Received: 8 มี.ค. 2569 **Revised:** 22 เม.ย. 2569 **Accepted:** 24 เม.ย. 2569

การวิเคราะห์เปรียบเทียบแบบจำลองการเรียนรู้ของเครื่องแบบรวมกลุ่มเพื่อการพยากรณ์เชิงพหุคูณมูลค่าการส่งออกยางพาราของประเทศไทยโดยอาศัยปัจจัยทางเศรษฐกิจ

Comparative Analysis of Ensemble Machine Learning Models for Multivariate Forecasting of Thailand's Natural Rubber Export Value Using Economic Factors

ฑิติยาพร ดรนาม[1], ณิชารีย์ สุโพธิ์แสน[1], เลอศักดิ์ โพธิ์ทอง[1] และ พรภิยา ม่วงเขียว[1]*
[1]สาขาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

Thitiyaporn Dornnam[1], Nicharee Suphosaen[1], Lersak Phothong[1] and Phonphiya Muangkeaw[1]*
[1]Department of Business Computer, Mahasarakham Business School, Mahasarakham University
*ผู้นิพนธ์ประสานงาน: jintaporn.m@mbs.msu.ac.th

Abstract

Natural rubber is a key agricultural commodity in Thailand, with export values subject to significant volatility driven by global prices, exchange rates, crude oil prices, and macroeconomic conditions. Despite extensive studies on commodity forecasting, prior research often relies on traditional econometric models or single machine learning techniques, with limited integration of multi-factor economic variables and ensemble learning approaches for improving predictive robustness. This study aims to develop and compare machine learning models for forecasting Thailand's natural rubber export value using monthly secondary data from 2012 to 2022 (132 observations, 14 variables). The analysis follows the CRISP-DM framework and employs four techniques: k-Nearest Neighbors (k-NN), Random Forest (RF), Neural Network (NN), and a Voting Ensemble model. The dataset is split into 70% for training and 30% for testing, and model performance is evaluated using MSE, RMSE, MAE, MAPE, and R^2. The results show that the Voting Ensemble model outperforms other models, achieving the lowest prediction errors (MSE = 3,723,671.862; RMSE = 1,929.6818; MAE = 1,502.4775; MAPE = 0.1083) and the highest R^2 (0.8259). This study contributes by demonstrating the effectiveness of ensemble learning in integrating heterogeneous

economic indicators to enhance forecasting accuracy. The findings provide a robust data-driven framework to support strategic decision-making in export planning, production management, and policy formulation for Thailand's natural rubber industry.

Keywords: *natural rubber export; machine learning; economic factors; forecasting; ensemble learning*

บทคัดย่อ

ยางพาราเป็นสินค้าเกษตรสำคัญของประเทศไทย โดยมูลค่าการส่งออกมีความผันผวนตามปัจจัยทางเศรษฐกิจหลายด้าน เช่น ราคายางในตลาดโลก อัตราแลกเปลี่ยน ราคาน้ำมันดิบ และภาวะเศรษฐกิจมหภาค อย่างไรก็ตาม งานวิจัยที่ผ่านมาในด้านการพยากรณ์สินค้าโภคภัณฑ์มักใช้แบบจำลองเศรษฐมิติแบบดั้งเดิมหรือเทคนิคการเรียนรู้ของเครื่องเพียงวิธีเดียว และยังขาดการบูรณาการตัวแปรเศรษฐกิจหลายมิติร่วมกับเทคนิคแบบรวมกลุ่มเพื่อเพิ่มความแม่นยำในการพยากรณ์ งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบแบบจำลองการเรียนรู้ของเครื่องสำหรับพยากรณ์มูลค่าการส่งออกยางพาราของประเทศไทย โดยใช้ข้อมูลรายเดือนช่วงปี พ.ศ. 2555–2565 จำนวน 132 ตัวอย่าง และ 14 ตัวแปร ดำเนินการตามกรอบ CRISP-DM และประยุกต์ใช้เทคนิค k-Nearest Neighbors, Random Forest, Neural Network และ Voting Ensemble โดยแบ่งข้อมูลเป็นชุดฝึก 70% และชุดทดสอบ 30% และประเมินผลด้วย MSE, RMSE, MAE, MAPE และ R^2 ผลการศึกษาพบว่าแบบจำลอง Voting Ensemble มีประสิทธิภาพสูงสุด โดยให้ค่าความคลาดเคลื่อนต่ำที่สุด (MSE = 3,723,671.862; RMSE = 1,929.6818; MAE = 1,502.4775; MAPE = 0.1083) และค่า R^2 สูงสุด (0.8259) งานวิจัยนี้แสดงให้เห็นถึงศักยภาพของเทคนิคแบบรวมกลุ่มในการบูรณาการปัจจัยเศรษฐกิจหลายมิติเพื่อเพิ่มความแม่นยำในการพยากรณ์ และสามารถใช้เป็นแนวทางสนับสนุนการตัดสินใจด้านการส่งออก การวางแผนการผลิต และการกำหนดนโยบายในอุตสาหกรรมยางพาราของประเทศไทย

คำสำคัญ: *การส่งออกยางพารา; การเรียนรู้ของเครื่อง; ปัจจัยทางเศรษฐกิจ; การพยากรณ์; การเรียนรู้แบบรวมกลุ่ม*

บทนำ

ยางพาราเป็นพืชเศรษฐกิจที่มีบทบาทสำคัญต่อระบบเศรษฐกิจของประเทศไทยมาอย่างยาวนาน เนื่องจากประเทศไทยมีสภาพภูมิอากาศที่เหมาะสมต่อการเพาะปลูก ทำให้สามารถผลิตและส่งออกยางพาราได้ในปริมาณมาก และเป็นหนึ่งในประเทศผู้ผลิตและผู้ส่งออกยางธรรมชาติรายสำคัญของโลก โดยมีปริมาณการส่งออกเฉลี่ยประมาณ 270,000 ตันต่อเดือน ซึ่งสะท้อนถึงความสำคัญของยางพาราในฐานะแหล่งรายได้หลักของเกษตรกร และเป็นวัตถุดิบสำคัญของอุตสาหกรรมต่อเนื่องทั้งในประเทศและต่างประเทศ (ไชยวัฒน์ เสาวเจริญสุข, 2565)

ในขณะเดียวกัน งานวิจัยสมัยใหม่ให้ความสำคัญกับการพัฒนาแบบจำลอง Machine Learning และ Ensemble Learning เช่น XGBoost และ LightGBM ซึ่งสามารถรวมจุดเด่นของหลายอัลกอริทึมเพื่อเพิ่มความแม่นยำของการพยากรณ์ (Zhang, 2023) โดยเฉพาะในบริบทของข้อมูลเศรษฐกิจและสินค้าโภคภัณฑ์ที่ได้รับอิทธิพลจากหลายปัจจัย รวมถึงสถานการณ์ภูมิรัฐศาสตร์และการแข่งขันจากประเทศผู้ผลิตรายอื่น ซึ่งส่งผลต่อความสามารถในการแข่งขันของอุตสาหกรรมยางพาราในตลาดโลก (ไชยวัฒน์ เสาวเจริญสุข, 2565)

จากสถานการณ์ดังกล่าว การพยากรณ์มูลค่าการส่งออกยางพาราจึงมีความสำคัญ เนื่องจากสามารถใช้สนับสนุนการวางแผนการผลิต การบริหารการส่งออก และการกำหนดนโยบายได้อย่างมีประสิทธิภาพ ในอดีตมีการใช้วิธีทางสถิติ เช่น ค่าเฉลี่ยเคลื่อนที่ การปรับเรียบแบบเอ็กซ์โปเนนเชียล และวิธีบ็อกซ์–เจนกินส์ (Box–Jenkins Method) รวมถึงการวิเคราะห์การถดถอยเชิงซ้อน (Multiple Regression Analysis) ในการศึกษาปัจจัยที่ส่งผลต่อการส่งออก (จารุวรรณ สิงห์ม่วง และ ธิดาพร ศุภภากร, 2563; อภิสิทธิ์ หะยีอุมา และ ธนวิทย์ บุญสิทธิ์, 2563)

ในระดับสากล งานวิจัยจำนวนมากมุ่งเน้นการพยากรณ์ราคาหรือแนวโน้มสินค้าเกษตร โดยใช้ทั้งแบบจำลองทางสถิติและ Machine Learning เช่น แบบจำลอง ARIMA (Cherdchoongam & Rungreunganun, 2016) และแบบจำลอง Machine Learning ที่ใช้ตัวชี้วัดตลาดสินค้าโภคภัณฑ์ (Nyondo & Varghese, 2024) ซึ่งให้ความแม่นยำสูงกว่าวิธีดั้งเดิม นอกจากนี้ แนวโน้มงานวิจัยยังชี้ให้เห็นถึงการเพิ่มขึ้นของ Hybrid Models และ Ensemble Models เพื่อยกระดับประสิทธิภาพของแบบจำลอง (Guindani et al., 2024; Makridakis et al., 2022)

อย่างไรก็ตาม งานวิจัยส่วนใหญ่ยังมุ่งเน้นการพยากรณ์ราคา หรือปริมาณการผลิต ขณะที่การพยากรณ์มูลค่าการส่งออก ซึ่งเป็นตัวแปรเชิงประกอบที่สะท้อนปัจจัยหลายมิติยังมีจำกัด โดยเฉพาะการประยุกต์ใช้ Machine Learning แบบเปรียบเทียบหลายเทคนิคและการใช้ Ensemble Learning เพื่อเพิ่มความแม่นยำ

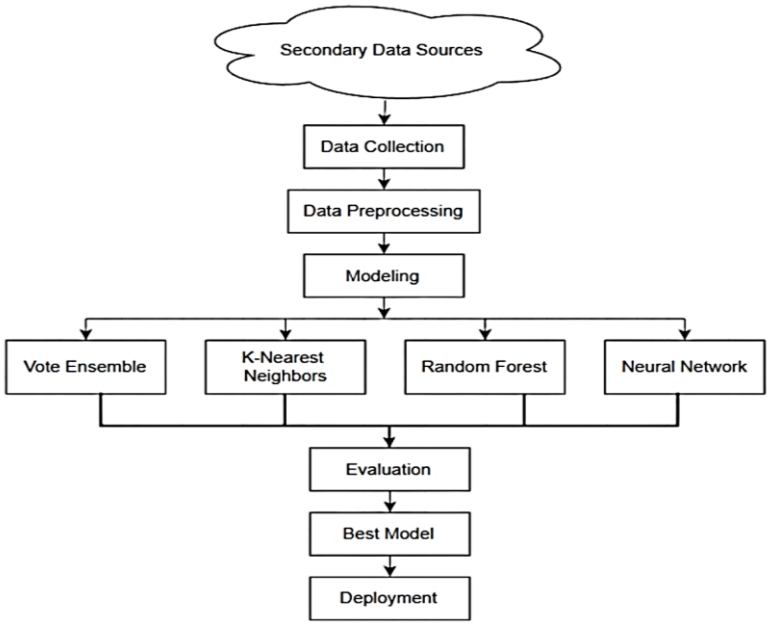
ดังนั้น งานวิจัยนี้จึงมุ่งศึกษาการพยากรณ์มูลค่าการส่งออกยางพาราของประเทศไทย ซึ่งมีความซับซ้อนกว่าการพยากรณ์ราคา เนื่องจากได้รับอิทธิพลจากปริมาณการส่งออก อุปสงค์ในตลาดโลก นโยบายการค้า อัตราแลกเปลี่ยน และต้นทุนโลจิสติกส์ รวมถึงความสัมพันธ์เชิงพหุคูณของตัวแปรเศรษฐกิจหลายมิติ โดยพิจารณาปัจจัยสำคัญ ได้แก่ ปริมาณการผลิต ปริมาณการส่งออก ปริมาณการใช้ในประเทศ อัตราแลกเปลี่ยน ราคาน้ำมัน ดัชนีราคาผู้บริโภค ราคายางในตลาดโลกและในประเทศ การเปลี่ยนแปลงราคาน้ำมัน เนื้อที่เก็บเกี่ยว และผลผลิตที่เก็บเกี่ยวได้ และประยุกต์ใช้เทคนิค Machine Learning จำนวน 4 วิธี ได้แก่ เทคนิคเพื่อนบ้านใกล้ที่สุด (k-Nearest Neighbors: k-NN) เทคนิคต้นไม้ป่าสุ่ม (Random Forest: RF) เทคนิคโครงข่ายประสาทเทียม (Neural Network: NN) และเทคนิควิธีการรวมกลุ่ม (Voting Ensemble: VE) เพื่อพัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลอง

ผลการศึกษาที่ได้สามารถใช้เป็นแนวทางในการสนับสนุนการตัดสินใจเชิงนโยบาย การวางแผนการผลิต และการกำหนดกลยุทธ์ทางเศรษฐกิจ เพื่อเสริมสร้างศักยภาพการแข่งขันของอุตสาหกรรมยางพาราของประเทศไทยในตลาดโลก

วิธีการทดลอง

การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบแบบจำลอง Machine Learning สำหรับการพยากรณ์มูลค่าการส่งออกยางพาราของประเทศไทย โดยใช้ข้อมูลทุติยภูมิรายเดือนช่วงปี พ.ศ. 2555–2565 จำนวน 132 ระเบียน และ 14 ตัวแปร ครอบคลุมด้านการผลิต การค้า และปัจจัยทางเศรษฐกิจ ดังแสดงในภาพที่ 1

ในการสร้างแบบจำลอง ผู้วิจัยประยุกต์ใช้เทคนิค 4 วิธี ได้แก่ k-NN, RF, NN และ VE โดยแบ่งข้อมูลเป็นชุดฝึก (training set) ร้อยละ 70 และชุดทดสอบ (testing set) ร้อยละ 30 ภายใต้ hold-out validation และกำหนดค่า random seed เท่ากับ 1992 เพื่อให้ผลการทดลองสามารถทำซ้ำได้



ภาพที่ 1 กรอบแนวคิดของขั้นตอนการวิจัย

1.การทำความเข้าใจเกี่ยวกับธุรกิจ (Business Understanding)

อุตสาหกรรมยางพารามีบทบาทสำคัญต่อเศรษฐกิจของประเทศไทย โดยมูลค่าการส่งออกมีความผันผวนจากปัจจัยทางเศรษฐกิจหลายด้าน เช่น อุปสงค์–อุปทานของตลาดโลก อัตราแลกเปลี่ยน และราคาน้ำมันดิบ ซึ่งส่งผลต่อรายได้ของภาคเกษตรและอุตสาหกรรมต่อเนื่อง (จิดาภา ลู่วิโรจน์, 2566) อีกทั้งลักษณะของปัญหาเป็นการพยากรณ์เชิงพหุคูณ (multivariate regression) ที่มีความสัมพันธ์ทั้งเชิงเส้นและไม่เป็นเชิงเส้น จึงเหมาะสมต่อการใช้เทคนิค Machine Learning เพื่อเพิ่มความแม่นยำในการพยากรณ์

2.การทำความเข้าใจเกี่ยวกับข้อมูล (Data Understanding)

ข้อมูลที่ใช้ในการวิจัยครั้งนี้เป็นข้อมูลทุติยภูมิที่รวบรวมจากแหล่งข้อมูลที่มีความน่าเชื่อถือหลายแหล่ง โดยประกอบด้วยข้อมูลด้านอุตสาหกรรมยางพาราของประเทศไทย ได้แก่ ปริมาณการผลิต ปริมาณการส่งออก และปริมาณการใช้ภายในประเทศ จากฐานข้อมูลสถิติยางพารา (กรมวิชาการเกษตร, n.d.; การยางแห่งประเทศไทย, n.d.) นอกจากนี้ยังรวมถึงตัวแปรทางเศรษฐกิจที่คาดว่าจะส่งผลต่อมูลค่าการส่งออก เช่น อัตราแลกเปลี่ยนเงินบาท ราคาน้ำมันดิบ ดัชนีราคาผู้บริโภค ราคายางในตลาดโลกและในประเทศ รวมถึงราคาทองคำ ซึ่งรวบรวมจากเว็บไซต์ Investing (Investing.com, n.d.-a) และข้อมูลด้านผลผลิตจากกรมส่งเสริมการเกษตร (กรมส่งเสริมการเกษตร, n.d.)

ชุดข้อมูลประกอบด้วยข้อมูลรายเดือนจำนวน 132 ระเบียน และตัวแปร 14 ตัวแปร ซึ่งครอบคลุมทั้งด้านการผลิต การค้า และปัจจัยเศรษฐกิจ เช่น อัตราแลกเปลี่ยน ราคาน้ำมัน ดัชนีราคาผู้บริโภค และราคายางในตลาดโลกและในประเทศ (อภิสิทธิ์ หะยีอุมา และ ธนวิทย์ บุญสิทธิ์, 2563)

แม้ว่าข้อมูลจะมีลักษณะเชิงเวลา แต่ในการศึกษานี้ไม่ได้ใช้เทคนิคอนุกรมเวลาโดยตรง โดยกำหนดตัวแปร Month/Year เป็นตัวระบุ (ID) และ Rubber_Export เป็นตัวแปรเป้าหมาย (Label)

3.การเตรียมข้อมูล (Data Preparation)

ข้อมูลได้รับการตรวจสอบความครบถ้วนและไม่พบค่าที่สูญหาย (missing values) จากนั้นดำเนินการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรด้วยเมทริกซ์สหสัมพันธ์ (Correlation Matrix) และกำจัดตัวแปรที่มีความสัมพันธ์สูง ($|r| \geq 0.7$) เพื่อลดปัญหาพหุสหสัมพันธ์ (multicollinearity) โดยผลจาก Correlation Matrix ถูกนำไปใช้ในขั้นตอนคัดเลือกตัวแปร ซึ่งตัวแปรที่มีความสัมพันธ์สูงถูกตัดออกด้วยคำสั่ง Remove Correlated Attributes ก่อนนำข้อมูลเข้าสู่ขั้นตอนการสร้างแบบจำลอง

ถัดมาได้ปรับมาตรฐานข้อมูลด้วยวิธี Z-score normalization เพื่อให้ตัวแปรมีค่าเฉลี่ย (mean) เท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐาน (standard deviation) เท่ากับ 1 โดยคำนวณจากสมการ

$$z = \frac{x-\mu}{\sigma} \qquad (1)$$

เมื่อ x คือค่าข้อมูลเดิม, μ คือค่าเฉลี่ย และ σ คือส่วนเบี่ยงเบนมาตรฐานของตัวแปร ตัวอย่างเช่น หากค่าราคาน้ำมันดิบของเดือนหนึ่งมีค่าสูงกว่าค่าเฉลี่ยของชุดข้อมูล ค่าที่ได้หลังการปรับมาตรฐานจะมีค่าเป็นบวก ซึ่งสะท้อนว่าข้อมูลดังกล่าวอยู่สูงกว่าค่าเฉลี่ยของชุดข้อมูล

การปรับมาตรฐานดังกล่าวช่วยลดผลกระทบจากความแตกต่างของช่วงข้อมูล (scale differences) และเพิ่มประสิทธิภาพของแบบจำลอง โดยเฉพาะ k-Nearest Neighbors (k-NN) และ Neural Network (NN) ทั้งนี้ การประมวลผลข้อมูลดำเนินการด้วยโปรแกรม Altair AI Studio (Educational Version 2025.0.1)

ชุดข้อมูลประกอบด้วยตัวแปรจำนวน 14 ตัวแปร ครอบคลุมด้านการผลิต การค้า และปัจจัยทางเศรษฐกิจ โดยรายละเอียดและบทบาทของตัวแปรแสดงในตารางที่ 1 ซึ่งกำหนดให้ Month/Year เป็นตัวระบุข้อมูล (ID) และ Rubber_Export เป็นตัวแปรเป้าหมาย (Label) ขณะที่ตัวแปรอื่นเป็นตัวแปรอิสระ (Input)

ตารางที่ 1 โครงสร้างตัวแปร แหล่งที่มา และบทบาทของข้อมูล

แหล่งที่มา	ชื่อตัวแปร	ชนิดข้อมูล	คำอธิบาย	บทบาท
กรมวิชาการเกษตรและการยางแห่งประเทศไทย (2025)	Month/Year	Date	เดือน/ปี	ID
	Rubber_Export	Real	มูลค่าการส่งออกยางพารา (ล้านบาท)	Label
	Production_Volume	Integer	ปริมาณการผลิต (ตัน)	Input
	Export_Volume	Integer	ปริมาณการส่งออก (ตัน)	Input
	Domestic_Consumption	Integer	ปริมาณการใช้ในประเทศ (ตัน)	Input
Investing.com (n.d.-e)	Exchange_Rate	Real	อัตราแลกเปลี่ยนเงินบาท/ดอลลาร์	Input
Investing.com (n.d.-b)	Oil_Price	Real	ราคาน้ำมันดิบ (ดอลลาร์)	Input
	Change_Oil	Real	การเปลี่ยนแปลงราคาน้ำมัน (%)	Input
สำนักงานนโยบายและยุทธศาสตร์การค้า (n.d.)	CPI	Real	ดัชนีราคาผู้บริโภค	Input
Investing.com (n.d.-c)	World_Rubber_Price	Real	ราคายางในตลาดโลก (USD/kg)	Input
Investing.com (n.d.-d)	Thai_Rubber_Price	Real	ราคายางในตลาดไทย (บาท/kg)	Input
Investing.com (n.d.-a)	Gold_Price	Real	ราคาทองคำ (ดอลลาร์)	Input
กรมส่งเสริมการเกษตร (n.d.)	Harvested_Area	Real	เนื้อที่เก็บเกี่ยว (ไร่)	Input
	Harvested_Output	Real	ผลผลิตที่เก็บเกี่ยว (กิโลกรัม)	Input

4. การสร้างแบบจำลอง (Modeling)

ภายหลังจากการเตรียมข้อมูล ผู้วิจัยได้นำชุดข้อมูลที่ผ่านการประมวลผลแล้วมาพัฒนาแบบจำลองการพยากรณ์เชิงประมาณค่า (Regression Models) โดยใช้เทคนิคการเรียนรู้ของเครื่อง

(Machine Learning) เพื่อให้สามารถเปรียบเทียบประสิทธิภาพของแบบจำลองได้อย่างเป็นระบบ ผู้วิจัยได้ใช้กระบวนการทำสำเนาชุดข้อมูล (Multiply) เพื่อสร้างชุดข้อมูลสำหรับการฝึกและทดสอบในรูปแบบเดียวกัน จากนั้นจึงพัฒนาแบบจำลองจำนวน 4 วิธี ได้แก่ k-Nearest Neighbors (k-NN), Random Forest (RF), Neural Network (NN) และ Voting Ensemble (VE)

การสร้างแบบจำลองดำเนินการด้วยโปรแกรม Altair AI Studio (Educational Version 2025.0.1) โดยใช้ชุดคำสั่งสำคัญ เช่น Correlation Matrix และ Remove Correlated Attributes สำหรับคัดเลือกตัวแปรเพื่อลดปัญหา multicollinearity, Normalize สำหรับปรับมาตราส่วนข้อมูล และ Split Data ร่วมกับ Cross Validation สำหรับแบ่งข้อมูลและประเมินแบบจำลอง นอกจากนี้ได้กำหนดค่า random seed เท่ากับ 1992 เพื่อให้ผลการทดลองสามารถทำซ้ำได้และลดความแปรปรวนของผลลัพธ์ (อนุพงศ์ สุขประเสริฐ, 2568)

4.1 เทคนิคเพื่อนบ้านใกล้ที่สุด (k-Nearest Neighbors: k-NN)

เทคนิค k-NN เป็นวิธีการพยากรณ์แบบไม่อาศัยพารามิเตอร์ (non-parametric) ซึ่งไม่ต้องกำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรล่วงหน้า โดยอาศัยหลักการค้นหาข้อมูลในชุดฝึกที่มีลักษณะใกล้เคียงกับข้อมูลเป้าหมายมากที่สุดจำนวน k ตัวอย่าง (nearest neighbors) และคำนวณค่าเฉลี่ยของตัวแปรตามเพื่อใช้เป็นค่าพยากรณ์

ในการศึกษานี้ ผู้วิจัยได้ทำการปรับแต่งค่าพารามิเตอร์ k ในช่วง 1–15 โดยใช้วิธีการค้นหาเชิงเส้น (Linear Search) และประเมินประสิทธิภาพของแบบจำลองด้วยวิธีการตรวจสอบไขว้ (Cross Validation) เพื่อเลือกค่าที่ให้ค่าความคลาดเคลื่อนต่ำที่สุด

เทคนิค k-NN มีความสามารถในการจัดการกับข้อมูลที่มีความสัมพันธ์ไม่เชิงเส้นได้ดี อย่างไรก็ตาม ประสิทธิภาพของแบบจำลองขึ้นอยู่กับการเลือกค่า k และวิธีการวัดระยะห่าง ซึ่งหากกำหนดไม่เหมาะสมอาจนำไปสู่ปัญหาการเรียนรู้มากเกินไป (overfitting) หรือการเรียนรู้น้อยเกินไป (underfitting) (Hastie, Tibshirani, & Friedman, 2009) โดยผลการทดลองพบว่า ค่า k = 3 เป็นค่าที่เหมาะสมที่สุด เนื่องจากให้ค่าความคลาดเคลื่อนของการพยากรณ์ต่ำที่สุด

4.2 เทคนิคต้นไม้ป่าสุ่ม (Random forest)

เทคนิค RF เป็นอัลกอริทึมการเรียนรู้แบบรวมกลุ่ม (Ensemble Learning) ที่พัฒนาจากวิธีแบกกิง (Bagging) โดยสร้างต้นไม้ตัดสินใจหลายต้นจากการสุ่มตัวอย่างข้อมูลแบบบูตสแตรป (Bootstrap Sampling) และการสุ่มเลือกตัวแปรในแต่ละโหนด (Random Feature Selection) จากนั้นรวมผลลัพธ์เพื่อเพิ่มความเสถียรของการพยากรณ์ โดยในปัญหาการประมาณค่า (Regression) จะใช้ค่าเฉลี่ยของผลลัพธ์จากต้นไม้ทั้งหมด (Breiman, 2001)

เทคนิค RF เหมาะสมกับข้อมูลเศรษฐกิจที่มีความสัมพันธ์ไม่เชิงเส้นและมีหลายตัวแปร เนื่องจากสามารถลดปัญหาการเรียนรู้มากเกินไป (overfitting) เมื่อเทียบกับต้นไม้ตัดสินใจเดี่ยว และช่วยเพิ่มความแม่นยำของแบบจำลอง (Liaw & Wiener, 2002)

ในการศึกษานี้ ผู้วิจัยได้ปรับแต่งค่าพารามิเตอร์สำคัญ โดยทดลองจำนวนต้นไม้ที่ 100, 200 และ 300 ต้น และพบว่า 200 ต้นให้ค่าความคลาดเคลื่อนต่ำที่สุด จึงเลือกใช้เป็นค่าที่เหมาะสม นอกจากนี้ได้กำหนดเกณฑ์การแบ่งโหนดแบบผลรวมกำลังสองน้อยที่สุด (Least Squares) และกำหนดความลึกสูงสุดของต้นไม้ (Maximal Depth) เท่ากับ 10 ระดับ เพื่อควบคุมความซับซ้อนและลดโอกาสการเกิด overfitting ซึ่งช่วยให้แบบจำลองสามารถเรียนรู้ความสัมพันธ์ของข้อมูลได้อย่างมีประสิทธิภาพและเพิ่มความแม่นยำในการพยากรณ์

4.3 เทคนิคโครงข่ายประสาทเทียม (Neural Network: NN)

เทคนิค NN เป็นวิธีการเรียนรู้ของเครื่อง (Machine Learning) ที่สามารถเรียนรู้ความสัมพันธ์เชิงซับซ้อนและไม่เป็นเชิงเส้นของข้อมูล โดยมีโครงสร้างเลียนแบบระบบประสาทมนุษย์ ประกอบด้วยโหนด (neurons) ที่เชื่อมโยงกันผ่านค่าน้ำหนัก (weights) และจัดเรียงเป็นชั้นต่าง ๆ ได้แก่ ชั้นนำเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และชั้นผลลัพธ์ (Output Layer) โดยใช้ฟังก์ชันกระตุ้น (activation function) ในการประมวลผลข้อมูล (อนุพงศ์ สุขประเสริฐ, 2568)

ในการศึกษานี้ ผู้วิจัยใช้โครงข่ายแบบหลายชั้น (Multilayer Perceptron: MLP) ซึ่งเป็นโครงข่ายแบบส่งต่อ (feedforward neural network) โดยกำหนดชั้นซ่อน (Hidden Layer) จำนวน 1 ชั้น และมีจำนวนโหนด 6 โหนด เพื่อควบคุมความซับซ้อนและลดความเสี่ยงของการเรียนรู้มากเกินไป (overfitting)

สำหรับการปรับแต่งพารามิเตอร์ (hyperparameters) ได้ทดลองจำนวนรอบการเรียนรู้ (Training Cycles) ที่ระดับ 200, 300 และ 400 รอบ และพบว่า 300 รอบให้ผลลัพธ์ดีที่สุด จึงเลือกใช้เป็นค่าที่เหมาะสม นอกจากนี้กำหนดอัตราการเรียนรู้ (Learning Rate) เท่ากับ 0.01 และโมเมนตัม (Momentum) เท่ากับ 0.9 เพื่อเพิ่มความเสถียรของกระบวนการเรียนรู้และช่วยให้แบบจำลองลู่เข้าสู่ค่าที่เหมาะสมได้อย่างมีประสิทธิภาพ ซึ่งส่งผลให้แบบจำลองสามารถเรียนรู้รูปแบบความสัมพันธ์ของข้อมูลและเพิ่มความแม่นยำในการพยากรณ์ได้อย่างมีประสิทธิภาพ

4.4 เทคนิควิธีการรวมกลุ่ม (Voting Ensemble: VE)

เทคนิค VE เป็นวิธีการเรียนรู้ของเครื่องแบบรวมกลุ่ม (Ensemble Learning) ที่ผสานแบบจำลองหลายตัวเพื่อเพิ่มความแม่นยำของการพยากรณ์ โดยอาศัยความหลากหลายของ

แบบจำลองพื้นฐานเพื่อลดความแปรปรวน (variance) และเพิ่มความน่าเชื่อถือของผลลัพธ์ ซึ่งโดยทั่วไปให้ประสิทธิภาพดีกว่าการใช้แบบจำลองเดี่ยว (อนุพงศ์ สุขประเสริฐ, 2568)

ในการศึกษานี้ แบบจำลอง VE ถูกพัฒนาจากแบบจำลองพื้นฐาน 3 วิธี ได้แก่ k-Nearest Neighbors (k-NN), Random Forest (RF) และ Neural Network (NN) โดยใช้ค่าพารามิเตอร์ที่เหมาะสมที่สุดจากการปรับแต่งในแต่ละเทคนิค (อรรศฎาวุธ เรืองสวัสดิ์ และคณะ, 2566) กล่าวคือ k-NN ใช้ค่า k = 3, RF ใช้จำนวนต้นไม้ 200 ต้น และความลึกสูงสุด 10 ระดับ และ NN ใช้โครงสร้างแบบหลายชั้น (Multilayer Perceptron: MLP) ที่มีชั้นซ่อน (Hidden Layer) 1 ชั้น จำนวน 6 โหนด พร้อมจำนวนรอบการเรียนรู้ (Training Cycles) เท่ากับ 300 รอบ อัตราการเรียนรู้ (Learning Rate) เท่ากับ 0.01 และโมเมนตัม (Momentum) เท่ากับ 0.9

ในการรวมผลลัพธ์ ผู้วิจัยใช้วิธีการเฉลี่ยค่าพยากรณ์แบบอย่างง่าย (simple averaging) ซึ่งเหมาะสมกับปัญหาการพยากรณ์เชิงประมาณค่า (regression) และช่วยลดความแปรปรวนของผลลัพธ์โดยไม่เพิ่มความซับซ้อนของแบบจำลอง เมื่อเปรียบเทียบกับวิธีการถ่วงน้ำหนัก (weighted averaging) ที่ต้องกำหนดค่าน้ำหนักเพิ่มเติม

5. การประเมินผล (Evaluation)

ภายหลังการพัฒนาแบบจำลอง ได้ทำการประเมินประสิทธิภาพเพื่อพิจารณาความสามารถในการพยากรณ์ข้อมูลใหม่ โดยใช้วิธี Split Validation แบ่งข้อมูลเป็นชุดฝึก (70%) และชุดทดสอบ (30%) เพื่อประเมินความสามารถในการ generalization ของแบบจำลอง ประสิทธิภาพของแบบจำลองถูกวัดด้วยตัวชี้วัดมาตรฐานสำหรับปัญหาการพยากรณ์เชิงประมาณค่า ได้แก่ Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) และ Coefficient of Determination (R^2) โดยตัวชี้วัดกลุ่มความคลาดเคลื่อน (MSE, RMSE, MAE และ MAPE) ใช้สะท้อนความแตกต่างระหว่างค่าจริงและค่าที่พยากรณ์ ซึ่งค่าที่ต่ำบ่งชี้ถึงความแม่นยำที่สูง ขณะที่ R^2 ใช้วัดความสามารถของแบบจำลองในการอธิบายความแปรปรวนของข้อมูล โดยค่าที่เข้าใกล้ 1 แสดงถึงประสิทธิภาพที่ดียิ่งขึ้น (Montgomery et al., 2012)

6. การนำไปใช้ (Deployment)

ขั้นตอนการนำไปใช้ (Deployment) เป็นขั้นตอนสุดท้ายของกระบวนการ CRISP-DM ซึ่งมุ่งนำแบบจำลองที่ผ่านการประเมินแล้วไปประยุกต์ใช้ในบริบทจริง เพื่อสนับสนุนการวิเคราะห์และการตัดสินใจเชิงนโยบายและเชิงธุรกิจ ในการศึกษานี้ ได้คัดเลือกแบบจำลองที่มีประสิทธิภาพสูงสุด

จากการเปรียบเทียบตัวชี้วัดความคลาดเคลื่อนและค่า R^2 เพื่อใช้เป็นเครื่องมือพยากรณ์มูลค่าการส่งออกยางพาราของประเทศไทย

แบบจำลองที่ได้สามารถใช้ในการวิเคราะห์แนวโน้มการส่งออกในอนาคต และสนับสนุนการวางแผนเชิงกลยุทธ์ในด้านการผลิต การบริหารจัดการอุปทาน และการกำหนดนโยบายการส่งออก โดยเฉพาะในสภาวะที่ตลาดมีความผันผวนจากปัจจัยเศรษฐกิจมหภาค นอกจากนี้ ผลลัพธ์เชิงพยากรณ์ยังสามารถใช้เป็นข้อมูลประกอบการตัดสินใจสำหรับหน่วยงานที่เกี่ยวข้อง เช่น การยางแห่งประเทศไทย ผู้ประกอบการ และหน่วยงานด้านนโยบาย เพื่อเพิ่มประสิทธิภาพในการบริหารความเสี่ยงและการวางแผนระยะยาวของอุตสาหกรรม

ในเชิงวิชาการ แนวทางการพัฒนาแบบจำลองด้วยเทคนิค Machine Learning และ Ensemble Learning ในการศึกษานี้ แสดงให้เห็นถึงศักยภาพในการบูรณาการข้อมูลเศรษฐกิจหลายมิติ เพื่อเพิ่มความแม่นยำของการพยากรณ์ และสามารถประยุกต์ใช้กับสินค้าเกษตรหรือสินค้าโภคภัณฑ์อื่นที่มีลักษณะข้อมูลคล้ายคลึงกันได้ ซึ่งช่วยขยายขอบเขตของการประยุกต์ใช้แบบจำลองเชิงพยากรณ์ในบริบทเศรษฐกิจที่หลากหลาย

ผลการทดลอง

1.การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปร

ผู้วิจัยได้วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระด้วยเมทริกซ์สหสัมพันธ์ (Correlation Matrix) เพื่อประเมินโครงสร้างข้อมูลและลดปัญหาพหุสหสัมพันธ์ (multicollinearity) ก่อนการสร้างแบบจำลอง โดยพบว่าค่าสหสัมพันธ์อยู่ในช่วง -0.698 ถึง 0.918 แสดงถึงความสัมพันธ์ทั้งเชิงบวกและเชิงลบในระดับที่แตกต่างกัน

เพื่อควบคุมความซ้ำซ้อนของตัวแปร ผู้วิจัยกำหนดเกณฑ์ว่าเมื่อค่าสหสัมพันธ์มากกว่า 0.7 หรือน้อยกว่า -0.7 จะถือว่ามีความสัมพันธ์ในระดับสูงและควรคัดเลือกตัวแปร โดยพบว่าตัวแปรราคายางในตลาดโลกและราคายางในตลาดประเทศไทยมีค่าสหสัมพันธ์สูง (0.918) จึงเลือกใช้เพียงตัวแปรเดียวและตัดอีกตัวแปรออกด้วยวิธีการกำจัดตัวแปรที่มีความสัมพันธ์สูง (Remove Correlated Attributes)

การดำเนินการดังกล่าวช่วยลดความซ้ำซ้อนของข้อมูล ลดผลกระทบต่อการเรียนรู้ของแบบจำลอง และเพิ่มประสิทธิภาพของอัลกอริทึม โดยเฉพาะ k-NN และ NN ซึ่งมีความไวต่อปัญหาพหุสหสัมพันธ์ ทั้งนี้ ผลการวิเคราะห์ถูกนำไปใช้ในขั้นตอนคัดเลือกตัวแปรก่อนการสร้างแบบจำลอง ดังแสดงในภาพที่ 2

Attribu...	ปริมาณการผ	ปริมาณส่งออ	ใช้ในประเท	อัตราแลกเป	ราคาน้ำมันดิ	ดัชนีราคาผู้	ราคายางในเ	ราคายางในเ	ราคาทองคำ	เปอร์เซนต์ค	เนื้อที่เก็บเกี่	ผลผลิตที่เก็
ปริมาณการผ	1	0.642	0.405	0.230	-0.130	0.316	-0.312	-0.321	-0.028	-0.072	0.198	0.267
ปริมาณส่งออ	0.642	1	0.358	0.281	-0.162	0.413	-0.425	-0.445	-0.126	-0.086	0.146	0.108
ใช้ในประเทศ	0.405	0.358	1	0.270	0.085	0.606	-0.244	-0.384	0.324	-0.010	0.129	0.240
อัตราแลกเปลี่	0.230	0.281	0.270	1	-0.217	0.385	-0.356	-0.187	-0.325	-0.151	0.462	0.134
ราคาน้ำมันดิบ	-0.130	-0.162	0.085	-0.217	1	0.067	0.607	0.479	0.239	0.083	-0.327	-0.119
ดัชนีราคาผู้บริ	0.316	0.413	0.606	0.385	0.067	1	-0.572	-0.698	0.359	-0.052	0.264	0.327
ราคายางในต	-0.312	-0.425	-0.244	-0.356	0.607	-0.572	1	0.918	0.129	0.048	-0.376	-0.283
ราคายางในต	-0.321	-0.445	-0.384	-0.187	0.479	-0.698	0.918	1	-0.149	-0.006	-0.210	-0.254
ราคาทองคำส	-0.028	-0.126	0.324	-0.325	0.239	0.359	0.129	-0.149	1	0.165	-0.197	0.134
เปอร์เซนต์กา	-0.072	-0.086	-0.010	-0.151	0.083	-0.052	0.048	-0.006	0.165	1	-0.028	0.021
เนื้อที่เก็บเกี่ย	0.198	0.146	0.129	0.462	-0.327	0.264	-0.376	-0.210	-0.197	-0.028	1	0.706
ผลผลิตที่เก็บ	0.267	0.108	0.240	0.134	-0.119	0.327	-0.283	-0.254	0.134	0.021	0.706	1

ภาพที่ 2 การวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรด้วย Correlation Matrix

2.การเปรียบเทียบประสิทธิภาพของแบบจำลอง

ผู้วิจัยได้เปรียบเทียบประสิทธิภาพของแบบจำลองโดยใช้เทคนิค k-NN, RF, NN และ VE โดยประเมินด้วยตัวชี้วัดมาตรฐาน ได้แก่ MSE, RMSE, MAE, MAPE และ R^2 เพื่อสะท้อนทั้งความคลาดเคลื่อนของการพยากรณ์และความสามารถในการอธิบายความแปรปรวนของข้อมูล ดังแสดงในตารางที่ 2

ผลการทดลองพบว่า VE ให้ประสิทธิภาพดีที่สุด โดยมีค่าความคลาดเคลื่อนต่ำที่สุดในทุกตัวชี้วัด (MSE = 3,723,671.862, RMSE = 1,929.6818, MAE = 1,502.4775, MAPE = 0.1083) และมีค่า R^2 สูงสุด (0.8259) แสดงถึงความสามารถในการอธิบายความแปรปรวนของข้อมูลได้เหนือกว่าแบบจำลองอื่น รองลงมาคือ RF (R^2 = 0.8161) ซึ่งสะท้อนถึงประสิทธิภาพของแบบจำลองประเภท tree-based ในการจัดการความสัมพันธ์เชิงไม่เป็นเส้นตรง ขณะที่ NN และ k-NN ให้ผลลัพธ์ต่ำกว่า โดยเฉพาะ k-NN (R^2 = 0.7565) ซึ่งอาจได้รับผลกระทบจากลักษณะข้อมูลที่มีความซับซ้อนและหลายมิติ

ผลลัพธ์ดังกล่าวยืนยันว่า VE สามารถเพิ่มความแม่นยำของการพยากรณ์ได้อย่างมีประสิทธิภาพ เนื่องจากเป็นการรวมจุดเด่นของแบบจำลองหลายประเภทและลดข้อจำกัดของแบบจำลองเดี่ยว โดยเฉพาะในบริบทของข้อมูลเศรษฐกิจที่มีความสัมพันธ์ไม่เชิงเส้นและมีความผันผวนสูง

ตารางที่ 2 เปรียบเทียบประสิทธิภาพของการประมาณค่าข้อมูลของตัวแบบเชิงพยากรณ์

Model	MSE	RMSE	MAE	MAPE	R^2
k-NN	5,274,055.411	2,296.5312	1,726.0823	0.1275	0.7565
RF	4,043,493.084	2,010.8439	1,566.0495	0.1140	0.8161
NN	4,514,462.136	2,124.7645	1,628.2451	0.1162	0.7869
VE	**3,723,671.862**	**1,929.6818**	**1,502.4775**	**0.1083**	**0.8259**

อภิปรายผล

ผลการศึกษาพบว่า Voting Ensemble (VE) ให้ประสิทธิภาพสูงสุด โดยมีค่าความคลาดเคลื่อนต่ำที่สุด (MSE, RMSE, MAE, MAPE) และมีค่า R^2 สูงสุด (0.8259) สะท้อนถึงความสามารถของแบบจำลองในการอธิบายความแปรปรวนของข้อมูลได้เหนือกว่าแบบจำลองเดี่ยว ผลลัพธ์ดังกล่าวยืนยันว่าเทคนิค Ensemble Learning สามารถเพิ่มความแม่นยำและความเสถียรของการพยากรณ์ได้ โดยอาศัยการผสานจุดเด่นของแบบจำลองหลายประเภท และลดข้อจำกัดด้านความเอนเอียง (bias) และความแปรปรวน (variance) โดยเฉพาะในบริบทข้อมูลเศรษฐกิจที่มีความสัมพันธ์ไม่เชิงเส้นและหลายมิติ

ในเชิงโครงสร้างข้อมูล มูลค่าการส่งออกยางพาราได้รับอิทธิพลจากปัจจัยเศรษฐกิจหลายด้าน เช่น อัตราแลกเปลี่ยน ราคาน้ำมันดิบ และราคายางในตลาดโลก ซึ่งมีความสัมพันธ์เชิงซับซ้อนและเปลี่ยนแปลงตามภาวะเศรษฐกิจโลก ส่งผลให้แบบจำลองที่สามารถเรียนรู้ความสัมพันธ์แบบไม่เชิงเส้นได้อย่างยืดหยุ่น เช่น VE มีประสิทธิภาพเหนือกว่าแบบจำลองเดี่ยว

ผลการศึกษานี้สอดคล้องกับ Makridakis et al. (2022) ที่ชี้ว่า Machine Learning โดยเฉพาะแบบจำลองเชิงผสมสามารถให้ความแม่นยำสูงในงานพยากรณ์ที่มีความซับซ้อน และสอดคล้องกับ Zhou (2019) ที่ระบุว่า Ensemble Learning สามารถเพิ่มประสิทธิภาพของแบบจำลองผ่านการรวมผลลัพธ์ของหลายอัลกอริทึม นอกจากนี้ ในมุมมองเศรษฐศาสตร์ ผลลัพธ์ยังสะท้อนถึงความเชื่อมโยงของการส่งออกยางพารากับพลวัตของเศรษฐกิจโลก เช่น อุตสาหกรรมยานยนต์ ราคาพลังงาน และต้นทุนยางสังเคราะห์ ซึ่งเป็นปัจจัยสำคัญต่ออุปสงค์ของยางธรรมชาติ (International Rubber Study Group, 2022)

ดังนั้น งานวิจัยนี้แสดงให้เห็นถึงศักยภาพของ Machine Learning โดยเฉพาะ VE ในการพยากรณ์ข้อมูลเศรษฐกิจเชิงพหุคูณ และสามารถนำไปใช้สนับสนุนการวิเคราะห์แนวโน้ม การบริหารความเสี่ยง และการกำหนดนโยบายได้อย่างมีประสิทธิภาพ

เป็นข้อเสนอแนะ

ผลการวิจัยพบว่าแบบจำลอง Voting Ensemble (VE) มีประสิทธิภาพสูงสุดในการพยากรณ์มูลค่าการส่งออกยางพาราของประเทศไทย โดยให้ค่าความคลาดเคลื่อนต่ำและมีค่า R^2 ในระดับสูง สะท้อนถึงศักยภาพของเทคนิค Ensemble Learning ในการจัดการข้อมูลเศรษฐกิจที่มีความซับซ้อนและหลายมิติได้อย่างมีประสิทธิภาพ

อย่างไรก็ตาม งานวิจัยนี้ยังมีข้อจำกัด ได้แก่ ขนาดชุดข้อมูลที่ค่อนข้างจำกัด ซึ่งอาจส่งผลต่อความสามารถในการนำแบบจำลองไปใช้กับข้อมูลใหม่ (generalization) และเพิ่มความเสี่ยงต่อการเรียนรู้มากเกินไป (overfitting) รวมถึงตัวแปรที่ใช้ยังไม่ครอบคลุมปัจจัยเศรษฐกิจมหภาคบางประการ เช่น อุปสงค์โลกหรือภาวะเศรษฐกิจระหว่างประเทศ นอกจากนี้ แม้ว่าข้อมูลจะมีลักษณะเชิงเวลา แต่การศึกษานี้ไม่ได้ประยุกต์ใช้เทคนิคการวิเคราะห์อนุกรมเวลา (time-series modeling) เช่น ตัวแปรหน่วงเวลา (lag features) หรือโครงสร้างเชิงลำดับของข้อมูล ซึ่งอาจทำให้แบบจำลองไม่สามารถสะท้อนพลวัตเชิงเวลาได้อย่างเต็มที่

ดังนั้น งานวิจัยในอนาคตควรมุ่งเน้นการขยายช่วงข้อมูลและเพิ่มตัวแปรเศรษฐกิจมหภาค เช่น ดัชนีเศรษฐกิจโลก รวมถึงพัฒนาแบบจำลองให้รองรับข้อมูลเชิงเวลา โดยใช้เทคนิค เช่น Time Series Cross-Validation, Walk-Forward Validation และแบบจำลองเฉพาะทาง เช่น ARIMA หรือ LSTM นอกจากนี้ ควรศึกษาความสำคัญของตัวแปร (feature importance) และพัฒนา Ensemble ขั้นสูง เช่น weighted voting หรือ stacking เพื่อยกระดับประสิทธิภาพของแบบจำลองในบริบทข้อมูลเศรษฐกิจที่มีความผันผวนสูง

กิตติกรรมประกาศ

คณะผู้วิจัยขอขอบคุณหน่วยงานที่ให้บริการข้อมูลสถิติและข้อมูลเศรษฐกิจที่ใช้ในการศึกษาครั้งนี้ ได้แก่ กรมวิชาการเกษตร สำนักงานนโยบายและยุทธศาสตร์การค้า การยางแห่งประเทศไทย และฐานข้อมูลด้านการลงทุนที่เกี่ยวข้อง นอกจากนี้ ขอขอบคุณคณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม ที่ให้การสนับสนุนการดำเนินงานวิจัยจนสำเร็จลุล่วงด้วยดี

เอกสารอ้างอิง

กรมส่งเสริมการเกษตร. (n.d.). เนื้อที่เก็บเกี่ยว ผลผลิตที่เก็บเกี่ยวได้ และราคาขายเฉลี่ย. สืบค้นเมื่อ February 9, 2025, จาก https://production.doae.go.th/site/login

กองการยางกรมวิชาการเกษตร. (n.d.). สถิติยางพาราในประเทศไทย. สืบค้นเมื่อ October 18, 2025, จาก https://www.doa.go.th/rubber/?p=2140

กองการยางแห่งประเทศไทย. (n.d.). ข้อมูลยางพารา. สืบค้นเมื่อ February 9, 2025, จาก https://raot.co.th/main.php?filename=index

จิดาภา ลู่วิโรจน์. (2566). เศรษฐศาสตร์อุตสาหกรรม: เศรษฐศาสตร์การเกษตรยางพาราไทย. สืบค้นเมื่อ February 15, 2025, จาก https://setthasarn.econ.tu.ac.th/blog/detail/661

จารุวรรณ สิงห์ม่วง และ ธิดาพร ศุภภากร. (2563). ตัวแบบพยากรณ์มูลค่าการส่งออกยางพาราของประเทศไทย. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครสวรรค์, 12(15), 58–82. https://ph02.tci-thaijo.org/index.php/JSTNSRU/article/view/231007

ไชยวัฒน์ เสาวเจริญสุข. (2565). แนวโน้มอุตสาหกรรมยางพารา ปี 2022–2024. Krungsri Research, 65(1), 3–23.

อภิสิทธิ์ หะยีอุมา และ ธนวิทย์ บุญสิทธิ์. (2563). การวิเคราะห์ปัจจัยที่ส่งผลต่อมูลค่าและปริมาณการส่งออกยางพาราของไทยในประเทศสิงคโปร์. วารสารวิชาการนอร์ทเทิร์น, 7(1), 65–78. สืบค้นจาก https://so09.tci-thaijo.org/index.php/AJntc/article/view/1349

อนุพงศ์ สุขประเสริฐ. (2568). คู่มือการทำเหมืองข้อมูลด้วยโปรแกรม RapidMiner Studio (พิมพ์ครั้งที่ 5). มหาวิทยาลัยมหาสารคาม.

อรรศฎาวุธ เรืองสวัสดิ์, อนุพงศ์ สุขประเสริฐ, ทิวา สินธุกูฏ, และ ศิริลักษณ์ ไกยวินิจ. (2566). การเปรียบเทียบประสิทธิภาพของตัวแบบสำหรับการพยากรณ์โรคมะเร็งปอด. วารสารวิทยาศาสตร์ เทคโนโลยี และนวัตกรรม มหาวิทยาลัยกาฬสินธุ์, 2(2), 39–52. https://doi.org/10.14456/ksti.2023.8

Cherdchoongam, S., & Rungreunganun, V. (2016). Forecasting the price of natural rubber in Thailand using the ARIMA model. *Applied Science and Engineering Progress*, 9(4). https://ph02.tci-thaijo.org/index.php/ijast/article/view/72567

Guindani, L. G., Oliveira, G. A., Ribeiro, M. H. D. M., Gonzalez, G. V., & de Lima, J. D. (2024). Exploring current trends in agricultural commodities forecasting methods through text mining: Developments in statistical and artificial intelligence methods. *Heliyon*, 10(23), e40568. https://doi.org/10.1016/j.heliyon.2024.e40568

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

International Rubber Study Group. (2022). *Rubber statistical bulletin*. https://www.rubberstudy.org/reports

Investing.com. (n.d.-a). ข้อมูลราคาน้ำมันดิบย้อนหลัง. สืบค้นเมื่อ November 17, 2025, จาก https://th.investing.com/commodities/crude-oil-historical-data

Investing.com. (n.d.-b). ข้อมูลราคาทองคำย้อนหลัง (XAU/USD). สืบค้นเมื่อ December 27, 2025, จาก https://th.investing.com/currencies/xau-usd-historical-data

Investing.com. (n.d.-c). ข้อมูลราคายางพารา TSR20 ย้อนหลัง. สืบค้นเมื่อ October 13, 2025, จาก https://th.investing.com/commodities/rubber-tsr20-futures-historical-data

Investing.com. (n.d.-d). ข้อมูลราคายางพารา TOCOM. สืบค้นเมื่อ October 13, 2025, จาก https://th.investing.com/commodities/tocom-rubber-futures

Investing.com. (n.d.-e). ข้อมูลอัตราแลกเปลี่ยน USD/THB ย้อนหลัง. สืบค้นเมื่อ September 26, 2025, จาก https://th.investing.com/currencies/usd-thb-historical-data

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18–22. https://cran.r-project.org/doc/Rnews/Rnews_2002-3.pdf

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). The M5 accuracy competition: Results, findings and conclusions. *International Journal of Forecasting*, 38(4), 1346–1364. https://doi.org/10.1016/j.ijforecast.2021.11.013

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). Wiley.

Nyondo, P., & Varghese, R. (2024). Forecasting natural rubber prices using commodity market indicators: A machine learning approach. International Journal of Revenue Management, 14(3), 221–252. https://doi.org/10.1504/IJRM.2024.142325

Zhang, Y., Wang, J., & Liu, X. (2023). Forecasting commodity prices using XGBoost model. *Resources Policy*, 81, 103321. https://doi.org/10.1016/j.resourpol.2023.103321

Zhou, Z.-H. (2019). Ensemble learning. *National Science Review*, 6(4), 801–802. https://doi.org/10.1093/nsr/nwz068s