื่องจากข้อมูลมีการกระจายตัวของตัวแปรอิสระและตัวแปรเป้าหมายไม่ปกติ จึงใช้ สัมประสิทธิ์สหสัมพันธ์อันดับของสเปียร์แมน (Spearman Correlation) [14] ในการวิเคราะห์ ผลลัพธ์แสดงดังนี้

Feature: WITHDRAW_SUM_08:00_12:00	Correlation: 0.6746
Feature: WITHDRAW_SUM_08:00_11:00	Correlation: 0.5807
Feature: WITHDRAW_COUNT_08:00_12:00	Correlation: 0.4996
Feature: WITHDRAW_COUNT_08:00_11:00	Correlation: 0.4712
Feature: WITHDRAW_SUM_08:00_10:00	Correlation: 0.4524
Feature: WITHDRAW_COUNT_08:00_10:00	Correlation: 0.4233
Feature: WITHDRAW_COUNT_PREVIOUS	Correlation: 0.2828
Feature: WITHDRAW_SUM_PREVIOUS	Correlation: 0.2794
Feature: WITHDRAW_SUM_08:00_09:00	Correlation: 0.2535
Feature: WITHDRAW_COUNT_08:00_09:00	Correlation: 0.2456
Feature: YEAR_NUMBER	Correlation: -0.0088
Feature: DAY_OF_WEEK	Correlation: -0.0942
Feature: MOUNTH_NUMBER	Correlation: -0.1639
Feature: WEEK_OF_MONTH	Correlation: -0.2403
Feature: DAY_NUMBER	Correlation: -0.2479

รูปที่ 3 ค่าสหสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรกลุ่มเป้าหมาย (0-3 ล้าน, มากกว่า 3 ล้าน)

จากรูปที่ 3 พบว่าตัวแปรที่เกี่ยวข้องกับยอดถอนเงินสดในช่วงเวลาต่าง ๆ ในวันเดียวกัน เช่น ยอดการถอนเงินสดช่วงเวลา 08:00–12:00 และ 08:00–11:00 มีความสัมพันธ์เชิงบวกที่สูงที่สุดกับตัวแปรเป้าหมาย แสดงให้เห็นถึงความสำคัญของข้อมูลการถอนเงินสดในช่วงเช้าในการพยากรณ์แนวโน้มยอดถอนเงินสดของวันนั้น ซึ่งยืนยันว่าข้อมูลยอดถอนเงินสดในช่วงเวลาที่กว้างขึ้นในตอนเช้ามีความสัมพันธ์กับพฤติกรรมการถอนเงินสดของวันนั้นมากที่สุด นอกจากนี้ ค่าสหสัมพันธ์ยังแสดงให้เห็นว่าตัวแปรที่เกี่ยวข้องกับวันที่ เช่น วันในสัปดาห์ และ หมายเลขเดือน มีความสัมพันธ์เชิงลบที่ต่ำ ส่วนตัวแปรอิสระยอดถอนเงินและจำนวนครั้งของวันก่อนหน้า มีค่าความสัมพันธ์เชิงลบที่ต่ำเช่นเดียวกัน แสดงให้เห็นว่า ปริมาณการถอนเงินในวันก่อนหน้าไม่ได้ส่งผลต่อปริมาณการถอนในวันถัดไปอย่างมีนัยสำคัญ บ่งชี้ว่าลักษณะของข้อมูลนี้ไม่ใช่ปัญหาแบบอนุกรมเวลา

3.3 การจัดโครงสร้างชุดข้อมูลตัวแปรอิสระตามช่วงเวลา
เพื่อให้สามารถเข้าใจถึงความสัมพันธ์ระหว่างข้อมูลในช่วงเวลาแต่ละช่วงกับตัวแปรเป้าหมาย (ยอดถอนเงินไม่เกิน 3 ล้านบาท และ ยอดถอนเงินมากกว่า 3 ล้านบาท) ได้ชัดเจนยิ่งขึ้น การทดลองจึงถูกแบ่งชุดข้อมูลที่เป็นตัวแปรอิสระออกเป็น 5 ชุด คือข้อมูลวันก่อนหน้า (Previous Day) และข้อมูลในช่วงเวลาช่วงเช้าของวันปัจจุบัน ได้แก่ 08:00–09:00, 08:00–10:00, 08:00–11:00 และ 08:00–12:00 โดยในแต่ละชุดช่วงเวลา ได้รวบรวมยอดถอนเงินสดรวม จำนวนครั้งของธุรกรรม และตัวแปรที่เกี่ยวข้องกับวันที่ เช่น หมายเลขวันในเดือน, หมายเลขเดือน และสัปดาห์ของเดือน รายละเอียดการกำหนดชุดตัวแปรอิสระ แสดงดังตารางที่ 5

ตารางที่ 5 ตัวแปรอิสระที่ใช้ในการศึกษาตามช่วงเวลา

ตัวแปรอิสระ/ ชุดตัวแปรอิสระ	Previous Day	08:00– 09:00	08:00– 10:00	08:00– 11:00	08:00– 12:00
ยอดถอนและจำนวน ครั้งของวันก่อนหน้า	✓				
ยอดถอนและจำนวน ครั้ง 1 ชั่วโมงแรก		✓			
ยอดถอนและจำนวน ครั้ง 2 ชั่วโมงแรก			✓		
ยอดถอนและจำนวน ครั้ง 3 ชั่วโมงแรก				✓	
ยอดถอนและจำนวน ครั้ง 4 ชั่วโมงแรก					✓
หมายเลขวัน	✓	✓	✓	✓	✓
หมายเลขเดือน	✓	✓	✓	✓	✓
สัปดาห์ของเดือน	✓	✓	✓	✓	✓

3.4 การแบ่งชุดข้อมูลฝึกฝนและทดสอบ, การลบค่าผิดปกติ และการทำให้ข้อมูลอยู่ในช่วงมาตรฐาน
ในขั้นตอนนี้ ผู้วิจัยได้แบ่งข้อมูลออกเป็นชุดข้อมูลฝึกฝน 70% จำนวน 491 วัน ชุดข้อมูลทดสอบ 30% จำนวน 211 วัน เพื่อใช้ในการพัฒนาและประเมินผลแบบจำลอง
เมื่อแบ่งชุดข้อมูลเรียบร้อยแล้ว ขั้นตอนถัดมาคือ การลบข้อมูลที่มียอดการถอนเงินรายวันที่ผิดปกติหรือค่าผิดปกติ (Outliers) ออกจากชุดข้อมูลฝึกฝน โดยใช้ช่วงระหว่างควอไทล์ (Interquartile Range: IQR) เป็นวิธีที่ใช้ตรวจจับค่าผิดปกติที่อยู่ห่างจากช่วงค่าปกติ เพื่อปรับปรุงคุณภาพของข้อมูลที่จะใช้ใน

การฝึกแบบจำลอง [15] หลังจากนั้นจะทำการปรับขนาดชุดข้อมูลฝึกและชุดทดสอบด้วยวิธีการปรับข้อมูลให้อยู่ในมาตรฐานเดียวกัน (Data Normalization) [16] เพื่อให้ตัวแปรอิสระทั้งหมดอยู่ในช่วงค่าที่เหมาะสม ซึ่งช่วยลดอิทธิพลของค่าที่มีความแตกต่างกันมากระหว่างตัวแปร

ตารางที่ 6 จำนวนข้อมูลก่อนและหลังการลบข้อมูลผิดปกติ

ชุดข้อมูล	ก่อนลบ Outlier		หลังลบ Outlier	
	0-3M	>3M	0-3M	>3M
ชุดข้อมูลฝึกฝน	264	277	263	213
ชุดข้อมูลทดสอบ	108	103	108	103

จากตารางที่ 6 จำนวนข้อมูลในกลุ่มยอดการถอนมากกว่า 3 ล้านบาท ลดลงหลังจากการลบค่าผิดปกติ (Outliers) ซึ่งแสดงถึงความผิดปกติในข้อมูลของกลุ่มนี้ ส่วนชุดข้อมูลทดสอบ ยังคงจำนวนข้อมูลเท่าเดิม เนื่องจากขั้นตอนการลบข้อมูลผิดปกติ (Outliers) ทำในข้อมูลชุดฝึกฝนเท่านั้น

3.5 แบบจำลองการเรียนรู้ของเครื่อง

ผู้วิจัยได้นำเสนอแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) ได้แก่ Logistic Regression, Gradient Boosting และ Multi-Layer Perceptron (MLP) ใช้ในการพยากรณ์ยอดถอนเงินรายวัน โดยการสร้างแบบจำลองเหล่านี้ใช้ไลบรารี scikit-learn ซึ่งเป็นเครื่องมือมาตรฐานที่ได้รับการยอมรับอย่างแพร่หลายสำหรับงานด้านการเรียนรู้ของเครื่องในภาษา Python [17] แบบจำลองแต่ละตัวถูกกำหนดค่าพารามิเตอร์ผ่านวิธี Grid Search Cross-Validation [18] โดยกำหนดค่า $k=5$ ซึ่งหมายถึงการแบ่งชุดข้อมูลฝึกฝนออกเป็นห้าส่วนที่มีขนาดใกล้เคียงกัน โดยแต่ละรอบของการเรียนรู้ แบบจำลองจะใช้ข้อมูล 4 ส่วนสำหรับการฝึกฝน และ 1 ส่วนสำหรับการตรวจสอบความถูกต้อง ขั้นตอนนี้จะดำเนินการซ้ำทั้งหมด 5 ครั้ง และใช้ ค่าความแม่นยำ (Accuracy) เป็นตัวชี้วัดในการประเมินเพื่อค้นหาชุดค่าพารามิเตอร์ที่เหมาะสมสำหรับการพยากรณ์แต่ละกรณี ค่าของพารามิเตอร์ที่กำหนด แสดงดังตารางที่ 7

ตารางที่ 7 กำหนดค่าพารามิเตอร์ GRID SEARCH CV

แบบจำลอง	พารามิเตอร์	ค่าที่กำหนด
Gradient Boosting	learning_rate	[0.001, 0.01, 0.05, 0.1]
	n_estimators	[50, 75, 100, 150]
	max_depth	[3, 5, 10]
Logistic Regression	(ค่าพื้นฐาน)	ไม่ได้ ปรับพารามิเตอร์ เนื่องจากใช้ค่าพื้นฐาน
MLP (Multi-Layer Perceptron)	hidden_layer_sizes	[(50, 20), (60, 30), (30, 10)]
	solver	['adam', 'sgd']
	alpha	[0.0001, 0.001, 0.01, 0.05]
	max_iter	[100, 150, 300]

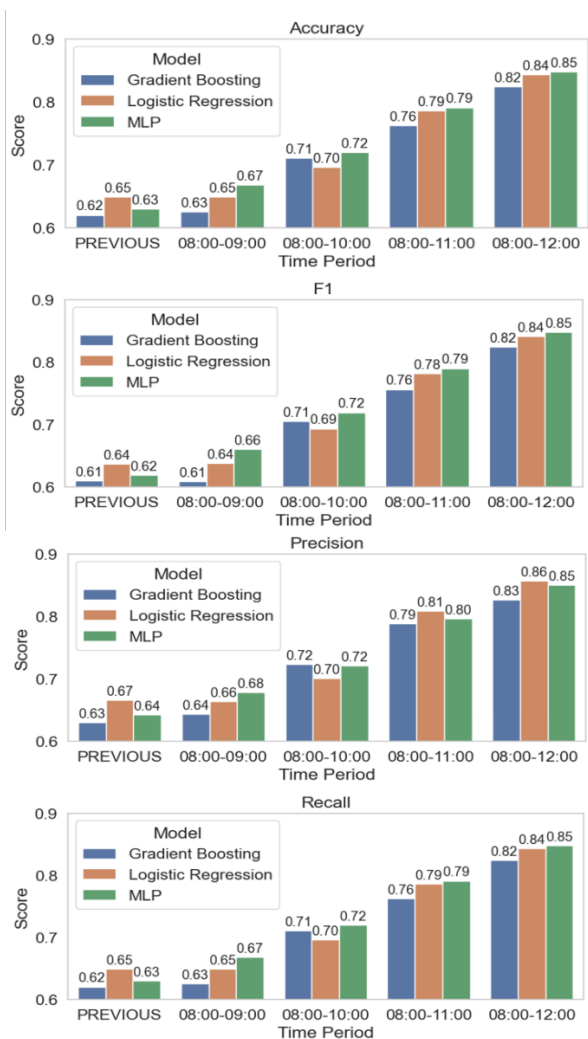
แบบจำลองทั้งหมดถูกฝึกแยกกันโดยใช้ข้อมูล 5 ชุดที่แบ่งตามช่วงเวลา ซึ่งค่าพารามิเตอร์ที่ดีที่สุดหลังจากผ่านขั้นตอน Grid Search Cross-Validation แสดงดังตารางที่ 8

ตารางที่ 8 พารามิเตอร์ที่ดีที่สุดของ Grid Search CV

ชุดตัวแปรอิสระ	Gradient Boosting (learning rate, max depth, estimators)	MLP (alpha, hidden layer, max iterations, solver)
Previous Day	(0.01, 3, 100)	(0.01, (60,30), 100, 'adam')
08:00–09:00	(0.01, 3, 75)	(0.01, (50,20), 150, 'adam')
08:00–10:00	(0.01, 3, 50)	(0.05, (60,30), 100, 'adam')
08:00–11:00	(0.01, 3, 50)	(0.001, (30,10), 150, 'adam')
08:00–12:00	(0.05, 3, 50)	(0.001, (50,20), 100, 'adam')

เมื่อได้ค่าพารามิเตอร์ที่เหมาะสมที่สุดจาก Grid Search Cross-Validation แล้ว พบว่าค่า alpha ของแบบจำลอง MLP มีแนวโน้มลดลงเมื่อใช้ชุดตัวแปรอิสระที่ครอบคลุมช่วงเวลานานขึ้น ซึ่งอาจเชื่อมโยงกับความแม่นยำที่เปลี่ยนไปตามลักษณะของข้อมูลฝึก ค่าพารามิเตอร์ดังกล่าวจะถูกนำไปใช้ในการฝึกฝนแบบจำลองด้วยชุดข้อมูลฝึกแต่ละช่วงเวลา

จากนั้นแบบจำลองที่ฝึกแล้วจะถูกประเมินกับชุดข้อมูลทดสอบ โดยใช้ตัวชี้วัด ได้แก่ ค่าความแม่นยำ (Accuracy), ความแม่นยำเฉพาะกลุ่ม (Precision), ความครอบคลุม (Recall) และ F1 (F1-score) เพื่อเปรียบเทียบประสิทธิภาพและผลของข้อมูลแต่ละช่วงเวลาต่อการพยากรณ์ ทั้งนี้ แบบจำลองบางประเภท เช่น MLP และ Gradient Boosting มีข้อจำกัดด้านการอธิบายผลลัพธ์ (interpretability) ซึ่งอาจไม่เหมาะกับบริบทที่ต้องการเหตุผลประกอบการตัดสินใจ

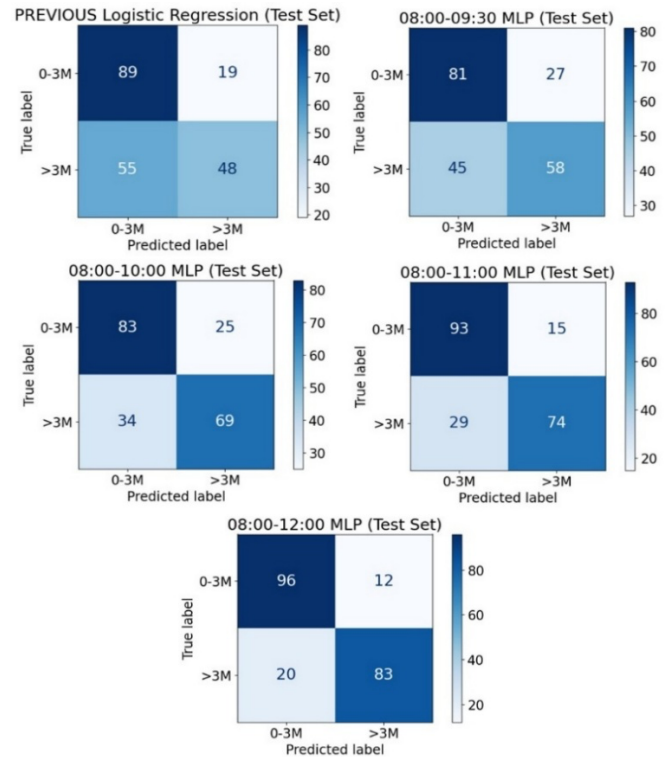


รูปที่ 3 กราฟผลลัพธ์ตัวชี้วัดของแบบจำลองในแต่ละช่วงเวลา

4. ผลการวิจัย

จากรูปที่ 3 ผลลัพธ์จากการทดลองแสดงให้เห็นว่าแบบจำลองทั้ง 3 แบบจำลอง ได้แก่ Gradient Boosting, Logistic Regression, และ MLP มีประสิทธิภาพที่แตกต่างกันเมื่อใช้ข้อมูลจากช่วงเวลา

ต่าง ๆ โดยทุกแบบจำลองแสดงผลลัพธ์ที่ดีที่สุดเมื่อใช้ข้อมูลช่วง 08:00-12:00 ซึ่งครอบคลุมข้อมูลในช่วงเวลาที่ยาวที่สุดในเช้าวันนั้น ค่าตัวชี้วัดทุกตัวมีค่าเพิ่มขึ้นอย่างชัดเจนเมื่อข้อมูลครอบคลุมช่วงเวลาที่นานขึ้น ส่วนตัวแปรอิสระที่ใช้อยอดถอนเงินสดวันก่อนหน้า (Previous Day) มีประสิทธิภาพต่ำที่สุดในทุกแบบจำลอง



รูปที่ 4 Confusion Matrix แบบจำลองที่ Accuracy สูงที่สุด

ในส่วน Confusion Matrix รูปที่ 4 แสดงถึงแนวโน้มเดียวกัน ยืนยันว่าการใช้ข้อมูลธุรกรรมที่ครอบคลุมช่วงเช้านั้นช่วยเพิ่มความแม่นยำในการพยากรณ์แนวโน้มการถอนเงินของทั้งวัน

แบบจำลอง Gradient Boosting มีผลลัพธ์ที่ดีขึ้นเรื่อย ๆ ตามช่วงเวลาที่ขยายออกไป โดยในช่วง 08:00-12:00 ค่าความแม่นยำ (Accuracy) และค่า F1 (F1-score) สูงสุดอยู่ที่ 0.82 แต่ค่าความแม่นยำเฉพาะกลุ่ม (Precision) และค่าความครอบคลุม (Recall) ยังคงค่อนข้างต่ำกว่าเมื่อเทียบกับแบบจำลองอื่น แบบจำลอง Logistic Regression มีประสิทธิภาพใกล้เคียงกันกับ Gradient Boosting ส่วนแบบจำลอง MLP มีประสิทธิภาพสูงสุดในทุกตัวชี้วัดเมื่อใช้ข้อมูลช่วง 08:00-12:00

จากการทดลองพบว่ามียอดเงินในชุดข้อมูลทดสอบที่ยอดถอนช่วงเช้าเกิน 3 ล้านบาท ได้แก่ ช่วงเวลา 08:00–10:00 (2 วัน), 08:00–11:00 (12 วัน) และ 08:00–12:00 (33 วัน) ข้อมูลเหล่านี้ช่วยเป็นสัญญาณสำคัญที่ระบุได้ว่า วันนั้นจะมียอดถอนเงินเกินกว่า 3 ล้านบาท จากการตรวจสอบผลการพยากรณ์พบว่า แบบจำลองทั้งหมดสามารถทำนายได้อย่างถูกต้องในทุกช่วงเวลา อย่างไรก็ตาม ในกรณีที่ยอดถอนช่วงเช้ายังไม่ถึงเกณฑ์แบบจำลองยังมีโอกาสทำนายผิดพลาด ซึ่งอาจส่งผลให้เกิดการทำนายผิดพลาดระหว่าง False Positive (คาดการณ์ว่าถอนเงินมากกว่า 3 ล้าน แต่จริง ๆ ต่ำกว่า 3 ล้าน) อาจทำให้สำรองเงินสดเกินความจำเป็น เพิ่มต้นทุนที่ไม่จำเป็น และ False Negative (คาดการณ์ว่าถอนเงินต่ำกว่า 3 ล้าน แต่จริง ๆ สูงกว่า 3 ล้าน) อาจนำไปสู่การขาดสภาพคล่องและกระทบต่อการให้บริการ ดังนั้น ตัวชี้วัดที่เหมาะสมที่สุดควรเลือกใช้ F1-score เนื่องจากช่วยรักษาสัมดุลระหว่าง Precision และ Recall ทำให้ช่วยลดความเสี่ยงจากการทำนายผิดทั้งสองประเภท

นอกจากนี้การเลือกช่วงเวลาที่ใช้ในการพยากรณ์ส่งผลโดยตรงต่อประสิทธิภาพในการบริหารเงินสดสำรอง แม้ว่าผลการทดลองจะแสดงให้เห็นว่า ช่วงเวลา 08:00–12:00 ให้ความแม่นยำสูงสุด แต่การรอนจนถึงเที่ยงอาจทำให้การตัดสินใจล่าช้าเกินไปเนื่องจากต้องประสานงานกับธนาคารเพื่อนำเงินสดมาส่ง ดังนั้น ควรเลือกช่วงเวลาที่สามารถพยากรณ์ได้เร็วขึ้น แม้จะความแม่นยำอาจลดลงเล็กน้อย แต่จะช่วยให้การสำรองเงินสดดำเนินการได้ทันเวลาและเหมาะสมกับข้อจำกัดในการปฏิบัติงานจริงของสหกรณ์

5. สรุปผลการวิจัย

การศึกษานี้มุ่งเน้นการใช้ยอดถอนเงินช่วงเช้าและยอดถอนเงินวันก่อนหน้า เพื่อพยากรณ์การสำรองเงินสดของสหกรณ์ออมทรัพย์ โดยใช้การเรียนรู้ของเครื่อง (Machine Learning) โดยเปรียบเทียบผลลัพธ์จาก 5 ชุดตัวแปรอิสระ ซึ่งครอบคลุมธุรกรรมในช่วงเช้าที่แตกต่างกัน (08:00–09:00, 08:00–10:00, 08:00–11:00 และ 08:00–12:00) รวมถึงข้อมูลยอดถอนของวันก่อนหน้า การฝึกแบบจำลองถูกดำเนินการแยกกันตามตัวแปรอิสระแต่ละชุด เพื่อ

จำแนกปริมาณการถอนเงินประจำวันให้อยู่ในช่วง 0-3 ล้านบาท หรือ มากกว่า 3 ล้านบาท

ผลการทดลองแสดงให้เห็นว่า ระยะเวลาของข้อมูลธุรกรรมช่วงเช้ามีผลโดยตรงต่อความแม่นยำของแบบจำลอง โดยเฉพาะ ช่วง 08:00–12:00 ซึ่งให้ผลลัพธ์ที่ดีที่สุดในทุกแบบจำลอง แบบจำลองที่ให้ความแม่นยำสูงสุดคือ MLP

อย่างไรก็ตาม ในบริบทของการนำไปใช้งานจริง การรอให้แบบจำลองได้รับข้อมูลถึง 08:00–12:00 อาจทำให้กระบวนการตัดสินใจล่าช้าเกินไปสำหรับการบริหารจัดการเงินสดของสหกรณ์ ดังนั้น ชุดตัวแปรอิสระช่วง 08:00–11:00 ซึ่งให้ผลลัพธ์ที่ใกล้เคียงกัน อาจเป็นทางเลือกที่เหมาะสมกว่า เนื่องจากช่วยให้สามารถตัดสินใจล่วงหน้าได้เร็วขึ้น งานวิจัยในอนาคตอาจพิจารณาปรับปรุงแบบจำลองให้สามารถพยากรณ์ได้ตั้งแต่ช่วงเวลาเช้าที่เร็วขึ้น หรือผสมผสานข้อมูลปัจจัยอื่น เช่น ข้อมูลของธุรกรรมฝากเงิน ข้อมูลเศรษฐกิจ หรือพฤติกรรมการทำธุรกรรมของสมาชิก เพื่อเพิ่มความแม่นยำและรองรับการใช้งานในสภาพแวดล้อมจริงอย่างมีประสิทธิภาพ

6. เอกสารอ้างอิง

- [1] ค. ชูชื่นมานะกิจ. (8 กรกฎาคม 2567). *ประเภทสหกรณ์*. [ระบบออนไลน์] แหล่งที่มา : <https://cpd.go.th/content-page/item/1741-coop-type.html>
- [2] D. L. Miller and C. Dou. “How Cooperative Are Savings and Credit Cooperatives in Developing Countries? An Analysis of Datasets from Uganda,” *Annals of Public & Cooperative Economics*, Vol. 88 (3), pp. 345–367, 2017. doi: 10.1111/apce.12155
- [3] O. Schmidt. *Machine Learning For Dummies®*. IBM Limited Edition. John Wiley & Sons, 2018.
- [4] S. Poorzaker Arabani and H. Ebrahimpour Komleh. “The Improvement of Forecasting ATMs Cash Demand of Iran Banking Network Using Convolutional Neural Network,” *Arabian Journal for Science & Engineering (Springer Science &*

- Business Media B.V.*), Vol. 44, pp. 3733–3743, 2019. doi: 10.1007/s13369-018-3647-7
- [5] A. Riabykh, I. Suleimanov, D. Surzhko, M. Konovalikhin, and V. Ryazanov. “ATM Cash Flow Prediction Using Local and Global Model Approaches in Cash Management Optimization,” *Pattern Recognition & Image Analysis*, Vol. 32, pp. 803–820, 2022. doi: 10.1134/S1054661822040113
- [6] A. Rajwani, T. Syed, B. Khan, and S. Iqbal. “ATM cash flow prediction: Exploratory Data Analysis and results from sequence-bases predictions,” *Elsevier Ltd*, Vol. 21, 2018.
- [7] M. Rafi, M. T. Wahab, M. B. Khan, and H. Raza. “Towards optimal ATM cash replenishment using time series analysis,” *Journal of Intelligent & Fuzzy Systems*, Vol. 41 (6), pp. 5915–5927, 2021. doi: 10.3233/JIFS-201
- [8] Z. Moubariki, L. Beljadid, M. El Haj Tirari, M. Kaicer, and R. O. H. Thami, “Enhancing cash management using machine learning”. *International Conference on Smart Systems and Data Science (ICSSD)*. 3-4 October. Rabat, Morocco : pp. 1–6, 2019. doi: 10.1109/ICSSD47982.2019.9002731
- [9] M. Cedolin and M. E. Genevois, “Cash Demand Prediction Problem using Econometric and Computational Intelligence Forecasting Models”. *International Conference on Control Decision and Information Technologies (CoDIT)*. 1-4 July. Valletta, Malta : pp. 103–108, 2024. doi: 10.1109/CoDIT62066.2024.10708473
- [10] P. S. Pratiwi, C. P. Utomo, and M. K. Amartya, “Machine Learning Model using Times Series Analytics for Prediction of ATM Transactions”. *International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. 8-9 December. Indonesia : pp. 775–780, 2022. doi: 10.1109/ISRITI56927.2022.10052900
- [11] M. Asad, M. Shahzaib, Y. Abbasi, and M. Rafi, “A Long-Short-Term-Memory Based Model for Predicting ATM Replenishment Amount”. *International Arab Conference on Information Technology (ACIT)*. 28-30 November. Zarqa, Jordan : pp. 1–6, 2020. doi: 10.1109/ACIT50332.2020.9300115
- [12] Z. Söylen, F. Mammadova, M. Fasounaki, A. İ. Özmen, and G. İnce, “Cash Flow Forecasting Based on Wavelet Transform and Neural Networks”. *International Conference on Computer Science and Engineering (UBMK)*, 13-15 September. Burdur, Türkiye : pp. 306–311, 2023. doi: 10.1109/UBMK59864.2023.10286711
- [13] C. Pruekpramool, N. Jaroentaku, and S. Srisuttiyakorn. “Efficiency of Pearson, Spearman and Kendall’s Correlation Coefficients When Data is Non-normal Distributed,” *An Online Journal of Education (OJED)*, Vol. 15 (2), pp. 1-16, 2020.
- [14] S. Wijitwanna. “Correlation in Statistics: How to Use,” *Journal of Humanities and Social Sciences, Rajapruk University*, Vol. 8 (2), pp. 1-15, 2022.
- [15] Y. Dodge. *The concise encyclopedia of statistics*. Springer, 2010.
- [16] Sankpal, Kalyani. “A Review on Data Normalization Techniques,” *International Journal of Engineering Research & Technology (IJERT)*, Vol. 9 (6), 2020.
- [17] J. Hao and T. K. Ho. “Machine Learning Made Easy: A Review of Scikit-learn Package in Python Programming Language,” *Journal of Educational and Behavioral Statistics*, Vol. 44 (3), pp. 348–361, 2019. doi: 10.3102/1076998619832
- [18] M. Adnan, A. A. S. Alarood, M. I. Uddin, and I. ur Rehman. “Utilizing grid search cross-validation with adaptive boosting for augmenting performance of machine learning models,” *PeerJ Comput Sci*, Vol. 8, 2022. doi: 10.7717/peerj-cs.803