COMPARATIVE STUDY OF ARIMA AND MACHINE LEARNING MODELS FOR BIOPROSTHETIC TISSUE HEART VALVE DEMAND FORECASTING AND ITS APPLICATION IN HIGH-VALUE MEDICAL SUPPLY INVENTORY MANAGEMENT

Waritnan PRUETHONG[1] and Panutporn RUANGCHOENGCHUM[1*]

1 College of Graduate Study in Management, Khon Kaen University, Thailand; rpanut@kku.ac.th (P. R.) (Corresponding Author)

ABSTRACT

This research aimed to (1) compare the forecasting performance of ARIMA and five Machine Learning (ML) models—Ridge Regression, Lasso Regression, Random Forest, Gradient Boosting, and XG-Boost together with two intermittent-demand benchmarks (Croston's Method and Syntetos-Boylan Approximation) for bioprosthetic tissue heart valve demand; (2) select the Best Model for nine valve items; and (3) develop an optimal inventory system and quantify financial savings at Queen Sirikit Heart Center of the Northeast. Monthly data for nine items (3 Porcine, 6 Bovine) over 36 months (1 January 2023-31 December 2025), totalling 324 data points, were purposively selected. Analyses employed ADF/KPSS stationarity tests, five-fold Walk-Forward Cross-Validation, Diebold-Mariano tests (baseline and pairwise), and MAE, RMSE, MASE metrics. Results indicated that all ML models significantly outperformed ARIMA (Pairwise DM Test, $p < .05$). Gradient Boosting achieved the lowest average MAE (0.354), whereas ARIMA recorded the highest (0.921). Per-SKU selection identified Ridge Regression as optimal for six items (66.67%), Gradient Boosting for two (22.22%), and XG-Boost for one (11.11%). At a 99% Service Level, optimal stock ranged from 36 to 59 units, reducing average inventory by **66.78%** (143 to 47.5 units), releasing approximately **9.81 million baht** in capital and saving an estimated **971,000 baht per year** in holding costs. This study presents the first empirical Best Model Selection framework in Thailand for high-value medical supply forecasting under intermittent demand.

Keywords: Demand Forecasting, Bioprosthetic Tissue Heart Valve, Machine Learning, ARIMA, Inventory Management

CITATION INFORMATION: Pruethong, W., & Ruangchoengchum, P. (2026). Comparative Study of ARIMA and Machine Learning Models for Bioprosthetic Tissue Heart Valve Demand Forecasting and Its Application in High-Value Medical Supply Inventory Management. *Procedia of Multidisciplinary Research, 4*(7), 52.

การศึกษาเปรียบเทียบประสิทธิภาพแบบจำลอง ARIMA และ Machine Learning ในการพยากรณ์ความต้องการลิ้นหัวใจเทียมชนิดเนื้อเยื่อ เพื่อเพิ่มประสิทธิ์ภาพการบริหารคลังเวชภัณฑ์มูลค่าสูง กรณีศึกษาศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ

วริทธิ์นันท์ ปรือทอง[1] และ ปณัทพร เรืองเชิงชุม[1*]

1 วิทยาลัยบัณฑิตศึกษาการจัดการ มหาวิทยาลัยขอนแก่น; rpanut@kku.ac.th (ปณัทพร) (ผู้ประพันธ์บรรณกิจ)

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์ 3 ประการ ได้แก่ (1) เปรียบเทียบประสิทธิภาพการพยากรณ์ระหว่างแบบจำลองค่าเฉลี่ยเคลื่อนที่ถดถอยปรับผสม (Autoregressive Integrated Moving Average: ARIMA) กับแบบจำลองการเรียนรู้ของเครื่อง (Machine Learning) 5 ประเภท ได้แก่ การถดถอยริดจ์ (Ridge Regression) การถดถอยลาสโซ่ (Lasso Regression) ป่าสุ่ม (Random Forest) การเพิ่มประสิทธิภาพแบบไล่ระดับ (Gradient Boosting) และ XG-Boost ร่วมกับแบบจำลองอ้างอิงเฉพาะทางอุปสงค์ไม่ต่อเนื่อง 2 ประเภท ได้แก่ วิธีของครอสตัน (Croston's Method) และ Syntetos-Boylan Approximation (SBA) สำหรับการพยากรณ์ความต้องการวัสดุปลูกถ่ายเนื้อเยื่อลิ้นหัวใจ (2) คัดเลือกแบบจำลองที่เหมาะสมที่สุดแยกตามรายการวัสดุ (Best Model Selection) สำหรับวัสดุ 9 รายการ และ (3) พัฒนาระบบพยากรณ์ความต้องการพร้อมกำหนดระดับสต็อกที่เหมาะสม กลุ่มตัวอย่างเป็นข้อมูลทุติยภูมิการเบิกใช้วัสดุปลูกถ่ายเนื้อเยื่อลิ้นหัวใจรายเดือน จำนวน 9 รายการ ประกอบด้วยกลุ่มเนื้อเยื่อหมู (Porcine) 3 รายการ และกลุ่มเนื้อเยื่อวัว (Bovine) 6 รายการ ครอบคลุม 36 เดือน (1 มกราคม 2566-31 ธันวาคม 2568) รวม 324 จุดข้อมูล คัดเลือกแบบเจาะจงจากศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ วิเคราะห์ด้วยการทดสอบความนิ่ง ADF/KPSS Walk-Forward Cross-Validation 5 รอบ Diebold-Mariano Test แบบ Baseline และ Pairwise และประเมินด้วย MAE, RMSE, MASE ผลการวิจัยพบว่า (1) แบบจำลองการเรียนรู้ของเครื่องทุกประเภทให้ผลการพยากรณ์แม่นยำกว่า ARIMA อย่างมีนัยสำคัญทางสถิติ (Pairwise DM Test, $p < .05$) โดย Gradient Boosting มีค่า MAE ต่ำสุด (0.354) ขณะที่ ARIMA มีค่า MAE สูงสุด (0.921) และไม่สามารถพยากรณ์ได้ดีกว่า Naive Forecast อย่างมีนัยสำคัญ ($p = .142$) (2) Ridge Regression เป็นแบบจำลองที่เหมาะสมที่สุดสำหรับ 6 รายการ (ร้อยละ 66.67) Gradient Boosting สำหรับ 2 รายการ (ร้อยละ 22.22) และ XG-Boost สำหรับ 1 รายการ (ร้อยละ 11.11) ยืนยันว่าไม่มีแบบจำลองใดเหมาะสมที่สุดกับทุกรายการวัสดุ และ (3) ระดับสต็อกที่เหมาะสมที่ระดับการให้บริการ (Service Level) ร้อยละ 99 อยู่ระหว่าง 36-59 ชิ้น ลดสต็อกจาก 143 ชิ้น เหลือ 47.5 ชิ้น (ร้อยละ 66.78) คิดเป็นมูลค่าสต็อกที่ลดลง 9.81 ล้านบาท และประหยัดต้นทุนการถือครองสินค้าคงคลังประมาณ 971,000 บาทต่อปี งานวิจัยนี้นำเสนอกรอบ Best Model Selection เป็นต้นแบบเชิงประจักษ์รายแรกของประเทศไทยสำหรับเวชภัณฑ์ฝังในร่างกายมูลค่าสูงประเภทอุปสงค์ไม่ต่อเนื่อง ซึ่งสามารถขยายผลไปยังศูนย์หัวใจอื่นและเวชภัณฑ์ฝังในร่างกายชนิดอื่นทั่วประเทศ

คำสำคัญ: การพยากรณ์อุปสงค์, ลิ้นหัวใจเทียมชนิดเนื้อเยื่อ, การเรียนรู้ของเครื่อง, อะรีมา, การบริหารคลังเวชภัณฑ์

ข้อมูลการอ้างอิง: วริทธิ์นันท์ ปรือทอง และ ปณัทพร เรืองเชิงชุม. (2569). การศึกษาเปรียบเทียบประสิทธิภาพแบบจำลอง ARIMA และ Machine Learning ในการพยากรณ์ความต้องการลิ้นหัวใจเทียมชนิดเนื้อเยื่อ เพื่อเพิ่มประสิทธิ์ภาพการบริหารคลังเวชภัณฑ์มูลค่าสูง กรณีศึกษาศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ. *Procedia of Multidisciplinary Research, 4*(7), 52.

บทนำ

โรคลิ้นหัวใจ (Valvular Heart Disease) เป็นปัญหาสาธารณสุขระดับโลกที่มีแนวโน้มรุนแรงขึ้น รายงาน Global Burden of Disease ปี 2564 ระบุว่ามีผู้ป่วยทั่วโลกกว่า 83.6 ล้านราย (GBD 2021 Valvular Heart Disease Collaborators, 2023) ในบริบทของประเทศไทยที่กำลังเข้าสู่สังคมผู้สูงอายุอย่างเต็มรูปแบบ อุบัติการณ์ของภาวะลิ้นหัวใจเสื่อมสภาพตามวัยเพิ่มขึ้นอย่างมีนัยสำคัญ ทำให้ลิ้นหัวใจเทียมชนิดเนื้อเยื่อ (Bioprosthetic Heart Valve Implants) กลายเป็นเวชภัณฑ์วิกฤตระดับความเสี่ยงสูงสุด (Level 5) ที่มีมูลค่า 65,000-240,000 บาทต่อชิ้น มีอายุการเก็บรักษาจำกัด และต้องควบคุมอุณหภูมิตลอดห่วงโซ่อุปทาน หากขาดแคลนหรือหมดอายุย่อมส่งผลกระทบโดยตรงต่อความปลอดภัยของผู้ป่วย (Axsäter, 2015; Otto et al., 2021)

ปัญหานี้ปรากฏชัดเจนที่ศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ ซึ่งเป็นโรงพยาบาลเฉพาะทางระดับตติยภูมิที่รองรับผู้ป่วยผ่าตัดลิ้นหัวใจลำดับที่ 2 ของประเทศ จำนวน 281 รายต่อปี (สมาคมศัลยแพทย์ทรวงอกแห่งประเทศไทย, 2566) ข้อมูลปี 2568 แสดงว่าหน่วยงานบริหารวัสดุดังกล่าว 9 รายการ มูลค่าทุนรวม 14.24 ล้านบาท มีภาระ Consignment 8.36 ล้านบาทต่อปี (ร้อยละ 58.7 ของมูลค่าคลัง) และต้นทุนการจัดเก็บสต็อก 1.44 ล้านบาทต่อปี (รายงานศูนย์หัวใจสิริกิติ์ฯ, 2568) ความซับซ้อนนี้มีรากฐานจาก Demand Uncertainty ใน 3 มิติ ได้แก่ (1) ระบบจัดซื้อที่พึ่งพา Consignment ก่อให้เกิด Bullwhip Effect และขาดแรงจูงใจในการพัฒนาระบบพยากรณ์ (Landry & Beaulieu, 2010) (2) ลักษณะอุปสงค์ไม่ต่อเนื่อง (Intermittent Demand) ที่เกิดจากดุลยพินิจของศัลยแพทย์ พยาธิสภาพเฉพาะรายผู้ป่วย และความพร้อมในการผ่าตัด (Kritchanchai & Meesamut, 2015) และ (3) ความแยกส่วนของข้อมูลจากทะเบียนผู้ป่วยรอผ่าตัด ระบบคลังเวชภัณฑ์ และบันทึกการใช้วัสดุในห้องผ่าตัด ซึ่งเป็นอุปสรรคต่อการพยากรณ์เชิงปริมาณ (Rezvova et al., 2023)

จากการทบทวนวรรณกรรมพบช่องว่างเชิงประจักษ์ 3 ประการ ได้แก่ (1) ยังไม่มีการเปรียบเทียบ ARIMA กับ ML ครบ 5 ประเภทสำหรับวัสดุปลูกถ่ายเนื้อเยื่อลิ้นหัวใจโดยเฉพาะ (2) ยังไม่มีการศึกษาในบริบทโรงพยาบาลภาครัฐไทยที่มีข้อจำกัดเชิงโครงสร้างเฉพาะ ได้แก่ ข้อมูลอนุกรมเวลาขนาดสั้น ระบบฝากสินค้า และดุลยพินิจทางคลินิก และ (3) ยังขาดกรอบการคัดเลือกแบบจำลองแบบรายรายการวัสดุอย่างเป็นระบบเพื่อนำไปกำหนดระดับสต็อกที่เหมาะสมในเชิงปฏิบัติ แม้ ARIMA จะให้ผลที่ยอมรับได้ในบริบทอุปสงค์สม่ำเสมอ (Kritchanchai & Meesamut, 2015) และ Croston's Method กับ SBA จะเหมาะสมกับอุปสงค์ไม่ต่อเนื่องในเชิงทฤษฎี (Syntetos & Boylan, 2005; Teunter et al., 2011) แต่ยังไม่มีงานวิจัยใดเปรียบเทียบทั้งสองแนวทางอย่างครอบคลุมในบริบทนี้

ด้วยเหตุนี้ ผู้วิจัยจึงสนใจศึกษาเปรียบเทียบประสิทธิภาพแบบจำลอง ARIMA และ Machine Learning ในการพยากรณ์ความต้องการลิ้นหัวใจเทียมชนิดเนื้อเยื่อ เพื่อเพิ่มประสิทธิ์ภาพการบริหารคลังเวชภัณฑ์มูลค่าสูง กรณีศึกษาศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ โดยหวังว่างานวิจัยนี้จะสนับสนุนการตัดสินใจในการบริคลังเวชภัณฑ์ บนพื้นฐานหลักฐานเชิงประจักษ์และเป็นต้นแบบสำหรับการขยายผลไปยังศูนย์หัวใจอื่นและเวชภัณฑ์ฝังในร่างกายชนิดอื่นในระบบสาธารณสุข

การทบทวนวรรณกรรม

แบบจำลองค่าเฉลี่ยเคลื่อนที่ถดถอยปรับผสม (ARIMA) ตามกรอบแนวคิดของ Box and Jenkins (Box et al., 2015) เป็นแบบจำลองอนุกรมเวลาเชิงเส้นที่ได้รับการยอมรับในการพยากรณ์เวชภัณฑ์ที่มีอุปสงค์สม่ำเสมอ (Kritchanchai & Meesamut, 2015) อย่างไรก็ตาม ARIMA มีข้อจำกัดสำคัญคือไม่สามารถจับความสัมพันธ์แบบไม่เป็นเชิงเส้นและไม่สามารถรวมตัวแปรภายนอกได้โดยตรง ทำให้ประสิทธิภาพลดลงอย่างชัดเจนเมื่อข้อมูลมีลักษณะอุปสงค์ไม่ต่อเนื่อง (Wang et al., 2024)

สำหรับเวชภัณฑ์ที่มีอุปสงค์ไม่ต่อเนื่อง Syntetos and Boylan (2005) พัฒนา Syntetos-Boylan Approximation (SBA) ซึ่งแก้ไขอคติเชิงบวกของ Croston's Method ด้วยตัวคูณปรับแก้ $(1 - \alpha/2)$ ทั้งสองแบบจำลองถือเป็นมาตรฐานอ้างอิง

สำหรับข้อมูลที่มีค่าศูนย์จำนวนมาก (Babai et al., 2022) อย่างไรก็ตาม แบบจำลองกลุ่มนี้ใช้เพียงข้อมูลอนุกรมเวลาเดียว จึงไม่สามารถบูรณาการตัวแปรเชิงบริบททางคลินิกได้ (Syntetos et al., 2023; Li et al., 2023)

การเรียนรู้ของเครื่องสำหรับการพยากรณ์อุปสงค์ การเรียนรู้ของเครื่องได้รับความสนใจเพิ่มขึ้นในการพยากรณ์ความต้องการทางการแพทย์ เนื่องจากสามารถจับความสัมพันธ์แบบไม่เป็นเชิงเส้นและบูรณาการตัวแปรหลายมิติได้ (Fang et al., 2022) Ridge Regression และ Lasso Regression เป็นเทคนิค Regularization ที่มีประสิทธิภาพในบริบทที่ข้อมูลมีขนาดเล็ก (Hoerl & Kennard, 1970; Tibshirani, 1996) Random Forest ซึ่งพัฒนาโดย Breiman (2001) เป็น Ensemble Method แบบ Bagging ที่ทนทานต่อค่าผิดปกติ โดย Thakur et al. (2024) ยืนยันว่า ML ให้ผลการพยากรณ์เวชภัณฑ์ดีกว่าแบบจำลองสถิติดั้งเดิมอย่างมีนัยสำคัญ Gradient Boosting และ XG-Boost เป็นเทคนิค Boosting ที่มีประสิทธิภาพสูง โดย Alim et al. (2020) แสดงให้เห็นว่า XG-Boost เหนือกว่า ARIMA ในบริบทเอเชียอย่างมีนัยสำคัญ

การประเมินประสิทธิภาพและการคัดเลือกแบบจำลอง สำหรับอุปสงค์ไม่ต่อเนื่องใช้ตัวชี้วัดหลัก ได้แก่ MAE, RMSE และ MASE โดย MASE เหมาะสมที่สุดสำหรับข้อมูลที่มีค่าศูนย์จำนวนมาก เนื่องจากไม่มีปัญหาตัวส่วนเป็นศูนย์ และค่า MASE น้อยกว่า 1 หมายความว่าแบบจำลองดีกว่าการพยากรณ์แบบ Naive (Hyndman & Koehler, 2006) Bentéjac et al. (2021) ยืนยันว่าไม่มีแบบจำลองใดเหนือกว่าทุกชุดข้อมูล ดังนั้นการคัดเลือกแบบรายรายการวัสดุ (Per-SKU Selection) จึงเหมาะสมกว่าการใช้แบบจำลองเดียวกับวัสดุทุกประเภท (Petropoulos et al., 2022) นอกจากนี้ การทดสอบ Diebold-Mariano (DM Test) ใช้ยืนยันนัยสำคัญทางสถิติของความแตกต่างระหว่างแบบจำลองโดยเฉพาะในกลุ่มตัวอย่างขนาดเล็ก (Coroneo & Iacone, 2020)

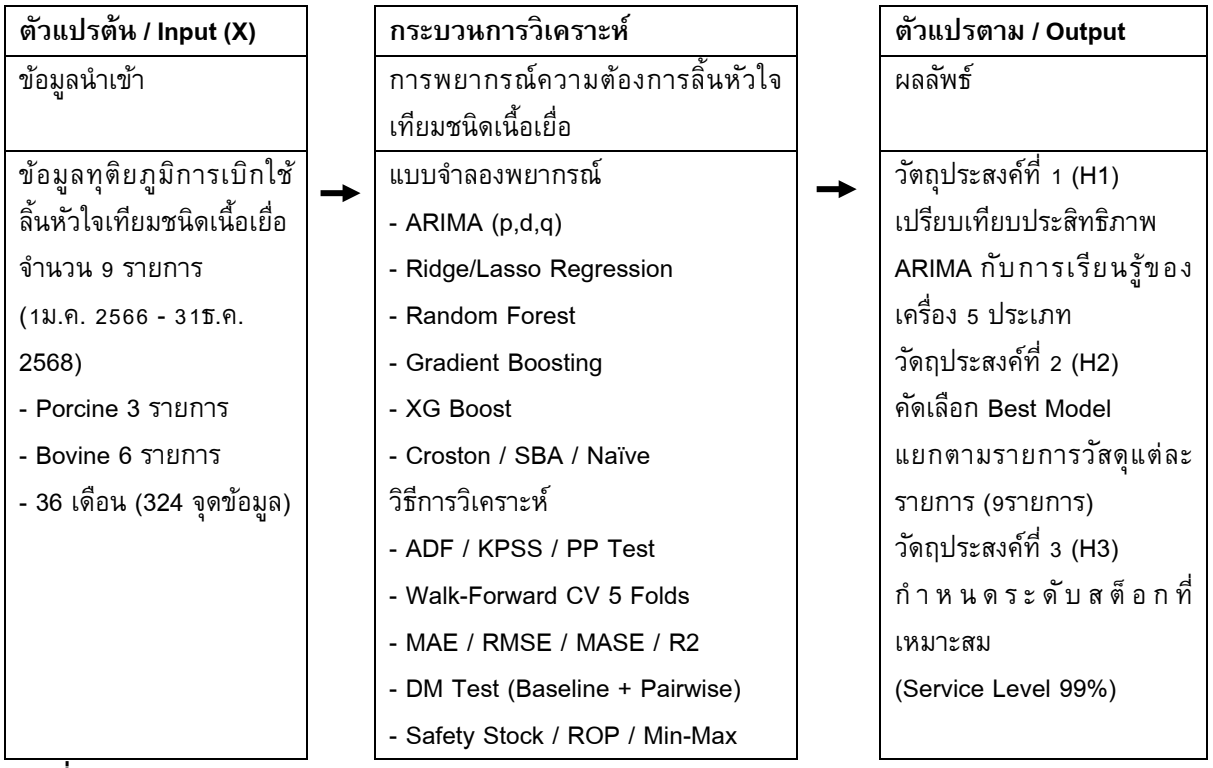
สมมติฐานการวิจัย

H1: แบบจำลอง ML อย่างน้อย 1 ประเภท ให้ประสิทธิภาพการพยากรณ์สูงกว่า ARIMA อย่างมีนัยสำคัญทางสถิติ เมื่อวัดด้วย MAE, RMSE และ MASE

H2: แบบจำลองที่เหมาะสมที่สุด (Best Model) สำหรับวัสดุแต่ละรายการมีความแตกต่างกัน กล่าวคือไม่มีแบบจำลองใดประเภทเดียวที่ให้ผลดีที่สุดกับทุกรายการวัสดุ

H3: ระบบพยากรณ์ที่พัฒนาขึ้นสามารถกำหนดระดับสต็อกที่เหมาะสมซึ่งต่ำกว่าระดับสต็อกในปัจจุบัน

กรอบแนวคิดการวิจัย



ภาพที่ 1 กรอบแนวคิด

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ผ่านการรับรองจากคณะกรรมการจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยขอนแก่น เลขที่ HE691189 เป็นการวิจัยเชิงปริมาณประเภทการวิจัยเชิงทำนาย (Predictive Research) ที่บูรณาการระเบียบวิธีทางสถิติ ML และการบริหารคลังสินค้า โดยใช้ข้อมูลทุติยภูมิจากระบบสารสนเทศคลังเวชภัณฑ์ของศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ ประชากรและกลุ่มตัวอย่าง ประชากร ได้แก่ ข้อมูลการเบิกใช้วัสดุปลูกถ่ายเนื้อเยื่อลิ้นหัวใจทุกรายการของศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ ระหว่างวันที่ 1 มกราคม 2566 ถึง 31 ธันวาคม 2568 กลุ่มตัวอย่างเลือกแบบเจาะจง (Purposive Sampling) จำนวน 9 รายการ ประกอบด้วยกลุ่มเนื้อเยื่อหมู (Porcine) 3 รายการ และกลุ่มเนื้อเยื่อวัว (Bovine) 6 รายการ ครอบคลุม 36 เดือน รวม 324 จุดข้อมูล เกณฑ์คัดเข้าคือมีการเบิกใช้ต่อเนื่องและมีข้อมูลครบ 36 เดือน เครื่องมือที่ใช้ในการวิจัย ใช้โปรแกรม IBM SPSS Statistics เวอร์ชัน 29 (ลิขสิทธิ์มหาวิทยาลัยขอนแก่น สำหรับสร้างแบบจำลอง ARIMA ผ่านโมดูล Expert Modeler ซึ่งคัดเลือกพารามิเตอร์ (p,d,q) โดยอัตโนมัติตามเกณฑ์ AIC และ BIC ต่ำสุด ภาษา Python เวอร์ชัน 3.10 ร่วมกับไลบรารี scikit-learn สำหรับแบบจำลอง ML, xgboost สำหรับ XG-Boost, stats models สำหรับ Croston's Method และ SBA และ shap สำหรับวิเคราะห์ความสำคัญของตัวแปร และ Microsoft Excel 365 สำหรับคำนวณ Safety Stock และ Reorder Pointการวิเคราะห์ข้อมูล ดำเนินการ 7 ขั้นตอน ได้แก่ ขั้นตอนที่ 1 ทดสอบความนิ่งด้วย ADF Test และ KPSS Test (α = .05) และ PP Test เสริมสำหรับรายการ Borderline ขั้นตอนที่ 2 พัฒนาแบบจำลอง ARIMA (p,d,q) คัดเลือกด้วยเกณฑ์ AIC และ BIC ต่ำสุด และทดสอบค่าตกค้างด้วย Ljung-Box Q-Test เพื่อยืนยัน White Noise ขั้นตอนที่ 3 สร้าง Feature Engineering 5 กลุ่ม รวม 15 ตัวแปร ได้แก่ Temporal (Lag, Rolling Mean), Trend (แนวโน้มอนุกรม), Calendar (เดือน ไตรมาส วันหยุดราชการ), Intermittent Demand (ADI, CV^2) และ SKU-level (ประเภทเนื้อเยื่อ Demand Profile) โดยสร้างภายในช่วงการฝึกของแต่ละ Fold เพื่อป้องกัน Data Leakage และพัฒนาแบบจำลอง ML 5 ประเภท ร่วมกับ Naive Forecast, Croston's Method และ SBA ขั้นตอนที่ 4 ทำ Walk-Forward Cross-Validation แบบ

Expanding Window 5 รอบ (Fold 1: t = 1-24/25-26 ถึง Fold 5: t = 1-35/36) ปรับพารามิเตอร์ด้วย Grid Search ภายในแต่ละ Fold ขั้นตอนที่ 5 ประเมินด้วย MAE, RMSE, MASE และ R^2 ทดสอบนัยสำคัญด้วย DM Test แบบ Baseline และ Pairwise คำนวณจาก Pooled Forecast Errors ทุก Fold (n = 81) ตามแนวทาง Fixed-Smoothing Asymptotics ขั้นตอนที่ 6 คัดเลือก Best Model แต่ละรายการด้วย MAE และ MASE เป็นเกณฑ์หลักจำแนกประเภทอุปสงค์ตามเกณฑ์ SBC Classification (ADI = 1.32, CV^2 = 0.49) ก่อนขั้นตอน Feature Engineering เพื่อนำ Demand Profile ไปใช้เป็น Item-Level Feature

ขั้นตอนที่ 7 กำหนดระดับสต็อกที่เหมาะสมโดยใช้วิธีการ Hybrid Distribution ที่สอดคล้องกับลักษณะอุปสงค์ของแต่ละรายการ แบ่งเป็น 2 กลุ่ม ได้แก่ (1) รายการที่มีอุปสงค์เฉลี่ยรายเดือน ($\bar{D} \geq 1$ ชิ้น/เดือน จำนวน 5 รายการ (Porcine_E_M, Porcine_O_H, Bovine_P_A, Bovine_P_M, Bovine_O_A) ใช้สูตร Normal Approximation: $SS = z \times \sqrt{(L \times \sigma D^2 + \bar{D}^2 \times \sigma L^2)}$ โดย z = 2.33 ที่ Service Level ร้อยละ 99 ตามนโยบาย Critical Item ของ WHO (2017), L = 1 เดือน, σL = 0.25 เดือน จากประวัติการส่งมอบจริงปี 2566-2568 เนื่องจากรายการเหล่านี้มีอุปสงค์สม่ำเสมอเพียงพอที่การแจกแจงแบบปกติสามารถประมาณได้อย่างสมเหตุสมผล และ (2) รายการที่มี $\bar{D} < 1$ ชิ้น/เดือน จำนวน 4 รายการ (Porcine_E_A, Bovine_O_M, Bovine_O_T, Bovine_O_InsA) ซึ่งมีลักษณะอุปสงค์ไม่ต่อเนื่องและมีค่าศูนย์จำนวนมาก (Zero-Inflated) ใช้การแจกแจงแบบปัวซง (Poisson Distribution) โดยกำหนด $\lambda = \bar{D} \times L$ และคำนวณ Reorder Point เป็นค่า k น้อยสุดที่ $P(X \leq k) \geq 0.99$ ตามแนวทางของ Axsäter (2015) และ Syntetos et al. (2023) ซึ่งยืนยันว่าการแจกแจงแบบปัวซงเหมาะสมกว่าการแจกแจงแบบปกติสำหรับข้อมูลที่มีอัตราการเกิดเหตุการณ์ต่ำและมีค่าศูนย์จำนวนมาก จากนั้นคำนวณ Safety Stock, Reorder Point และระดับ Min-Max ครบทั้ง 9 รายการ

ผลการวิจัย

จากตารางที่ 1 ผลการทดสอบความนิ่งของข้อมูลพบว่า ผลการทดสอบด้วย ADF Test และ KPSS Test พบว่าวัสดุทั้ง 9 รายการมีค่า p-value ของ ADF Test ต่ำกว่า .05 และค่า p-value ของ KPSS Test สูงกว่า .05 แสดงว่าข้อมูลทุกรายการมีคุณสมบัติ Stationary เหมาะสมสำหรับการสร้างแบบจำลอง ARIMA โดยตรง โดยไม่จำเป็นต้องผ่านขั้นตอน Differencing เพิ่มเติม สำหรับ Bovine_O_A ซึ่งให้ผล Borderline ได้ใช้ PP Test เสริมยืนยัน (p = .018)

จากตารางที่ 2 ผลการเปรียบเทียบประสิทธิภาพแบบจำลอง (วัตถุประสงค์ที่ 1- H1) พบว่า แบบจำลอง ML ทุกประเภทให้ผลการพยากรณ์ดีกว่า ARIMA อย่างชัดเจน โดย Gradient Boosting มีค่า MAE ต่ำสุด (0.354) และ MASE ต่ำสุด (0.218) ขณะที่ ARIMA มีค่า MAE สูงสุด (0.921) ที่น่าสังเกตคือ ARIMA ไม่สามารถพยากรณ์ได้ดีกว่า Naive Forecast อย่างมีนัยสำคัญ (p = .142) บ่งชี้ว่าไม่สามารถสกัดสัญญาณที่มีประโยชน์จากข้อมูล zero-inflated intermittent ได้

จากตารางที่ 3 ผลการทดสอบ Pairwise DM Test (วัตถุประสงค์ที่ 1 H1) พบว่า แบบจำลอง ML ทุกประเภทให้ผลดีกว่า ARIMA อย่างมีนัยสำคัญทางสถิติ (p < .05) สนับสนุน H1 โดย Gradient Boosting มี DM Statistic สูงสุด (-5.421, p < .001) สะท้อนความเหนือกว่า ARIMA มากที่สุด แม้แต่ Croston's Method และ SBA ซึ่งออกแบบมาเฉพาะสำหรับอุปสงค์ไม่ต่อเนื่องก็ยังดีกว่า ARIMA อย่างมีนัยสำคัญ

จากตารางที่ 4 ผลการคัดเลือก Best Model แยกตามรายการวัสดุ (วัตถุประสงค์ที่ 2 — H2) พบว่า Ridge Regression เป็น Best Model สำหรับ 6 รายการ (ร้อยละ 66.67) โดยเฉพาะรายการที่มี Demand Profile แบบ Intermittent ซึ่งมีค่าศูนย์จำนวนมากและอุปสงค์แปรปรวนต่ำ Gradient Boosting เหมาะสมกับ 2 รายการ (ร้อยละ 22.22) ที่มีรูปแบบ Erratic และ Smooth-Lumpy และ XG-Boost เหมาะสมกับ 1 รายการ (ร้อยละ 11.11) ที่มีอุปสงค์ Smooth ผลนี้ยืนยัน H2 ว่าไม่มีแบบจำลองใดเหมาะสมที่สุดกับทุกรายการ

จากตารางที่ 5 ผลการทดสอบความเสถียรของแบบจำลอง (สนับสนุนวัตถุประสงค์ที่ 1 และ 2) พบว่า ค่า SD ของ MAE รายรอบของ ML ทุกประเภทอยู่ในช่วง 0.006-0.016 สะท้อนความเสถียรสูงตลอดช่วงเวลาทดสอบ โดย Gradient

Boosting มี SD ต่ำสุด (0.006) ขณะที่ ARIMA มี SD สูงสุด (0.071) บ่งชี้ความไม่เสถียรในแต่ละรอบ ยืนยันว่าความเหนือกว่าของ ML เป็นผลสม่ำเสมอไม่ใช่ผลจากรอบใดรอบหนึ่ง

จากตารางที่ 6 ผลการกำหนดระดับสต็อกที่เหมาะสม (วัตถุประสงค์ที่ 3 — H3) พบว่า ระดับสต็อกที่เหมาะสมที่ Service Level ร้อยละ 99 รวมทุกรายการอยู่ระหว่าง Min 36 ชิ้น ถึง Max 59 ชิ้น ค่าเฉลี่ย 47.5 ชิ้นต่อเดือน ลดลงจากยอดนับคลังจริง ณ วันที่ 31 มีนาคม 2569 จำนวน 143 ชิ้น คิดเป็นการลดลงร้อยละ 66.78 มูลค่าสต็อกที่ลดลง 9.81 ล้านบาท และประหยัดต้นทุนการถือครองสินค้าคงคลังประมาณ 971,000 บาทต่อปี (อัตรา 10.1% ต่อปี)

ตารางที่ 1 ผลการทดสอบความนิ่งของข้อมูลด้วย ADF Test และ KPSS Test

รายการ	ADF Statistic	ADF p-value	KPSS Statistic	KPSS p-value	PP p-value	ผล
Porcine_E_A	-4.821	.0001**	0.213	>.10	-	ผ่าน
Porcine_E_M	-3.987	.0021**	0.342	>.10	-	ผ่าน
Porcine_O_H	-4.215	.0008**	0.287	>.10	-	ผ่าน
Bovine_P_A	-3.654	.0049**	0.391	>.10	-	ผ่าน
Bovine_P_M	-4.502	.0002**	0.245	>.10	-	ผ่าน
Bovine_O_M	-5.123	.0000**	0.198	>.10	-	ผ่าน
Bovine_O_A	-3.421	.0102*	0.412	.073	.018*	ผ่าน
Bovine_O_T	-3.876	.0028**	0.356	>.10	-	ผ่าน
Bovine_O_InsA	-4.298	.0005**	0.267	>.10	-	ผ่าน

** p < .01, * p < .05; = Stationary; PP Test รายงานเฉพาะรายการ Borderline

ตารางที่ 2 ผลการเปรียบเทียบประสิทธิภาพเฉลี่ยของแบบจำลองทั้งหมด

แบบจำลอง	Avg MAE	Avg RMSE	Avg MASE	DM Test (vs. Naive)
ARIMA	**0.921**	**1.204**	**0.892**	p=.142 (ns)
Naive Forecast	0.825	1.082	1.000	— (Baseline)
Croston's Method	0.612	0.798	0.687	p < .05*
SBA	0.587	0.756	0.658	p < .01**
Ridge Regression	0.401	0.488	0.236	p < .001***
Lasso Regression	0.467	0.562	0.298	p < .01**
Random Forest	0.512	0.643	0.341	p < .01**
Gradient Boosting	**0.354**	**0.435**	**0.218**	p < .001***
XG-Boost	0.389	0.471	0.256	p < .05*

*** p < .001, ** p < .01, * p < .05, ns = ไม่มีนัยสำคัญทางสถิติ; DM Test คำนวณจาก Pooled Forecast Errors รวมทุก Fold (n = 81); MASE < 1 = ดีกว่า Naive Forecast; MAPE ของ Croston's และ SBA ไม่รายงานเนื่องจากไม่นิยามเมื่อ Yt = 0; **แถวตัวหนา** = ประสิทธิภาพต่ำสุด (ARIMA) และสูงสุด (Gradient Boosting)

ตารางที่ 3 ผลการทดสอบ Diebold-Mariano Test เปรียบเทียบกับ ARIMA

คู่เปรียบเทียบ	DM Statistic	p-value	ผลการตัดสินใจ
Ridge Regression vs. ARIMA	-4.892	< .001***	ปฏิเสธ H_0
Lasso Regression vs. ARIMA	-4.213	< .001***	ปฏิเสธ H_0
Random Forest vs. ARIMA	-3.567	< .01**	ปฏิเสธ H_0
Gradient Boosting vs. ARIMA	-5.421	< .001***	ปฏิเสธ H_0
XG-Boost vs. ARIMA	-4.788	< .001***	ปฏิเสธ H_0
Croston's Method vs. ARIMA	-2.341	< .05*	ปฏิเสธ H_0
SBA vs. ARIMA	-2.687	< .01**	ปฏิเสธ H_0

*** p < .001, ** p < .01, * p < .05; DM Statistic ติดลบ = แบบจำลองในแถวมีความคลาดเคลื่อนต่ำกว่า ARIMA; ทดสอบแบบ One-sided ที่ระดับนัยสำคัญ .05; n = 81

ตารางที่ 4 Best Model แยกตามรายการลิ้นหัวใจเทียมชนิดเนื้อเยื่อ

รายการวัสดุ	Demand Profile	Best Model	MAE	MASE
Porcine_E_A	Intermittent	Ridge Regression	0.182	0.198
Porcine_E_M	Erratic	Gradient Boosting	0.387	0.245
Porcine_O_H	Erratic	Ridge Regression	0.421	0.267
Bovine_P_A	Smooth-Lumpy	Ridge Regression	0.654	0.298
Bovine_P_M	Smooth	XG-Boost	0.521	0.314
Bovine_O_M	Intermittent	Ridge Regression	0.098	0.156
Bovine_O_A	Smooth-Lumpy	Gradient Boosting	0.578	0.287
Bovine_O_T	Intermittent	Ridge Regression	0.165	0.187
Bovine_O_InsA	Intermittent	Ridge Regression	0.142	0.165

Demand Profile จำแนกตามแนวคิด Syntetos-Boylan-Croston (SBC) Classification โดยใช้เกณฑ์ ADI = 1.32 และ CV^2 = 0.49 ตาม Syntetos et al. (2005) ซึ่งได้รับการยืนยันโดย Petropoulos et al. (2022) ว่าเหมาะสมสำหรับข้อมูลรายเดือน; ADI = จำนวนงวดเฉลี่ยระหว่างการเบิกที่ไม่เป็นศูนย์; CV^2 = กำลังสองของสัมประสิทธิ์การแปรผัน คำนวณจากขนาดอุปสงค์ที่ไม่เป็นศูนย์เท่านั้น; เกณฑ์คัดเลือก = MAE และ MASE ต่ำสุดในแต่ละรายการ

ตารางที่ 5 ค่า MAE รายรอบ Cross-Validation ของแบบจำลองหลัก

แบบจำลอง	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	SD
ARIMA	0.875	1.012	0.945	0.834	0.939	0.921	**0.071**
Naive Forecast	0.815	0.832	0.830	0.819	0.829	0.825	0.008
Croston's Method	0.598	0.647	0.621	0.578	0.616	0.612	0.026
SBA	0.572	0.610	0.595	0.561	0.597	0.587	0.020

แบบจำลอง	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	SD
Ridge Regression	0.398	0.421	0.408	0.385	0.393	0.401	0.014
Lasso Regression	0.475	0.482	0.461	0.452	0.465	0.467	0.012
Random Forest	0.524	0.534	0.498	0.501	0.503	0.512	0.016
Gradient Boosting	0.348	0.362	0.351	0.358	0.351	0.354	**0.006**
XG-Boost	0.392	0.405	0.378	0.387	0.383	0.389	0.011

SD = Standard Deviation ของ MAE รายรอบ; **แถวตัวหนา** = ARIMA (SD สูงสุด) และ Gradient Boosting (SD ต่ำสุด)

ตารางที่ 6 ระดับสต็อกที่เหมาะสม (Safety Stock, Reorder Point, Min-Max) ที่ Service Level 99%

รายการวัสดุ	$\bar{D}$ (ชิ้น/เดือน)	σD	SS	ROP	Min (ชิ้น)	Max (ชิ้น)	Avg (ชิ้น)
Porcine_E_A	0.24	0.06	1.76	2.0	2	3	2.5
Porcine_E_M	1.73	0.16	1.07	2.80	3	5	4.0
Porcine_O_H	1.63	0.40	1.33	2.96	3	5	4.0
Bovine_P_A	4.83	0.99	3.64	8.47	9	14	11.5
Bovine_P_M	4.03	0.39	2.52	6.55	7	12	9.5
Bovine_O_M	0.07	0.06	0.93	1.0	1	2	1.5
Bovine_O_A	4.20	0.57	2.78	6.98	7	12	9.5
Bovine_O_T	0.20	0.05	1.80	2.0	2	3	2.5
Bovine_O_InsA	0.20	0.03	1.80	2.0	2	3	2.5
รวม	**17.13**	—	**17.63**	**37.76**	**36**	**59**	**47.5**

$\bar{D}$ = อุปสงค์เฉลี่ยรายเดือน; σD = ส่วนเบี่ยงเบนมาตรฐานของอุปสงค์; วิธีคำนวณ Hybrid: รายการที่ $\bar{D} \geq 1$ ใช้ Normal Approximation: $SS = z \times \sqrt{(L \times \sigma D^2 + \bar{D}^2 \times \sigma L^2)}$ โดย $z = 2.33$, $L = 1$, $\sigma L = 0.25$; รายการที่ $\bar{D} < 1$ ใช้ Poisson Distribution: $\lambda = \bar{D} \times L$ และ $ROP = k$ น้อยสุดที่ $P(X \leq k) \geq 0.99$ โดย $SS = k - \lambda$ (Axsäter, 2015; Syntetos et al., 2023); Min = $\lceil ROP \rceil$; Max = Min + $\lceil \bar{D} \rceil$; Avg = (Min + Max) / 2

สรุปและอภิปรายผลการวิจัย

สรุปผลการวิจัย

การวิจัยครั้งนี้บรรลุวัตถุประสงค์ทั้ง 3 ประการ ดังนี้ วัตถุประสงค์ที่ 1 ผลการเปรียบเทียบประสิทธิภาพยืนยัน H1 ว่าแบบจำลอง ML ทุกประเภทให้ผลการพยากรณ์ดีกว่า ARIMA อย่างมีนัยสำคัญทางสถิติ (Pairwise DM Test, $p < .05$) โดย Gradient Boosting มีค่า MAE ต่ำสุด (0.354) และ SD รายรอบต่ำสุด (0.006) สะท้อนทั้งความแม่นยำและความเสถียร ขณะที่ ARIMA มีค่า MAE สูงสุด (0.921) และไม่สามารถพยากรณ์ได้ดีกว่า Naive Forecast อย่างมีนัยสำคัญ ($p = .142$) วัตถุประสงค์ที่ 2 ยืนยัน H2 ว่า Ridge Regression เป็น Best Model สำหรับ 6 รายการ (ร้อยละ 66.67) Gradient Boosting 2 รายการ (ร้อยละ 22.22) และ XG-Boost 1 รายการ (ร้อยละ 11.11) ยืนยันว่าไม่มีแบบจำลองใดเหมาะสมที่สุดกับทุกรายการวัสดุ วัตถุประสงค์ที่ 3 ยืนยัน H3 ว่าระบบที่พัฒนาขึ้นสามารถกำหนดระดับสต็อกที่

เหมาะสมต่ำกว่าปัจจุบัน โดยระดับสต็อกเฉลี่ยที่แนะนำ 47.5 ชิ้น (Min 36 / Max 59 ชิ้น) ลดลงจากยอดนับจริง 143 ชิ้น คิดเป็นร้อยละ 66.78 มูลค่าสต็อกที่ลดลง 9.81 ล้านบาท และประหยัดต้นทุนการถือครองสินค้าคงคลังประมาณ 971,000 บาทต่อปี

อภิปรายผลการวิจัย

อภิปรายวัตถุประสงค์ที่ 1 ความเหนือกว่าของ ML เหนือ ARIMA ในบริบทอุปสงค์ไม่ต่อเนื่อง ML ทุกประเภทให้ผลการพยากรณ์ดีกว่า ARIMA อย่างมีนัยสำคัญ สอดคล้องกับ Alim et al. (2020) และ Thakur et al. (2024) โดย Gradient Boosting ลดค่า MAE ได้ร้อยละ 61.6 เนื่องจากสามารถจับความสัมพันธ์แบบไม่เป็นเชิงเส้นซึ่งเป็นข้อจำกัดเชิงโครงสร้างของ ARIMA (Wang et al., 2024) ข้อค้นพบสำคัญคือ ARIMA ไม่สามารถพยากรณ์ได้ดีกว่า Naive Forecast ($p = .142$) บ่งชี้ว่าไม่สามารถสกัดสัญญาณจากข้อมูล zero-inflated intermittent ได้ (Petropoulos et al., 2022) ความน่าเชื่อถือของข้อสรุปได้รับการยืนยันจากสองหลักฐาน ประการแรก การเปรียบเทียบเอื้อประโยชน์ต่อ ARIMA มากกว่า ML เนื่องจาก ARIMA ใช้ข้อมูลครบ 36 เดือน ขณะที่ ML ใช้เฉพาะช่วงการฝึกตาม Walk-Forward Cross-Validation การที่ ML ยังเหนือกว่าภายใต้เงื่อนไขนี้จึงยิ่งตอกย้ำความแข็งแกร่งของผลการวิจัย (Coroneo & Iacone, 2020) ประการที่สอง ค่า SD ของ MAE รายรอบของ ML ทุกประเภทอยู่ในช่วง 0.006-0.016 บ่งชี้ว่าความเหนือกว่าเป็นผลสม่ำเสมอตลอดทุก Fold ไม่ใช่ผลจากชุดข้อมูลใดชุดหนึ่ง การเลือก MASE เป็นตัวชี้วัดหลัก เป็นไปตาม Hyndman and Koehler (2006) เนื่องจากข้อมูล zero-inflated มี baseline variance ต่ำ ทำให้ R^2 สะท้อนประสิทธิภาพของแบบจำลองได้อย่างไม่เป็นธรรม นอกจากนี้ Croston's Method และ SBA ที่ออกแบบมาเฉพาะสำหรับอุปสงค์ไม่ต่อเนื่องก็ยังด้อยกว่า ML ทุกประเภท ชี้ให้เห็นว่า Feature Engineering ที่บูรณาการตัวแปรเชิงบริบทเป็นปัจจัยสำคัญไม่ใช่เพียงการเลือก Algorithm (Petropoulos et al., 2022)

อภิปรายวัตถุประสงค์ที่ 2 ความเหมาะสมของการคัดเลือกแบบจำลองแบบรายรายการวัสดุ Demand Profile ของวัสดุแต่ละรายการเป็นตัวกำหนดสำคัญในการเลือกแบบจำลอง Ridge Regression เป็น Best Model สำหรับ 6 รายการ Intermittent เนื่องจาก L2 Penalty ป้องกัน Overfitting ในชุดข้อมูลขนาดเล็กได้อย่างมีประสิทธิภาพ (Hoerl & Kennard, 1970) ขณะที่ Gradient Boosting และ XG-Boost เหมาะกับรายการ Erratic และ Smooth-Lumpy ที่มีความผันแปรสูงกว่า (Chen & Guestrin, 2016) ผลนี้ยืนยัน Bentéjac et al. (2021) ว่าไม่มีแบบจำลองใดเหนือกว่าทุกชุดข้อมูล และสนับสนุน Patharkar (2024) ที่เสนอ Per-SKU Selection สำหรับเวชภัณฑ์ฝังในร่างกายมูลค่าสูง

อภิปรายวัตถุประสงค์ที่ 3 ผลกระทบเชิงเศรษฐศาสตร์และนัยเชิงนโยบาย ระบบที่พัฒนาขึ้นลดสต็อกได้ร้อยละ 66.78 จาก 143 ชิ้น เหลือ 47.5 ชิ้น โดยคงรักษา Service Level ร้อยละ 99 มูลค่าสต็อกลดลง 9.81 ล้านบาท และประหยัด Holding Cost ประมาณ 971,000 บาทต่อปี คำนวณจาก Holding Cost Rate จริงร้อยละ 10.1 ต่อปี ซึ่งต่ำกว่าอัตราสมมติฐานทั่วไปร้อยละ 15 แสดงว่าผลประหยัดที่รายงานเป็น Conservative Estimate ที่น่าเชื่อถือ การปรับระดับสต็อกให้สอดคล้องกับปริมาณที่พยากรณ์ได้ยังช่วยลดความเสี่ยงจากวัสดุหมดอายุและลด Bullwhip Effect ในระบบ Consignment (Landry & Beaulieu, 2010) งานวิจัยนี้จึงนำเสนอกรอบ Best Model Selection เป็นต้นแบบเชิงประจักษ์รายแรกของประเทศไทยสำหรับเวชภัณฑ์ฝังในร่างกายมูลค่าสูงประเภทอุปสงค์ไม่ต่อเนื่อง ซึ่งสามารถขยายผลไปยังเวชภัณฑ์ชนิดอื่น เช่น Mechanical Valve, Pacemaker และ Knee Implant

ข้อจำกัดงานวิจัย

แม้ว่างานวิจัยครั้งนี้จะใช้ข้อมูลอนุกรมเวลา 36 จุด โดยมีชุดฝึกสอนในรอบแรกเพียง 24 เดือน แต่ในวิธีการวิจัยได้ผ่านการรับรองกลไกหลายประการ ในส่วนของแบบจำลอง ARIMA ที่สร้างผ่านโมดูล Expert Modeler ของโปรแกรม SPSS Statistics เวอร์ชัน 26 นั้น การคัดเลือกพารามิเตอร์ (p, d, q) ดำเนินการโดยอัตโนมัติตามเกณฑ์ AIC และ BIC ต่ำสุด พร้อมยืนยัน White Noise ด้วย Ljung-Box Q-Test กระบวนการดังกล่าวช่วยป้องกันการกำหนดพารามิเตอร์ตามดุลยพินิจของผู้วิจัยและลดความเสี่ยงต่อภาวะการเรียนรู้จำ (Overfitting) ที่อาจเกิดขึ้นจากการประมวลผลด้วยข้อมูลขนาด

เล็กได้อย่างมีประสิทธิภาพ สำหรับแบบจำลอง ML ผู้วิจัยใช้ L1/L2 Regularization และการปรับแต่งพารามิเตอร์ด้วย Grid Search ภายในแต่ละ Fold เพื่อควบคุม Overfitting เพิ่มเติม (Hoerl & Kennard, 1970)

ข้อเสนอแนะเชิงปฏิบัติ

ศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือควรนำระบบ Best Model Selection ไปบูรณาการกับระบบสารสนเทศคลังเวชภัณฑ์ที่มีอยู่ โดยกำหนดระดับสต็อก Min-Max ตามตารางที่ 6 ปรับข้อตกลง Consignment กับผู้จัดจำหน่ายให้สอดคล้องกับปริมาณที่พยากรณ์ได้ และ Retrain แบบจำลองทุก 6 เดือน พร้อมติดตาม Inventory Turnover, Stockout Rate และ Expiry Loss เป็นตัวชี้วัดประจำปี โรงพยาบาลอื่นสามารถนำกรอบนี้ไปปรับใช้กับเวชภัณฑ์ฝังในร่างกายชนิดอื่น เช่น Mechanical Valve, Pacemaker และ Knee Implant โดยเริ่มจากการจำแนกประเภทอุปสงค์ตามเกณฑ์ SBC Classification ก่อนเลือกแบบจำลองที่เหมาะสม

ข้อเสนอแนะในการวิจัยครั้งต่อไป

การวิจัยในอนาคตควรขยายขนาดข้อมูลเป็น 60 เดือนขึ้นไปและพัฒนาเป็น Multi-Center Study เพื่อเพิ่ม External Validity นอกจากนี้ควรบูรณาการตัวแปรภายนอก เช่น Surgical Backlog ผ่าน ARIMAX หรือ Multivariate ML ร่วมกับ Dynamic Lead Time Distribution และ Activity-Based Costing เพื่อประเมินต้นทุนที่แม่นยำยิ่งขึ้น

เอกสารอ้างอิง

ศูนย์หัวใจสิริกิติ์ภาคตะวันออกเฉียงเหนือ. (2568). *รายงานประจำปี 2567*. ขอนแก่น: มหาวิทยาลัยขอนแก่น.

สมาคมศัลยแพทย์ทรวงอกแห่งประเทศไทย. (2566). *รายงานสถิติการผ่าตัดโรคหัวใจประจำปี 2566*. กรุงเทพฯ: สมาคมศัลยแพทย์ทรวงอกแห่งประเทศไทย.

Alim, M., Ye, G. H., Guan, P., Huang, D. S., Zhou, B. S., & Wu, W. (2020). Comparison of ARIMA model and XG Boost model for prediction of human brucellosis in mainland China: A time-series study. *BMJ Open, 10*(12), e039676. https://doi.org/10.1136/bmjopen-2020-039676

Axsäter, S. (2015). *Inventory control* (3rd ed.). Cham, Switzerland: Springer.

Babai, M. Z., Boylan, J. E., & Rostami-Tabar, B. (2022). Demand forecasting in supply chains: A review of aggregation and hierarchical approaches. *International Journal of Production Research, 60*(1), 324-348. https://doi.org/10.1080/00207543.2021.1977858

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review, 54*(3), 1937-1967. https://doi.org/10.1007/s10462-020-09896-5

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). Hoboken, NJ: Wiley.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Chen, T., & Guestrin, C. (2016). XG Boost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). New York, NY: ACM. https://doi.org/10.1145/2939672.2939785

Coroneo, L., & Iacone, F. (2020). Comparing predictive accuracy in small samples using fixed-smoothing asymptotics. *Journal of Applied Econometrics, 35*(4), 391-409. https://doi.org/10.1002/jae.2754

Fang, T., Lahdelma, R., & Salminen, P. (2022). Demand forecasting for hospital pharmaceuticals: A comparative study of statistical and machine learning models. *Health Care Management Science, 25*(3), 419-435. https://doi.org/10.1007/s10729-022-09597-3

GBD 2021 Valvular Heart Disease Collaborators. (2023). Global, regional, and national burden of valvular heart disease, 1990-2019. *Lancet, 401*(10371), 147-164. https://doi.org/10.1016/S0140-6736(22)01693-1

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55-67. https://doi.org/10.1080/00401706.1970.10488634

Kritchanchai, D., & Meesamut, W. (2015). Developing inventory management in hospital. *International Journal of Supply Chain Management, 4*(2), 11-19.

Landry, S., & Beaulieu, M. (2010). Achieving lean healthcare by combining the two-bin Kanban replenishment system with RFID technology. *International Journal of Health Management and Information, 1*(1), 85-98.

Otto, C. M., Nishimura, R. A., Bonow, R. O., Carabello, B. A., Erwin, J. P., Gentile, F., ... & Toly, C. (2021). 2020 ACC/AHA guideline for the management of patients with valvular heart disease. *Journal of the American College of Cardiology, 77*(4), e25-e197. https://doi.org/10.1016/j.jacc.2020.11.018

Patharkar, A. (2024). Machine learning approaches for surgical implant demand forecasting in tertiary care hospitals. *Operations Research for Health Care, 42*, 100437. https://doi.org/10.1016/j.orhc.2024.100437

Petropoulos, F., Apiletti, D., Assimakopoulos, V., Babai, M. Z., Barrow, D. K., Taieb, S. B., ... & Ziel, F. (2022). Forecasting: Theory and practice. *International Journal of Forecasting, 38*(3), 705-871. https://doi.org/10.1016/j.ijforecast.2021.11.001

Rezvova, M. A., Klyshnikov, K. Y., Rezvov, A. A., & Ovcharenko, E. A. (2023). Polymer-based heart valve prostheses: A comprehensive review. *Polymers, 15*(3), 562. https://doi.org/10.3390/polym15030562

Syntetos, A. A., & Boylan, J. E. (2005). The accuracy of intermittent demand estimates. *International Journal of Forecasting, 21*(2), 303-314. https://doi.org/10.1016/j.ijforecast.2004.10.001

Syntetos, A. A., Babai, M. Z., & Gardner, E. S. (2023). Forecasting intermittent inventory demands: Simple parametric methods vs. bootstrapping. *Journal of Business Research, 156*, 113493. https://doi.org/10.1016/j.jbusres.2022.113493

Teunter, R. H., Syntetos, A. A., & Babai, M. Z. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research, 214*(3), 606-615. https://doi.org/10.1016/j.ejor.2011.05.018

Thakur, V., Anbanandam, R., & Mehta, R. (2024). Hospital supply chain management during COVID-19 using machine learning approach. *Annals of Operations Research, 335*(1), 487-513. https://doi.org/10.1007/s10479-022-04921-8

Wang, Y., Zhang, L., & Chen, H. (2024). A hybrid LSTM-ARIMA model for medical device demand forecasting. *Expert Systems with Applications, 238*, 121812. https://doi.org/10.1016/j.eswa.2023.121812

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



Copyright: © 2026 by the authors. This is a fully open-access article distributed under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).