

การเปรียบเทียบประสิทธิภาพของตัวแบบเชิงพยากรณ์สำหรับปริมาณการส่งออก มะพร้าวในประเทศไทย

Comparison of Predictive Model Performance for Coconut Export Volume in Thailand

ชนมนิภา แฝงเมืองคุก¹ กรพิน ภิรมย์¹ ณัฐพร คำวรรณ¹ พิกุล สัตย์ไธสง¹ อนุนพงศ์ สุขประเสริฐ^{2*}

Chonnipha Faengmueangkuk¹ Korapin Phirom¹ Natthaporn Khamwan¹

Phikool Sattaisong¹ Anupong Sukprasert^{2*}

¹นิสิตสาขาวิชาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

²อาจารย์ประจำสาขาวิชาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม

* Email: anupong.s@msu.ac.th

บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบและเปรียบเทียบประสิทธิภาพของตัวแบบสำหรับการพยากรณ์ปริมาณการส่งออกมะพร้าวในประเทศไทย โดยใช้ข้อมูลจำนวน 73 ระเบียบ และ 17 แอตทริบิวต์ ซึ่งนำมาวิเคราะห์ข้อมูลตามกระบวนการมาตรฐานการทำเหมืองข้อมูล และใช้เทคนิคการประมาณค่าข้อมูล (Regression Model) สำหรับการสร้างแบบจำลอง ในการวิจัยนี้ได้เลือกใช้เทคนิคการเรียนรู้ของเครื่อง 4 เทคนิค ได้แก่ เทคนิคสมการถดถอยเชิงเส้น (Linear Regression), เทคนิคเพื่อนบ้านใกล้ที่สุด (k-Nearest Neighbors), เทคนิคโครงข่ายประสาทเทียม (Neural Network), และ เทคนิคต้นไม้ตัดสินใจ (Decision Tree) จากผลการวิเคราะห์ พบว่า เทคนิคเพื่อนบ้านใกล้ที่สุด (k-Nearest Neighbors) เป็นเทคนิคที่มีความเหมาะสมที่สุด สำหรับการสร้างแบบจำลองการพยากรณ์ปริมาณการส่งออกมะพร้าวในประเทศไทย โดยให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Square Error: MSE) ต่ำที่สุดที่ 38,688,625,970,446.50, ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Squared Error: RMSE) เท่ากับ 6,216,200.551, ค่าความคลาดเคลื่อนสัมบูรณ์ (Absolute Error: AE) เท่ากับ 4,612,294.570, ค่าความคลาดเคลื่อนกำลังสอง (Squared Error: SE) เท่ากับ 45,260,457,821,710.650, และค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R-Squared: R^2) เท่ากับ 0.888 (88.80%) เนื่องจาก เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) มีค่าความคลาดเคลื่อนต่ำที่สุด ทำให้การพยากรณ์แม่นยำที่สุด ผลลัพธ์จากการวิเคราะห์ข้อมูลสามารถนำไปใช้ในการพัฒนาแบบจำลองสำหรับพยากรณ์ปริมาณการส่งออกมะพร้าวในประเทศไทย ซึ่งจะเป็นประโยชน์ต่อหน่วยงานที่เกี่ยวข้อง อาทิ กระทรวงเกษตรและสหกรณ์ ผู้ประกอบการส่งออก ตลอดจนเกษตรกรผู้เพาะปลูกมะพร้าว เพื่อให้สามารถวางแผนการผลิต การตลาด และการบริหารจัดการทรัพยากรได้อย่างมีประสิทธิภาพ ทั้งนี้ แบบจำลองดังกล่าวสามารถ

นำไปปรับปรุงและประยุกต์ใช้กับสินค้าเกษตรประเภทอื่น เพื่อเพิ่มความแม่นยำในการพยากรณ์และเสริมสร้างเสถียรภาพทางเศรษฐกิจของภาคเกษตรกรรมไทย นอกจากนี้ การใช้แบบจำลองพยากรณ์ช่วยให้หน่วยงานที่เกี่ยวข้องสามารถติดตามแนวโน้มการส่งออก และกำหนดนโยบายที่เหมาะสมกับสถานการณ์ทางเศรษฐกิจและการค้าระหว่างประเทศได้อย่างทันทั่วทั้งที่ อีกทั้งยังเอื้อประโยชน์ต่อผู้ประกอบการในการบริหารจัดการปริมาณสินค้า ลดความสูญเสียอันเนื่องมาจากปัญหาสินค้าล้นตลาดหรือขาดแคลน ตลอดจนเพิ่มขีดความสามารถในการแข่งขันในตลาดโลก นอกเหนือจากนั้น ข้อมูลที่ได้รับจากการวิเคราะห์สามารถนำมาใช้ในการพัฒนากลยุทธ์ทางการตลาด รวมถึงการกำหนดราคาสินค้าให้สอดคล้องกับกลไกอุปสงค์และอุปทาน ซึ่งจะช่วยให้เกษตรกรได้รับผลตอบแทนที่เหมาะสมและเป็นธรรม อันจะส่งเสริมให้ภาคเกษตรกรรมของประเทศไทยเติบโตอย่างมั่นคงและยั่งยืน

คำสำคัญ : ปริมาณการส่งออกมะพร้าว, เทคนิคสมการถดถอยเชิงเส้น, เทคนิคเพื่อนบ้านใกล้ที่สุด, เทคนิคโครงข่ายประสาทเทียม, เทคนิคต้นไม้ตัดสินใจ

Abstract

This research aims to develop and compare the performance of models for forecasting coconut export volumes in Thailand using 73 records and 17 attributes. The data was analyzed following standard data mining processes, employing regression techniques to construct predictive models. Four machine learning techniques were selected: Linear Regression, k-Nearest Neighbors (k-NN), Neural Network, and Decision Tree. The analysis results indicate that the k-NN technique is the most suitable for forecasting coconut export volumes, achieving the lowest error values Mean Squared Error (MSE) of 38,688,625,970,446.50, Root Mean Squared Error (RMSE) of 6,216,200.551, Absolute Error (AE) of 4,612,294.570, Squared Error (SE) of 45,260,457,821,710.650, and a multiple coefficient of determination (R^2) of 0.888 (88.80%) ensuring the highest prediction accuracy. The findings from this study can be applied to enhance forecasting models for coconut exports, benefiting relevant organizations such as the Ministry of Agriculture and Cooperatives, exporters, and coconut farmers by supporting effective production planning, marketing strategies, and resource management. Additionally, the developed model can be refined and adapted for other agricultural products to improve forecasting accuracy and strengthen the economic stability of Thailand's agricultural sector. Furthermore, using forecasting models enables authorities to monitor export trends, develop appropriate policies in response to economic and international trade conditions, and assist businesses in managing inventory efficiently. This helps reduce losses caused by market oversupply or shortages and enhances competitiveness in the global market. The insights derived from this analysis can also be utilized to develop marketing strategies and pricing mechanisms aligned with supply and demand dynamics, ensuring fair and reasonable returns for farmers. Ultimately, this contributes to the stable and sustainable growth of Thailand's agricultural sector.

Keywords: Coconut export volume, Linear Regression, k-Nearest Neighbors, Neural Network, Decision Tree

1. บทนำ

มะพร้าวไทย เป็นพืชที่ชาวไทยจากรุ่นสู่รุ่นนำมาใช้ประกอบอาหารในหลายรูปแบบ ทั้งอาหารคาวและอาหารหวาน นอกจากนั้นมะพร้าวยังเป็นหนึ่งในพืชเศรษฐกิจของไทยที่ทำเงินได้ไม่แพ้ราชาผลไม้อย่างทุเรียน นอกจากในเรื่องของการบริโภคผลแล้ว ส่วนของมะพร้าวสามารถนำไปใช้ประโยชน์ได้หลากหลายแทบทุกส่วน ไม่ว่าจะเป็น ลำต้น แกนใบ (ทางมะพร้าว) ใบและก้านใบ ราก ช่อดอก เปลือก (กาบมะพร้าว) และยอดมะพร้าว เป็นต้น ส่วนในมุมของการส่งออกมะพร้าวไทยไปต่างประเทศ การส่งออกมะพร้าวไทยอ่อนไปยังต่างประเทศ ปี 2566 มีมูลค่าการส่งออกทั้งสิ้น 288.04 ล้านดอลลาร์สหรัฐ เติบโตเพิ่มขึ้น 3.98% คิดเป็นปริมาณการส่งออกทั้งสิ้นถึง 432,492 ตัน [1]

จากข้อมูลสถิติพบว่าในปี 2560 มูลค่าของตลาดมะพร้าวอยู่ที่ 10,280 ล้านบาท และเพิ่มสูงถึง 14,440 ล้านบาท ในปี 2564 คาดการณ์ว่ามูลค่าดังกล่าวจะทะลุ 20,000 ล้านบาท ในปี 2569 ประเทศจีนบริโภคมะพร้าวสดประมาณ 2,600 ล้านลูกและแปรรูปมะพร้าวประมาณ 1,500 ล้านลูกต่อปี ในขณะที่ผลผลิตมะพร้าวของจีนส่วนใหญ่มาจากมณฑลไห่หนาน โดยมีผลผลิตต่อปีประมาณอยู่ที่ 250 ล้านลูกเท่านั้น สามารถตอบสนองความต้องการในการบริโภคและการแปรรูปได้เพียงร้อยละ 6 เท่านั้น และที่เหลือต้องนำเข้าจากต่างประเทศ โดยเฉพาะนำเข้าจากประเทศอาเซียน ได้แก่ ไทย อินโดนีเซีย เวียดนาม และฟิลิปปินส์ ประเทศไทยเป็นประเทศอันดับหนึ่งของจีนในการนำเข้ามะพร้าวสด โดยคิดเป็นร้อยละ 60 ของปริมาณการนำเข้าทั้งหมดของจีน ช่วง 8 เดือนแรกของปี 2567 จีนนำเข้ามะพร้าวสดจากต่างประเทศปริมาณ 558,450 ตัน มูลค่า 1,965 ล้านบาท ลดลงร้อยละ 32 เมื่อเทียบกับช่วงเวลาเดียวกัน สำหรับนำเข้าจากประเทศไทยมีปริมาณ 189,336 ตัน มูลค่า 1,261 ล้านบาท ลดลงร้อยละ 44 สำหรับประเทศอื่น ปัจจุบันธุรกิจมะพร้าวในตลาดจีนกำลังอยู่แนวโน้มการเติบโต แนวโน้มนี้ส่วนใหญ่เกิดจากผู้บริโภคให้ความสำคัญกับแนวคิดของการรับประทานอาหารเพื่อสุขภาพมากขึ้น ในขณะที่มะพร้าวมีคุณสมบัติตามธรรมชาติดีต่อสุขภาพและมีคุณค่าทางโภชนาการ ซึ่งได้รับความชื่นชอบของผู้บริโภคอย่างกว้างขวาง โดยเฉพาะในตลาดเครื่องดื่มน้ำมะพร้าวได้กลายเป็นช่องทางการบริโภคใหม่ ซึ่งได้ผลักดันให้ตลาดมะพร้าวเพิ่มขึ้นทุกปี [2]

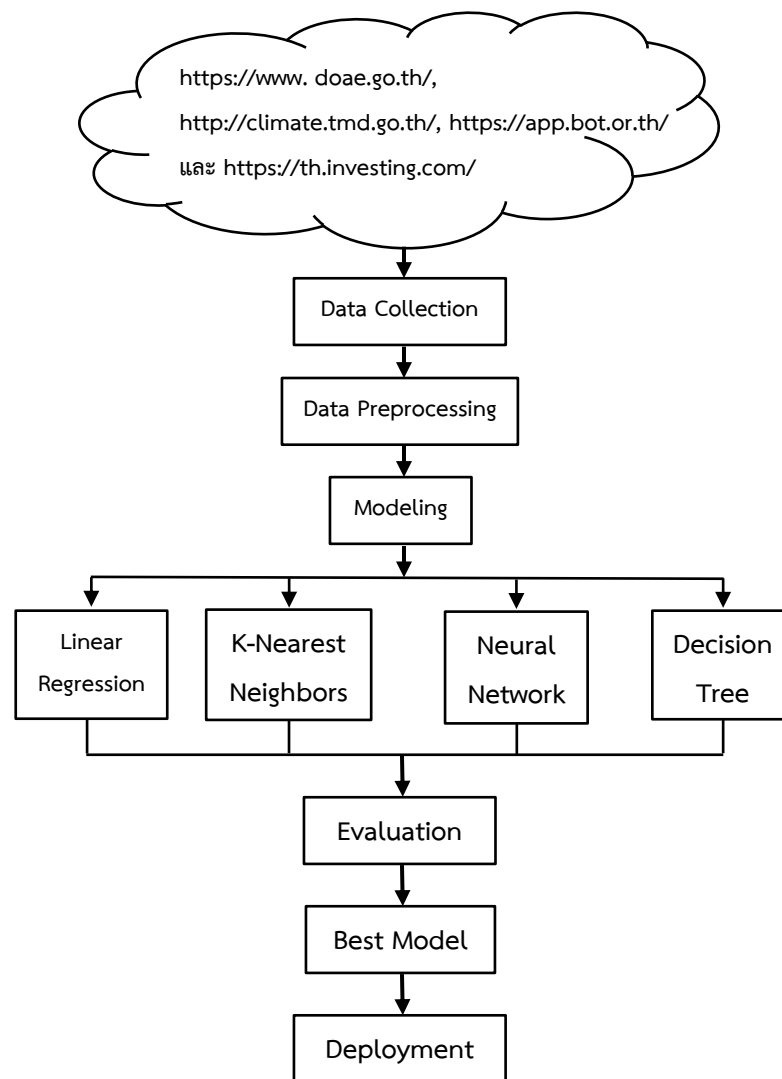
ดังนั้น ผู้วิจัยจึงมีแนวคิดที่จะศึกษาการสร้างตัวแบบและเปรียบเทียบประสิทธิภาพของตัวแบบด้วยเทคนิคการทำเหมืองข้อมูล เพื่อสร้างตัวแบบและเปรียบเทียบประสิทธิภาพในการส่งออกมะพร้าว เพื่อเพิ่มความสามารถในการวิเคราะห์การส่งออกให้แม่นยำ โดยใช้เทคนิคการจำแนกประเภทข้อมูลสำหรับการส่งออกมะพร้าวในครั้งนี้ ซึ่งจะส่งผลดีต่อการทำงานของหน่วยงานเศรษฐกิจ รวมไปถึงธุรกิจภาครัฐและเอกชน และในอนาคตยังสามารถนำไปประยุกต์ใช้ในการสร้างระบบสนับสนุนการตัดสินใจในส่วนของแนวทางการส่งออกมะพร้าวได้ต่อไป

2. วัตถุประสงค์

1. เพื่อพัฒนาตัวแบบพยากรณ์ปริมาณการส่งออกมะพร้าวของประเทศไทย
2. เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบพยากรณ์ ปริมาณการส่งออกมะพร้าวของประเทศไทย

3. วิธีการดำเนินงานวิจัย

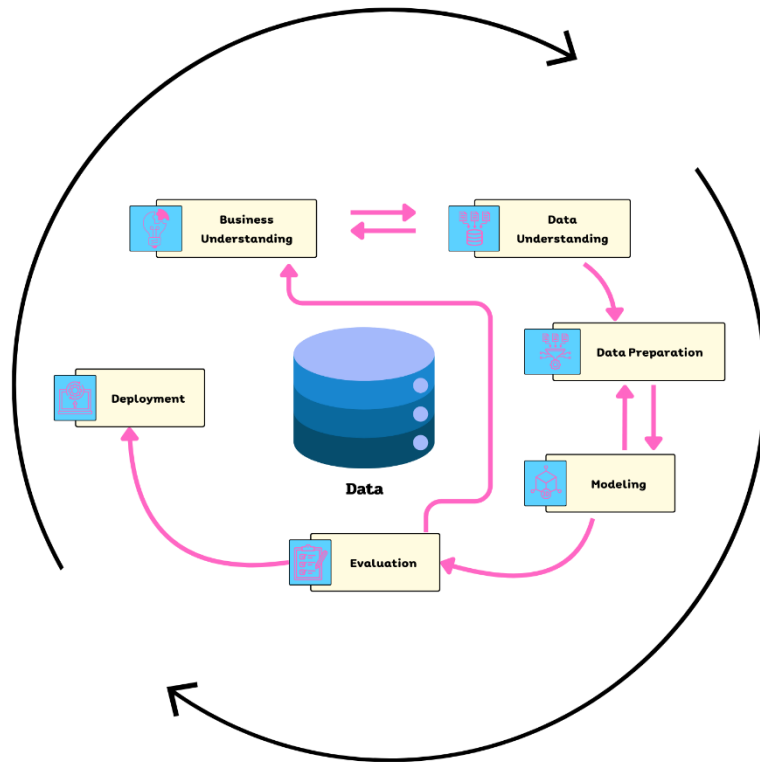
งานวิจัยนี้ ผู้วิจัยได้ทำการเก็บรวบรวมข้อมูลมาหลายแหล่งข้อมูลโดยนำมาจากเว็บไซต์ต่างๆ ดังนี้ <https://www.doae.go.th/>, <http://climate.tmd.go.th/>, <https://app.bot.or.th/> และ <https://th.investing.com/> เพื่อให้ได้ชุดข้อมูลที่เหมาะสมมีคุณภาพพร้อมในการทำเหมืองข้อมูล (Data Mining) จากนั้นเข้าสู่กระบวนการสร้างตัวแบบพยากรณ์ด้วยเทคนิคการจำแนกประเภทข้อมูล 4 เทคนิค ได้แก่ เทคนิคสมการถดถอยเชิงเส้น (Linear Regression) เทคนิคเพื่อนบ้านใกล้ที่สุด (k-Nearest Neighbors) เทคนิคโครงข่ายประสาทเทียม (Neural Network) และเทคนิคต้นไม้แบบสุ่ม (Decision Tree) จากนั้นทดสอบประสิทธิภาพตัวแบบการพยากรณ์ด้วยการแบ่งชุดข้อมูลด้วยวิธี Cross-Validation เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์ให้ได้ตัวแบบที่ดีที่สุด ดังภาพที่ 1 [3]



ภาพที่ 1 ขั้นตอนการดำเนินงานวิจัย

4. กระบวนการมาตรฐานในการทำเหมืองข้อมูล (Cross Standard Process for Data mining)

การใช้เทคนิคการทำเหมืองข้อมูลในการ การเปรียบเทียบประสิทธิภาพของตัวแบบเชิงพยากรณ์สำหรับปริมาณการส่งออกมะพร้าวในประเทศไทย ใช้กระบวนการมาตรฐานของการทำเหมืองข้อมูล CRISP-DM (Cross-Industry Standard Process for Data Mining) ประกอบด้วย 6 ขั้นตอน ดังภาพที่ 2



ภาพที่ 2 แสดงกระบวนการมาตรฐานการทำเหมืองข้อมูล (CRISP-DM)

4.1 การทำความเข้าใจกับปัญหา (Business Understanding)

การส่งออกมะพร้าวเป็นหนึ่งในอุตสาหกรรมสำคัญของประเทศไทย ซึ่งมีบทบาทต่อเศรษฐกิจและภาคการเกษตรของประเทศ อย่างไรก็ตาม ปริมาณการส่งออกมะพร้าวอาจได้รับผลกระทบจากหลายปัจจัย เช่น นโยบายการค้าระหว่างประเทศ ความต้องการของตลาดโลก สภาพอากาศ ภัยธรรมชาติ และต้นทุนการผลิต ในช่วงไม่กี่ปีที่ผ่านมา อุตสาหกรรมส่งออกมะพร้าวของไทยต้องเผชิญกับความท้าทายหลายประการ เช่น การแข่งขันจากประเทศผู้ผลิตรายอื่น เช่น อินโดนีเซีย ฟิลิปปินส์ และเวียดนาม ที่มีต้นทุนการผลิตต่ำกว่า มาตรการด้านการค้าและภาษีของประเทศคู่ค้า ที่อาจส่งผลต่อการส่งออก ปัญหาการขาดแคลนวัตถุดิบ เนื่องจากปัจจัยทางสภาพอากาศและการเปลี่ยนแปลงของพื้นที่เพาะปลูก ผลกระทบจากสถานการณ์โลก เช่น การแพร่ระบาดของโควิด-19 ที่ส่งผลกระทบต่อห่วงโซ่อุปทานและการขนส่ง [4]

ดังนั้น ผู้วิจัยจึงมีความสนใจศึกษาเรื่องการสร้างตัวแบบและเปรียบเทียบประสิทธิภาพของตัวแบบเชิงพยากรณ์สำหรับปริมาณการส่งออกมะพร้าวในประเทศไทย เพื่อเพิ่มขีดความสามารถในการวิเคราะห์การส่งออกมะพร้าวให้แม่นยำ ซึ่งเป็นประโยชน์ต่อภาครัฐ ภาคธุรกิจ และเกษตรกรในการวางแผนการผลิตและการส่งออกให้เหมาะสมกับสถานการณ์ตลาดโลก

4.2 การทำความเข้าใจเกี่ยวกับข้อมูล (Data Understanding)

ผู้วิจัยได้รวบรวมข้อมูลที่เกี่ยวข้อง เพื่อใช้ในการวิเคราะห์ด้วยเทคนิคการทำเหมืองข้อมูล โดยรวบรวมข้อมูลจากหลายๆ แหล่ง โดยรวบรวมจากเว็บไซต์ต่างๆ ดังนี้ <https://www.doe.go.th/>, <http://climate.tmd.go.th/>, <https://app.bot.or.th/> และ <https://th.investing.com/> ซึ่งมีจำนวนข้อมูลทั้งหมด 72 ชุด โดยมีตัวแปรที่ใช้สำหรับการศึกษา 17 แอทริบิวต์ ได้แก่ เดือน/ปี จำนวนมะพร้าวที่ส่งออก (กิโลกรัม) จำนวนเงินรวม (บาท) ปริมาณน้ำฝนเฉลี่ยต่อเดือน (มิลลิเมตร) อัตราแลกเปลี่ยนเงินตรา (ดอลลาร์)ราคาน้ำมันดิบล่าสุด (ดอลลาร์) ราคาเปิดน้ำมันดิบ (ดอลลาร์) ราคาน้ำมันดิบสูงสุด (ดอลลาร์) ราคาน้ำมันดิบต่ำสุด (ดอลลาร์) ปริมาณราคาน้ำมันดิบ (ดอลลาร์) เปอร์เซนต์เปลี่ยนน้ำมันดิบ ราคาทองคำล่าสุด (ดอลลาร์) ราคาเปิดทอง (ดอลลาร์) ราคาทองสูงสุด (ดอลลาร์) ราคาทองต่ำสุด (ดอลลาร์) ปริมาณราคาทอง (ดอลลาร์) เปอร์เซนต์เปลี่ยนทอง โดยข้อมูลชุดนี้ทำการรวบรวมและเก็บอยู่ในรูปแบบไฟล์ xlsx ไฟล์ ชื่อชุดข้อมูล พยากรณ์การส่งออกมะพร้าว

4.3 การเตรียมข้อมูล (Data Preparation)

เป็นการเตรียมข้อมูลซึ่งเป็นกระบวนการให้เกิดความมั่นใจคุณภาพของข้อมูลที่จะนำมาใช้ แสดงถึงความน่าเชื่อถือของข้อมูลมีทั้งหมด 2 ขั้นตอน ดังนี้

4.3.1 ผู้วิจัยได้ดำเนินการรวบรวมข้อมูลจากฐานข้อมูลปริมาณการส่งออกมะพร้าวในประเทศไทย โดยข้อมูลดังกล่าวครอบคลุมช่วงเวลาตั้งแต่เดือนมกราคม พ.ศ. 2562 ถึงเดือนธันวาคม พ.ศ. 2567 สำหรับข้อมูลที่น่ามาใช้ในการวิเคราะห์นี้ซึ่งแสดงรายละเอียดดังตารางที่ 1

ตารางที่ 1 แสดงลักษณะชุดข้อมูลและที่มาของข้อมูล “พยากรณ์การส่งออกมะพร้าว”

No	Name	Data Type	Description	แหล่งที่มา
1	M/Y	Date	เดือน/ปี	https://www.doe.go.th/
2	Product _amount (kg)	Integer	ปริมาณมะพร้าวที่ส่งออก (กิโลกรัม)	
3	Product (baht)	Integer	มูลค่าการส่งออก (บาท)	
4	Rainfall (mm)	real	ปริมาณน้ำฝนเฉลี่ยต่อเดือน (มิลลิเมตร)	http://climate.tmd.go.th/

5	Exchange rate for foreign currencies: Unit: Baht per 1 unit	real	อัตราแลกเปลี่ยน เงินตรา (ดอลลาร์)	https://app.bot.or.th/BTWS STAT/statistics/BOTWEBSTA T.aspx
6	latest:oil_Dollar	real	ราคาน้ำมันดิบ ล่าสุด (ดอลลาร์)	https://th.investing.com/
7	opening price:oil_Dollar	real	ราคาเปิด น้ำมันดิบ (ดอลลาร์)	
8	Max:oil_Dollar	real	ราคาน้ำมันดิบ สูงสุด (ดอลลาร์)	
9	min:oil_Dollar	real	ราคาน้ำมันดิบ ต่ำสุด (ดอลลาร์)	
10	Quantity:oil_Dollar	Integer	ปริมาณราคา น้ำมันดิบ (ดอลลาร์)	
11	% change_oil	real	เปอร์เซ็นต์เปลี่ยน น้ำมันดิบ	
12	latest:gold_Dollar	real	ราคาทองคำล่าสุด (ดอลลาร์)	https://th.investing.com/
13	opening price:gold_Dollar	real	ราคาเปิดทอง (ดอลลาร์)	
14	Max:gold_Dollar	real	ราคาทองสูงสุด (ดอลลาร์)	
15	min:gold_Dollar	real	ราคาทองต่ำสุด (ดอลลาร์)	
16	Quantity:gold_Dollar	Integer	ปริมาณราคาทอง (ดอลลาร์)	
17	% change_gold	real	เปอร์เซ็นต์เปลี่ยน ทอง	

ข้อมูลดังกล่าวถูกนำไปใช้ในการวิเคราะห์และพัฒนาแบบจำลองพยากรณ์ เพื่อศึกษาความสัมพันธ์ระหว่างปัจจัยที่เกี่ยวข้องกับปริมาณการส่งออกมะพร้าวในประเทศไทย รวมถึงการประเมินประสิทธิภาพของแบบจำลองในบริบทของการพยากรณ์การส่งออกมะพร้าวในประเทศไทยในอนาคต

4.3.2 การกำหนดหน้าที่ให้กับแอตทริบิวต์โดยกำหนดให้แอตทริบิวต์ที่ชื่อ จำนวนมะพร้าวที่ส่งออก (กิโลกรัม) Product _amount (kg) ทำหน้าที่เป็น Label และกำหนดหน้าที่ให้แอตทริบิวต์โดยกำหนดให้แอตทริบิวต์ที่ชื่อ เดือน/ปี (M/Y) โดยกำหนดให้เป็น ID ส่วนแอตทริบิวต์ที่เหลืออีก 15 แอตทริบิวต์ใช้สำหรับเป็นตัวแปรในการวิเคราะห์ซึ่งได้แสดงรายละเอียดดังตารางที่ 2

ตารางที่ 2 แสดงแอตทริบิวต์ข้อมูลสำหรับการวิเคราะห์

No.	Name	Data Type	Description
1	M/Y (ID)	Date	เดือน/ปี
2	Product _amount (kg) (Label)	Integer	จำนวนมะพร้าวที่ส่งออก (กิโลกรัม)
3	Product (baht)	Integer	จำนวนเงินรวม (บาท)
4	Rainfall (mm)	real	ปริมาณน้ำฝนเฉลี่ยต่อเดือน (มิลลิเมตร)
5	Exchange rate for foreign currencies: Unit: Baht per 1 unit	real	อัตราแลกเปลี่ยนเงินตรา (ดอลลาร์)
6	latest:oil_Dollar	real	ราคาน้ำมันดิบล่าสุด (ดอลลาร์)
7	opening price:oil_Dollar	real	ราคาเปิดน้ำมันดิบ (ดอลลาร์)
8	Max:oil_Dollar	real	ราคาน้ำมันดิบสูงสุด (ดอลลาร์)
9	min:oil_Dollar	real	ราคาน้ำมันดิบต่ำสุด (ดอลลาร์)
10	Quantity:oil_Dollar	Integer	ปริมาณราคาน้ำมันดิบ (ดอลลาร์)
11	% change_oil	real	เปอร์เซ็นต์เปลี่ยนน้ำมันดิบ
12	latest:gold_Dollar	real	ราคาทองคำล่าสุด (ดอลลาร์)
13	opening price:gold_Dollar	real	ราคาเปิดทอง (ดอลลาร์)

14	Max:gold_Dollar	real	ราคาทองสูงสุด (ดอลลาร์)
15	min:gold_Dollar	real	ราคาทองต่ำสุด (ดอลลาร์)
16	Quantity:gold_Dollar	Integer	ปริมาณราคาทอง (ดอลลาร์)
17	% change_gold	real	เปอร์เซ็นต์เปลี่ยนทอง

4.4 สร้างแบบจำลอง (Modeling)

ในขั้นตอนนี้ นำข้อมูลมาวิเคราะห์ตามอัลกอริทึม โดยใช้เทคนิคการทำเหมืองข้อมูล ด้วยเทคนิคการจำแนกประเภทข้อมูล (Classification) ทั้งหมด 4 เทคนิค ประกอบด้วย เทคนิคการถดถอยเชิงเส้น (Linear Regression) เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors : K-NN) เทคนิคโครงข่ายประสาทเทียม (Neural Network) และเทคนิคต้นไม้ตัดสินใจ (Decision Tree) ด้วยโปรแกรม Rapid Miner Studio

4.4.1 เทคนิคการถดถอยเชิงเส้น (Linear Regression) เป็นการศึกษาในรูปแบบความสัมพันธ์ของตัวแปร เพื่อใช้ทำนายค่าตัวแปรที่ต้องการศึกษา โดยอาศัยความรู้เกี่ยวกับค่าตัวแปรอื่นที่เกี่ยวข้อง (ตัวแปรที่กำหนด) ซึ่งอาจจะมีหนึ่งตัวแปรหรือมากกว่า ตัวแปรที่ต้องการศึกษาเรียกว่าตัวแปรตาม (Dependent variable) แทนด้วยตัวแปร Y ตัวแปรที่กำหนดให้ เรียกว่าตัวแปรอิสระ (Independent variable) แทนด้วยตัวแปร X

- ตัวแปรอิสระ (Independent variable) เป็นตัวแปรที่กำหนดค่าได้แน่นอน หรือ สามารถวัดค่าได้โดยไม่มีคลาดเคลื่อน
 - ตัวแปรตาม (Dependent variable) เป็นตัวแปรที่เกิดขึ้นโดยสุ่ม
- รูปแบบพื้นฐานที่ง่ายที่สุดของการวิเคราะห์การถดถอยโดยมีตัวแบบการถดถอย จากสูตรที่ 1 ดังนี้

$$\hat{Y} = +b_0 + b_1x_1 + \varepsilon$$

ตัวแบบการถดถอยเชิงพหุคูณ (Multiple Linear Regression Model) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรหลายตัว(k) กับตัวแปรตาม 1 ตัว เพื่อศึกษาว่ามีตัวแปรอิสระใดบ้างที่ร่วมกันพยากรณ์ สมมติมีตัวแปรอิสระ k ตัว คือ x_1, x_2, \dots, x_k ที่มีอิทธิพลต่อ Y ในรูปแบบเส้นตรง ดังนั้นตัวแบบถดถอยของ Y เมื่อกำหนด x_1, x_2, \dots, x_k จากสูตรที่ 2 ดังนี้

$$\hat{Y} = +b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \varepsilon$$

เรียกรูปแบบความสัมพันธ์ดังกล่าวนี้ว่า ตัวแบบถดถอยแบบเชิงเส้นตรงเชิงพหุของประชากร (Population multiple linear regression model) สามารถเขียนตัวแบบในรูปแบบของข้อมูลที่เก็บรวบรวมมา จากสูตรที่ 3 ดังนี้

$$\hat{Y} = +b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_kx_k + e$$

โดยที่ \hat{Y} คือ ค่าที่คาดการณ์ได้จากแบบจำลองการถดถอยเชิงเส้น (Linear Regression)

b_0 คือ ค่าคงที่หรือตัวตัดแกน (Intercept)

$b_1 - b_k$ คือ ค่าสัมประสิทธิ์ (Coefficients) ของตัวแปรอิสระ

$x_1 - x_k$ คือ ค่าของตัวแปรอิสระแต่ละตัวที่ส่งผลต่อ \hat{Y}

e คือ ค่าความคลาดเคลื่อนหรือ Residual (Error Term)

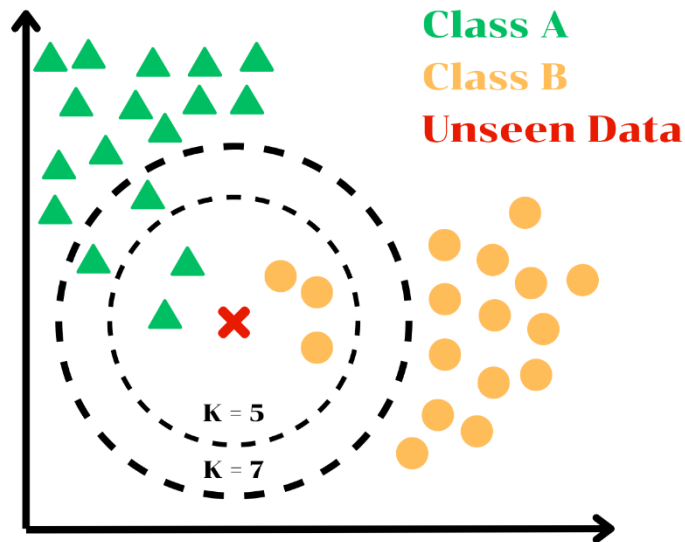
4.4.2 เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors : K-NN) เป็นวิธีการจำแนกประเภทข้อมูลวิธีหนึ่งที่ได้รับคามนิยมอย่างมาก ซึ่งสามารถนำไปประยุกต์ใช้งานได้หลากหลายลักษณะคือ ไม่มีการสร้างโมเดลในการคล้ายคลึงกันมาจำแนกประเภทข้อมูลใหม่ได้โดยอาศัยการตรวจสอบค่า k แบบไขว้ (Cross Validation) ใช้การคำนวณหาระยะห่างระหว่างข้อมูลตัวอย่างที่สนใจกับข้อมูลอื่นๆซึ่งมีระยะห่างกันมาก แสดงว่ามีความคล้ายกันน้อย ทุกตัวระยะห่างยูคลิเดียน (Euclidian Distance) จากสูตรที่ 4 ดังนี้

$$d(i, j) = \sqrt{\sum_{n=1}^N (x_{i,n} - x_{j,n})^2}$$

โดยที่ $d(i, j)$ คือ ระยะห่างระหว่างจุดที่ i และ j

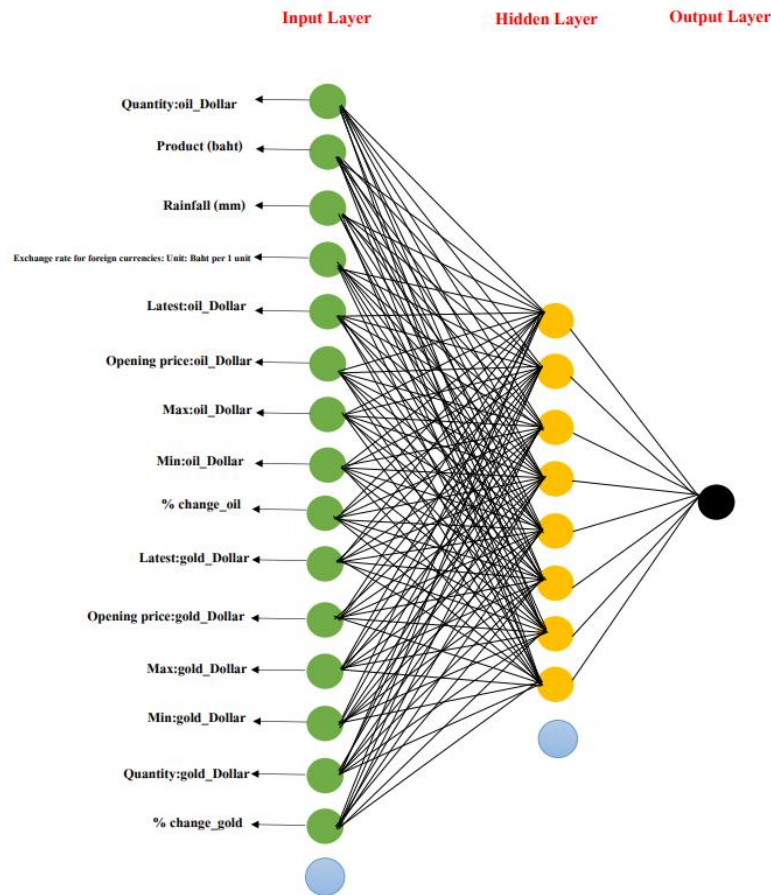
$x_{i,n}$ คือ ค่าของจุด i ในมิติที่ n

$x_{j,n}$ คือ ค่าของจุด j ในมิติที่ n



ภาพที่ 4 k-NN Algorithm Concept

4.4.3 เทคนิคโครงข่ายประสาทเทียม (Neural Network) โครงข่ายประสาทเทียมประกอบด้วยหลายชั้น (layers) ซึ่งแต่ละชั้นประกอบไปด้วยโหนด (nodes) หรือเรียกว่าเซลล์ประสาท (neurons) ซึ่งแต่ละโหนดจะรับข้อมูลจากโหนดในชั้นก่อนหน้าและส่งผลลัพธ์ไปยังโหนดในชั้นถัดไป โหนดในชั้นแรกที่รับข้อมูลเรียกว่า Input Layer และโหนดในชั้นสุดท้ายที่ส่งผลลัพธ์ออกมาเรียกว่า Output Layer ซึ่งแต่ละโหนดใน Output Layer จะแสดงผลลัพธ์ในรูปแบบของค่าน้ำหนัก (weights) ของแต่ละหมวดหมู่ (class) โดยมีค่าน้ำหนักเพิ่มขึ้นตามความเหมาะสมของโมเดลในการจำแนกแต่ละหมวดหมู่ นอกจาก Input Layer และ Output Layer ยังมี Hidden Layer ซึ่งเป็นชั้นที่อยู่ระหว่าง Input Layer และ Output Layer ซึ่งชั้นนี้จะมีผลต่อประสิทธิภาพของโมเดลเป็นอย่างมาก การเพิ่มหรือลดจำนวนชั้นและโหนดในแต่ละชั้นจะมีผลต่อการประมวลผลและประสิทธิภาพของโมเดลด้วย การสร้างโครงข่ายประสาทเทียมยังต้องระบุฟังก์ชัน Activation ที่ใช้ในการคำนวณผลลัพธ์ของแต่ละโหนดในโครงข่าย และการเลือกฟังก์ชัน Activation ที่เหมาะสมสามารถทำให้โมเดลมีประสิทธิภาพมากขึ้นด้วย โดยมีลักษณะโครงสร้าง ดังภาพที่ 5



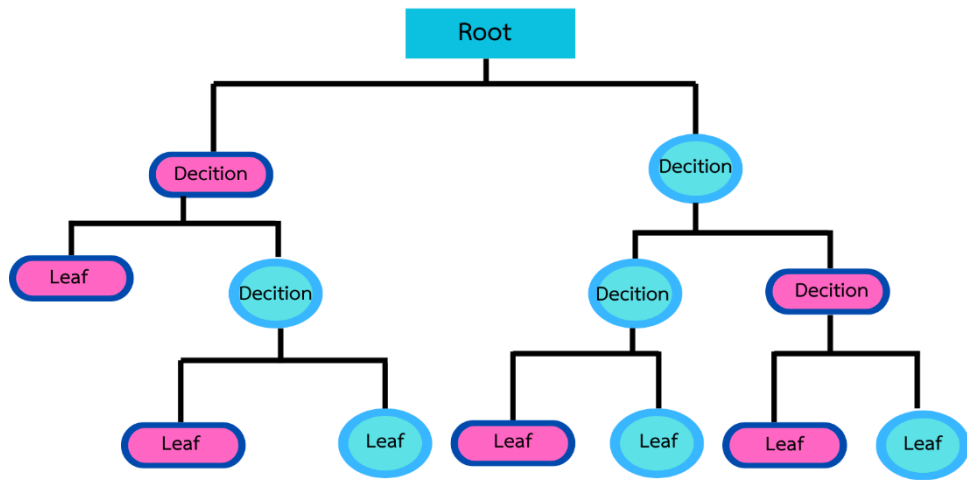
ภาพที่ 5 neural network algorithm

4.4.4 เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เป็นการจำแนกโดยการแบ่งข้อมูลออกเป็นแขนงสาขา ซึ่งค่อนข้างสอดคล้องกับการตัดสินใจจริงของมนุษย์และถูกแสดงเป็นภาพที่เข้าใจง่าย และสามารถทำงานได้กับข้อมูลที่มีความเกี่ยวข้องกันของข้อมูลสูง และอาจมีหลายลำดับชั้น โดยมีอัลกอริทึมที่นิยมนำมาใช้ในการสร้างแผนการตัดสินใจ (Decision Tree) ซึ่งเป็นการสร้างแผนการตัดสินใจ (Decision Tree) จากข้อมูลที่ต้องการทำการเรียนรู้จะต้องมีการกำหนดคุณลักษณะ (Attribute) และการแบ่งโหนด (Node or split) ของตัวแบบแผน (Tree) โดยใช้หลักการต้นไม้ในการค้นหาคุณลักษณะที่ดีที่สุดจากบนลงล่างมาสร้างเป็น โหนดราก (Root Node) ใช้ค่า Gain Ratio ที่สูงที่สุดเป็นโหนดรากและโหนดถัดไป ซึ่งค่า Entropy Information Gain และ Split Information ซึ่งค่า Information Gain (IG) หาได้จากสูตรที่ 5 ดังนี้

สูตร $IG(\text{parent}, \text{child}) = \text{entropy}(\text{parent}) - [p(c_1) \times \text{entropy}(c_1) + p(c_2) \times \text{entropy}(c_2) + \dots]$

โดยที่ $\text{entropy}(c_1) = -p(c_1) \log p(c_1)$

$p(c_1)$ คือ ค่าความน่าจะเป็นของ c_1



ภาพที่ 6 Decision Tree Classification Algorithm

4.5 การประเมินผล (Evaluation)

หลังจากที่ได้สร้างตัวแบบขึ้นมาได้แล้ว ขั้นตอนถัดมาจะต้องทำการวัดประสิทธิภาพการจำแนกประเภทข้อมูลของตัวแบบที่สร้างได้ ก่อนที่จะนำตัวแบบไปใช้ในการพยากรณ์ สำหรับงานวิจัยนี้ ผู้วิจัยได้ใช้เกณฑ์ในการวัดค่าความแม่นยำของตัวแบบซึ่งมี 5 วิธี ประกอบด้วย

4.5.1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) เป็นการวัดค่าความคลาดเคลื่อน โดยการนำค่าความคลาดเคลื่อนมา ยกกำลังแล้วนำไปหาค่าเฉลี่ย ในการวัดค่าความแม่นยำจากวิธีการนี้ยิ่งค่าที่ได้มีค่าน้อยแสดงว่าตัวแบบที่ได้จะมีความแม่นยำมาก จากสูตรที่ 6

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

\hat{Y}_t คือ ค่าที่ได้จากการพยากรณ์ที่เวลา t ใด

4.5.2 ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Squared Error: RMSE) เป็นวิธีการวัดค่าความคลาดเคลื่อนแบบมาตรฐาน ซึ่งนิยมใช้กันอย่างแพร่หลาย โดยมีสมการที่ใช้ในการวัดค่าความแม่นยำจากวิธีการนี้ยิ่งค่า RMSE ที่ได้ค่าน้อยแสดงว่าตัวแบบที่ได้จะมีความแม่นยำมาก จากสูตรที่ 7

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

\hat{Y}_t คือ ค่าที่ได้จากการพยากรณ์ที่เวลา t ใด

4.5.3 ค่าคลาดเคลื่อนสัมบูรณ์ (Absolute Error) ค่าคลาดเคลื่อนสัมบูรณ์ (Absolute Error: AE) การนำค่าความคลาดเคลื่อนมาหาค่าสัมบูรณ์ จากสูตรที่ 8

$$AE = |Y_t - \hat{Y}_t|$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

\hat{Y}_t คือ ค่าที่ได้จากการพยากรณ์ที่เวลา t ใด

4.5.4 ค่าความคลาดเคลื่อนยกกำลังสอง (Square Error) ค่าคลาดเคลื่อนยกกำลังสอง (Square Error: SE) คือการนำค่าความคลาดเคลื่อนมายกกำลังสอง จากสูตรที่ 9

$$SE = (|Y_t - \hat{Y}_t|)^2$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

\hat{Y}_t คือ ค่าที่ได้จากการพยากรณ์ที่เวลา t ใด

4.5.5 สัมประสิทธิ์การตัดสินใจ (Coefficient of Determination: R^2) สัมประสิทธิ์การตัดสินใจ (Coefficient of Determination: R^2) คือ สัดส่วนที่ตัวแปรอิสระสามารถอธิบายการเปลี่ยนแปลงของตัวแปรตามได้ ดังนั้นถ้า R^2 มีค่ามากแสดงว่าตัวแปรตามและตัวแปรอิสระมีความสัมพันธ์กันมากหรือแสดงว่าตัวแปรอิสระของสมการความถดถอยนั้นสามารถอธิบายการเปลี่ยนแปลงของตัวแปรตามได้มาก จากสูตรที่ 10

$$R^2 = 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \bar{Y})^2}$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่เวลา t ใด ๆ

\hat{Y}_t คือ ค่าที่ได้จากการพยากรณ์ที่เวลา t ใด

4.6 การนำไปใช้ (Deployment)

เมื่อได้ทำการวิเคราะห์ครบ 5 ขั้นตอนเสร็จเรียบร้อยแล้ว จะได้ผลลัพธ์ที่เหมาะสมสำหรับการสร้างกลยุทธ์ในการส่งออกมะพร้าวไทยในอนาคต โดยการใช้ข้อมูลจากการวิเคราะห์และการพยากรณ์เพื่อประเมินแนวโน้มและความต้องการของตลาดต่างประเทศ เพื่อให้สามารถปรับกลยุทธ์การผลิตและการจัดจำหน่ายมะพร้าวไทยได้อย่างมีประสิทธิภาพ การพยากรณ์นี้จะเป็นประโยชน์ในการวางแผนการส่งออกมะพร้าวไทย และสนับสนุนนโยบายการพัฒนาตลาดต่างประเทศในอนาคต พร้อมกับการคำนึงถึงการตอบสนองต่อความต้องการและแนวโน้มที่เกิดขึ้นในตลาดโลก

5. ผลการศึกษา

ผู้วิจัยได้ทำการเปรียบเทียบค่าที่ทดสอบประสิทธิภาพของตัวแบบการพยากรณ์จำนวนการให้บริการผู้ป่วยจิตเวช ด้วยเทคนิคการเรียนรู้ของเครื่องทั้ง 4 เทคนิค ได้แก่ เทคนิคการถดถอยเชิงเส้น (Linear Regression) เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) เทคนิคโครงข่ายประสาทเทียม (Neural Network) และเทคนิคต้นไม้ตัดสินใจ (Decision Tree) และทำการทดสอบประสิทธิภาพของการประมาณค่าด้วยค่า ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error) ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Squared Error) ค่าความคลาดเคลื่อนสัมบูรณ์ (Absolute Error) ค่าความคลาดเคลื่อนกำลังสอง (Squared Error) และให้ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ (R -Squared) ซึ่งผลการวิเคราะห์ข้อมูลได้ผลแสดงดังตารางที่

ตารางที่ 3 เปรียบเทียบค่าทดสอบประสิทธิภาพของการจำแนกประเภทข้อมูลของตัวแบบเชิงพยากรณ์ สำหรับปริมาณการส่งออกมะพร้าวในประเทศไทยทั้ง 4 ชนิด

เทคนิคการ จำแนก ประเภท ข้อมูล	ค่าทดสอบประสิทธิภาพของการประมาณค่าประเภทข้อมูลของตัวแบบเชิงพยากรณ์				
	Mean Square	Root Mean	Absolute	Squared Error	R ²
	Error	Squared	Error		
		Error			
Linear Regression	54,191,237,807, 710.26	7,361,469. 813	5,746,810. 570	58,377,200,021, 927.810	73.5%
K-Nearest Neighbors*	38,688,625,970, 446.50	6,216,200. 551	4,612,294. 570	45,260,457,821, 710.650	88.8%
Neural Network	44,889,809,371, 802.38	6,699,985. 774	5,359,367. 378	50,693,163,274, 153.240	86.9%
Decision Tree	52,906,116,936, 372.17	7,266,404. 465	5,895,114. 404	58,612,508,571, 301.350	76.9%

*คือ เทคนิคที่มีความเหมาะสมสำหรับนำมาสร้างตัวแบบการพยากรณ์ปริมาณการส่งออกมะพร้าวในประเทศไทย

จากตารางที่ 3 พบว่าเทคนิคที่มีความเหมาะสมที่สุดสำหรับการสร้างตัวแบบพยากรณ์จำนวนการให้บริการผู้ป่วยจิตเวช คือ เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) โดยให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย Mean Square Error เท่ากับ 38,688,625,970,446.50 ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง Root Mean Squared Error เท่ากับ 6,216,200.551 ค่าความคลาดเคลื่อนสัมบูรณ์ Absolute Error เท่ากับ 4,612,294.570 ค่าความคลาดเคลื่อนกำลังสอง Squared Error เท่ากับ 45,260,457,821,710.650 ซึ่งมีค่าความคลาดเคลื่อนน้อยที่สุดเมื่อเทียบกับเทคนิคอื่นๆ และให้ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ R² มีค่าเท่ากับ 0.888 หรือ 88.80% เนื่องจากเทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) มีค่าความคลาดเคลื่อนต่ำที่สุด ทำให้การพยากรณ์แม่นยำที่สุด รองลงมาคือ (1) เทคนิคโครงข่ายประสาทเทียม (Neural Network) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย Mean Square Error เท่ากับ 44,889,809,371,802.38 ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง Root Mean Squared Error เท่ากับ 6,699,985.774 ค่าความคลาดเคลื่อนสัมบูรณ์ Absolute Error เท่ากับ 5,359,367.378 ค่าความคลาดเคลื่อนกำลังสอง

Squared Error เท่ากับ 50,693,163,274,153.240 และให้ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ R^2 มีค่าเท่ากับ 0.869 หรือ 86.90% (2) เทคนิคต้นไม้ตัดสินใจ (Decision Tree) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย Mean Square Error เท่ากับ 52,906,116,936,372.17 ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง Root Mean Squared Error เท่ากับ 7,266,404.465 ค่าความคลาดเคลื่อนสัมบูรณ์ Absolute Error เท่ากับ 5,895,114.404 ค่าความคลาดเคลื่อนกำลังสอง Squared Error เท่ากับ 58,612,508,571,301.350 และให้ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ R^2 มีค่าเท่ากับ 0.769 หรือ 76.90% (3) เทคนิคการถดถอยเชิงเส้น (Linear Regression) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย Mean Square Error เท่ากับ 54,191,237,807,710.26 ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง Root Mean Squared Error เท่ากับ 7,361,469.813 ค่าความคลาดเคลื่อนสัมบูรณ์ Absolute Error เท่ากับ 5,746,810.570 ค่าความคลาดเคลื่อนกำลังสอง Squared Error เท่ากับ 58,377,200,021,927.810 และให้ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ R^2 มีค่าเท่ากับ 0.735 หรือ 73.50% ตามลำดับ

6. อภิปรายและสรุปผล

จากการเปรียบเทียบประสิทธิภาพของตัวแบบที่ใช้การเปรียบเทียบประสิทธิภาพของตัวแบบเชิงพยากรณ์สำหรับปริมาณการส่งออกมะพร้าวในประเทศไทย โดยมีการหาค่าที่เหมาะสมที่สุด (Optimization) พบว่า เทคนิคเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbors) โดยให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย Mean Square Error เท่ากับ 38,688,625,970,446.50 ค่ารากที่สองของความคลาดเคลื่อนเฉลี่ยกำลังสอง Root Mean Squared Error เท่ากับ 6,216,200.551 ค่าความคลาดเคลื่อนสัมบูรณ์ Absolute Error เท่ากับ 4,612,294.570 ค่าความคลาดเคลื่อนกำลังสอง Squared Error เท่ากับ 45,260,457,821,710.650 ซึ่งมีค่าความคลาดเคลื่อนน้อยที่สุดเมื่อเทียบกับเทคนิคอื่นๆ และให้ค่าสัมประสิทธิ์การตัดสินใจพหุคูณ R^2 มีค่าเท่ากับ 0.888 หรือ 88.80% จึงเหมาะสำหรับนำมาเปรียบเทียบประสิทธิภาพของตัวแบบที่ใช้การเปรียบเทียบประสิทธิภาพของตัวแบบเชิงพยากรณ์สำหรับปริมาณการส่งออกมะพร้าวในประเทศไทย ซึ่งพบว่าแตกต่างจากงานวิจัยของ Kavitha M G, Raviprakash M L, Kassem AL-Attabi, B. Sathyavani, S Meenakshi Sundaram การวิเคราะห์เชิงพยากรณ์โดยใช้ Modified Inception Net และ Support Vector Machine สำหรับการทำนายผลผลิตมะพร้าวแห้งในเกษตรกรกรม โดยใช้เทคนิคการเรียนรู้ของเครื่อง 2 ประเภท ได้แก่ Modified Inception Net และ Support Vector Machine จากผลการทดลอง พบว่า Modified Inception Net ให้ค่า สัมประสิทธิ์การตัดสินใจพหุคูณ R^2 สูงที่สุดที่ 94.10% เนื่องจากมีตัวแปรที่ใช้ในการศึกษาแตกต่างกัน ซึ่งการกำหนดตัวแปรต้นและตัวแปรตามให้มีความสัมพันธ์กัน และผู้วิจัยไม่ได้มีการหาค่าพารามิเตอร์ที่เหมาะสมในการสร้างแบบจำลอง

ดังนั้น คุณค่าจากการวิจัยนี้ จึงเหมาะสำหรับนำไปสร้างตัวแบบที่ใช้การเปรียบเทียบประสิทธิภาพของตัวแบบเชิงพยากรณ์สำหรับปริมาณการส่งออกมะพร้าวในประเทศไทย ซึ่งเป็นประโยชน์ต่อภาครัฐ ภาคธุรกิจ และเกษตรกรในการวางแผนการผลิตและการส่งออกให้เหมาะสมกับสถานการณ์ตลาดโลก

ทั้งนี้ การวิจัยนี้สามารถนำข้อมูลที่ได้ไปประยุกต์ใช้เพื่อการพัฒนาในด้านการพยากรณ์จำนวนการส่งออกมะพร้าวในอนาคตอาจต้องใช้ข้อมูลหลายตัวแปร เช่น การเปลี่ยนแปลงของสภาพภูมิอากาศที่อาจส่งผลกระทบต่อการผลิตมะพร้าว หรือการเปลี่ยนแปลงในความต้องการของตลาดต่างประเทศ ที่อาจมีปัจจัยทั้งทางเศรษฐกิจและสังคมเข้ามามีบทบาท

7. กิตติกรรมประกาศ

งานวิจัยนี้สำเร็จได้ด้วยการเอื้อเฟื้อข้อมูลที่เป็นประโยชน์จากการรวบรวมข้อมูลจากเว็บไซต์ <https://www.doae.go.th/> (ข้อมูลปริมาณการส่งออกมะพร้าวและมูลค่าการส่งออกมะพร้าว), <http://climate.tmd.go.th/> (ข้อมูลปริมาณน้ำฝน), <https://app.bot.or.th/> (ข้อมูลอัตราแลกเปลี่ยนเงินตราดอลลาร์) และ <https://th.investing.com/> (ข้อมูลน้ำมันดิบและทองคำ) ที่ได้นำความรู้ที่มีความน่าสนใจมาวิเคราะห์ในงานวิจัยครั้งนี้ และขอขอบพระคุณ ดร.อนุพงศ์ สุขประเสริฐ อาจารย์ประจำสาขาวิชาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม ที่ให้ความอนุเคราะห์ชี้แนะแนวทางให้คำแนะนำต่างๆ อันเป็นประโยชน์และมอบโอกาส และให้การสนับสนุนในการทำงานวิจัยในครั้งนี้ จนทำให้สำเร็จลุล่วงไปได้ด้วยดี คณะผู้วิจัยหวังเป็นอย่างยิ่งว่างานวิจัยนี้จะก่อให้เกิด ประโยชน์ทั้งในภาครัฐ ภาคธุรกิจ เกษตรกร และผู้ที่สนใจทุกท่าน

8. ข้อเสนอแนะ

8.1 การเตรียมข้อมูลมีความสำคัญมากในการทำเหมืองข้อมูล ผู้วิจัยควรศึกษาข้อมูลให้ครบถ้วนและทำความเข้าใจในข้อมูล ความสำคัญของแอตทริบิวต์ต่าง ๆ ที่มีความสำคัญในการนำมาวิเคราะห์ข้อมูลในการทำเหมืองข้อมูล

8.2 ในการทำตัวแบบโมเดลต้องเตรียมความพร้อมของข้อมูลเป็นอย่างดีให้พร้อมต่อการประมวลผลการกำหนดตัวแปรต้นและตัวแปรตามให้มีความสัมพันธ์และถูกต้อง การกำหนดชนิดของข้อมูลให้ถูกต้องกับ ข้อมูลเพื่อจะทำให้ไม่มีปัญหาในการประมวลผลตัวแบบสำหรับการพยากรณ์

8.3 ควรศึกษาเกี่ยวกับข้อมูลอื่นที่ใช้ในการทำเหมืองข้อมูลที่อาจไม่เพียงพอต่อการทำนาย เพื่อให้ สามารถนำไปประยุกต์ใช้ในการเพิ่มประสิทธิภาพของงานวิจัยได้มากยิ่งขึ้น

9. เอกสารอ้างอิง

- [1] Praornpit Katchwattana, “มะพร้าวไทย กับโอกาสเติบโตในเวทีการค้าโลก”, salika [ออนไลน์]. เว็บไซต์: <https://www.salika.co/2024/02/26/thai-coconut-in-world-trading/> (เข้าถึงเมื่อ 22 มีนาคม 2568).
- [2] “โอกาสของมะพร้าวไทยในตลาดจีน”, กรมส่งเสริมการค้าระหว่างประเทศ กระทรวงพาณิชย์ [ออนไลน์]. เว็บไซต์: <https://www.ditp.go.th/post/185050> (เข้าถึงเมื่อ 22 มีนาคม 2568).
- [3] อนุพงศ์ สุขประเสริฐ, คู่มือการทำเหมืองข้อมูลด้วยโปรแกรม RapidMiner Studio, พิมพ์ครั้งที่ 5. มหาสารคาม, 2563.
- [4] “ข้อมูลตลาดและแนวโน้มการค้าสินค้าเกษตร”, องค์การอาหารและเกษตรแห่งสหประชาชาติ (FAO) [ออนไลน์]. เว็บไซต์: <https://www.fao.org/home/en/> (เข้าถึงเมื่อ 22 มีนาคม 2568).
- [5] กรมส่งเสริมการเกษตร, “การส่งออกมะพร้าวของไทย”, [ออนไลน์]. เว็บไซต์: <https://www.doae.go.th/> (เข้าถึงเมื่อ 22 มีนาคม 2568).
- [6] กรมอุตุนิยมวิทยา, “ปริมาณน้ำในประเทศไทย”, [ออนไลน์]. เว็บไซต์: <http://climate.tmd.go.th/> (เข้าถึงเมื่อ 22 มีนาคม 2568).
- [7] ธนาคารแห่งประเทศไทย, “อัตราแลกเปลี่ยนเฉลี่ยของธนาคารพาณิชย์ในกรุงเทพมหานคร (2545-ปัจจุบัน)”, [ออนไลน์]. เว็บไซต์: <https://app.bot.or.th/> (เข้าถึงเมื่อ 22 มีนาคม 2568).
- [8] Investing, “สัญญาซื้อขายล่วงหน้าน้ำมันดิบ WTI และสัญญาซื้อขายล่วงหน้าทองคำ”, [ออนไลน์]. เว็บไซต์: <https://th.investing.com/> (เข้าถึงเมื่อ 22 มีนาคม 2568).

[9] Kavitha M G, Raviprakash M L, Kassem AL-Attabi, B. Sathyavani, S Meenakshi Sundaram, “Predictive Analysis Using Modified Inception Net and Support Vector Machine for Forecasting Copra Yield in Agriculture,” การคำนวณแบบกระจายและวงจรไฟฟ้าและอิเล็กทรอนิกส์ (ICDCECE), ปีที่ 2024, DOI: 10.1109/ICDCECE60827.2024.10548986

[10] บัณฑิตา พิณดิษฐ์, รินลดา กองพิธิ, กฤษณพล พูลกลาง, ปิ่นปรางกรม ทับทิม, สิริมา วัฒนศิริ, วีรยาภรณ์ ยอดจันทร์, อนุพงศ์ สุขประเสริฐ. (2025). การเปรียบเทียบประสิทธิภาพของตัวแบบสำหรับการคัดกรองผู้ป่วยโรคถุงน้ำในรังไข่หลายใบ Comparison of Model Efficiency Screening Patients with Polycystic Ovary Syndrome. มหาสารคาม: มหาวิทยาลัยมหาสารคาม