

**การคาดการณ์ปริมาณผลผลิตเงาะในภูมิภาคเหนือ และ ใต้ ของประเทศไทยโดย
เปรียบเทียบปริมาณน้ำฝน พื้นที่เพาะปลูกต่อไร่ ผลผลิตต่อไร่ ค่าความชื้น และ
อุณหภูมิ ในปี (2561-2562)**

Estimating the amount of rambutan harvest of the northern and southern
shields by analyzing the amount of Plantation area , the yield per rai
,Production ,Humidity and temperature in (2018-2019)

ทรงพล สิงขรรัตน์

Songpon Singkhonrat

สาขาวิทยาการคอมพิวเตอร์และนวัตกรรมข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏสวนสุนันทา

Department of Computer Science and Data Innovation

Faculty of Science and Technology

Suan Sunandha Rajabhat University

บทคัดย่อ

การศึกษานี้มีวัตถุประสงค์เพื่อประมาณการผลผลิตของเงาะในภูมิภาคเหนือและใต้ของประเทศไทยโดย
การวิเคราะห์ข้อมูลเกี่ยวกับพื้นที่ปลูก ผลผลิตต่อไร่ การผลิต ค่าความชื้น และอุณหภูมิจากปี พ.ศ. 2561-2562
การวิจัยใช้ข้อมูลเกี่ยวกับพื้นที่ปลูก ผลผลิตต่อไร่ การผลิต ค่าความชื้น และอุณหภูมิเพื่อสร้างโมเดลที่ทำนาย
ผลผลิตของเงาะในภูมิภาคเหล่านี้ การวิเคราะห์พิจารณาปัจจัยต่าง ๆ เช่น พื้นที่ปลูก ผลผลิตต่อไร่ การผลิต สภาพ
พื้นที่ ค่าความชื้น และการแปรผันของอุณหภูมิ ซึ่งเป็นปัจจัยสำคัญสำหรับผลผลิตทางการเกษตร โดยการรวม
ผลผลิต การวิจัยให้ข้อมูลเชิงลึกเกี่ยวกับผลผลิตที่อาจเกิดขึ้นในภูมิภาคต่างๆของประเทศไทย เพื่อช่วยในการ
วางแผนการเกษตรและการจัดสรรทรัพยากรให้เหมาะสม

คำสำคัญ : พยากรณ์ ผลผลิตเงาะ ใต้ เหนือ

Abstract

This study aimed to estimate the rambutan harvest in the northern and southern regions of Thailand by analyzing plantation area, yield per rai, production, ground level, and temperature data from 2018 to 2019. The research utilized data on plantation area, yield per rai, production, ground level, and temperature to create models predicting rambutan harvest in these regions. The analysis considered various factors such as plantation area, yield per rai, production, ground level conditions, and temperature variations, which are crucial for agricultural productivity. By integrating these factors, the study provides insights into potential rambutan harvests in different regions of Thailand, aiding agricultural planning and resource allocation.

Keyword : Predict Rambutan Production South North

บทนำ

การเก็บเกี่ยวเงาะ (Nephelium lappaceum) เป็นหนึ่งในผลไม้เมืองร้อนที่สำคัญที่ปลูกในประเทศไทย มีบทบาทสำคัญในเศรษฐกิจและเศรษฐกิจเกษตรกรรมของประเทศ ภูมิภาคทางภูมิศาสตร์ที่หลากหลายของประเทศ โดยเฉพาะภูมิภาคเหนือและใต้ มีการเสนอเงื่อนไขอากาศและภูมิประเทศที่แตกต่างกัน มีผลต่อการผลิตเงาะที่แตกต่างกันอย่างมีนัยสำคัญ การเข้าใจเกี่ยวกับดีเอ็นเอของการเก็บเกี่ยวเงาะในภูมิภาคเหล่านี้มีความสำคัญสำหรับการวางแผนการเกษตร เปลี่ยนที่และความรู้สู่มั่นคงทางเศรษฐกิจ การศึกษานี้เน้นการประมาณการการเก็บเกี่ยวเงาะในภูมิภาคเหนือและใต้ของประเทศไทยในช่วงปี พ.ศ. 2561-2562 การประมาณการนี้ขึ้นอยู่กับการวิเคราะห์อย่างละเอียดของหลายปัจจัยเช่น พื้นที่ปลูก ผลผลิตต่อไร่ การผลิต สภาพพื้นที่ ความชื้น และการแปรผันของอุณหภูมิ โดยการตรวจสอบข้อมูลเหล่านี้ การวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาโมเดลที่สามารถให้ข้อมูลเชิงลึกเกี่ยวกับผลผลิตที่เป็นไปได้ของเงาะในภูมิภาคที่เกี่ยวข้องโดยใช้ 2 โมเดลในการเปรียบเทียบคือ การวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis) และ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด หรือ K-Nearest Neighbors โครงการนี้ มีเงื่อนไขสภาพแวดล้อมที่แตกต่างกัน ซึ่งเช่นการแปรผันของอุณหภูมิ ระบบฝนและประเภทดิน ที่มีผลโดยตรงต่อผลผลิตการเกษตร นอกจากนี้ ความแตกต่างในการปฏิบัติการปลูกและเทคนิคการจัดการยังมีส่วนร่วม

ในการทำให้มีความแตกต่างในผลผลิตของเงาะระหว่างภูมิภาคเหล่านี้ ดังนั้น การวิเคราะห์อย่างละเอียดที่คำนึงถึงปัจจัยเหล่านี้เป็นสิ่งจำเป็นสำหรับการประมาณการการเก็บเกี่ยวเงาะอย่างแม่นยำ นอกจากนี้ ด้วยความต้องการเพิ่มขึ้นของเงาะทั้งในประเทศและต่างชาติ การทำให้มีผลผลิตอย่างสม่ำเสมอและเหมาะสมเป็นสิ่งสำคัญ โดยใช้ข้อมูลประวัติศาสตร์และเทคนิคการวิเคราะห์ขั้นสูง การศึกษานี้มุ่งหวังที่จะให้ข้อมูลที่มีคุณค่าเกี่ยวกับปัจจัยที่นำมาซึ่งผลผลิตเงาะในภูมิภาคเหนือและใต้ของประเทศไทย ข้อมูลเหล่านี้สามารถใช้ในการตัดสินใจของนักวิชาการด้านการเกษตร นักปฏิบัติการเกษตร และผู้เกี่ยวข้องในการจัดการแปลงเงาะ

1.วัตถุประสงค์การวิจัย

1.1 การวิจัยนี้มีวัตถุประสงค์หลักเพื่อให้ข้อมูลเชิงลึกเกี่ยวกับปัจจัยที่มีผลต่อการผลิตเงาะในประเทศไทย โดยเฉพาะปริมาณน้ำฝน, พื้นที่เพาะปลูกเงาะ ผลผลิตต่อไร่ ความชื้น และอุณหภูมิ ในปี 2018 และ 2019 เพื่อนำข้อมูลนี้มาช่วยในการจัดทรัพยากรและกลยุทธ์ที่ส่งเสริมอุตสาหกรรมเงาะในประเทศไทย

1.2 การวิจัยนี้มีเป้าหมายเพื่อพัฒนาระบบคาดการณ์ผลผลิตเงาะที่มีความแม่นยำในประเทศไทย โดยใช้ข้อมูลปริมาณน้ำฝน, พื้นที่เพาะปลูก, ผลผลิตต่อไร่ ความชื้น และอุณหภูมิ จากยุคก่อนเป็นข้อมูลพื้นฐาน การพัฒนาระบบคาดการณ์นี้จะช่วยการเกษตรกรในการวางแผนการผลิต, การจัดทรัพยากร, และการลดความเสี่ยง ทั้งนี้เพื่อเพิ่มประสิทธิภาพในการเพาะปลูกเงาะและเพิ่มผลผลิตในอนาคตให้มีความคงทนและสร้างรายได้สูงขึ้นสำหรับทั้งพื้นที่และประเทศในเครือข่ายรางวัลเกษตรกรอย่างยั่งยืน.

2.เอกสารและงานวิจัยที่เกี่ยวข้อง

ชนิษฐา กุลนาวิน และคณะ (2565) การออกแบบแบบจำลองการพยากรณ์ผลผลิตมันสำปะหลังในเขตพื้นที่จังหวัดนครราชสีมาด้วยเทคนิคเหมืองข้อมูล (Applied Data Mining for forecasting cassava yield in Nakhon Ratchasima) เป็นหัวข้อของบทความวิชาการซึ่งเป็นผลงานวิจัยที่มีวัตถุประสงค์เพื่อออกแบบแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังในพื้นที่เขตพื้นที่ปลูกของจังหวัดนครราชสีมา โดยการใช้เทคนิคเหมืองข้อมูล (Data Mining) เพื่อทำนายผลผลิตอย่างมีประสิทธิภาพ วัตถุประสงค์ของงานวิจัยที่จะออกแบบแบบจำลองการพยากรณ์โดยใช้ข้อมูลของผลผลิตมันสำปะหลังในจังหวัดนครราชสีมาในช่วงปี พ.ศ. 2555-2559

ซึ่งประกอบด้วยข้อมูลทั้งหมด 8 คุณลักษณะ หรือตัวแปรที่สำคัญ ผลจากขั้นตอนการคัดเลือกคุณลักษณะที่สำคัญจากข้อมูล (Feature selection) จากนั้นได้นำข้อมูลที่ได้มาสร้างแบบจำลองการพยากรณ์โดยใช้เทคนิคเหมือนข้อมูล โดยมีการเลือกใช้โมเดลการพยากรณ์ทั้งหมด 4 แบบได้แก่ Decision Tree, Artificial Neural Network, Naive Bayes, และ Support Vector Machine เพื่อทำการเปรียบเทียบประสิทธิภาพของแต่ละโมเดล โดยมีวัตถุประสงค์ที่จะหาโมเดลที่มีประสิทธิภาพสูงสุดในการทำนายผลผลิตมันสำปะหลังในพื้นที่ปลูกของจังหวัดนครราชสีมา ผลลัพธ์จากงานวิจัยนี้ได้แสดงให้เห็นถึงความแม่นยำและประสิทธิภาพของแบบจำลองการพยากรณ์ โดยพบว่าโมเดล Naive Bayes มีประสิทธิภาพสูงสุดในการทำนายผลผลิตมันสำปะหลังในพื้นที่ปลูกของจังหวัดนครราชสีมา โดยมีค่าความแม่นยำ (accuracy) เฉลี่ยอยู่ที่ 0.673, ค่าความแม่นยำในการทำนายผลลัพธ์ที่เชื่อถือได้ (precision) อยู่ที่ 0.606, ค่าความระลึก (recall) อยู่ที่ 0.685 และค่า F-measure อยู่ที่ 0.629 นอกจากนี้ยังมีการวัดประสิทธิภาพของโมเดลโดยใช้หลักการของ Root Mean Square Error และ Mean Absolute Error ซึ่งได้ค่าเป็น 0.303 และ 0.201 ตามลำดับ นำมาเปรียบเทียบกับโมเดลอื่น ๆ ที่ใช้ในการทำนายผลผลิตมันสำปะหลังในพื้นที่เดียวกัน บทความนี้มีความสำคัญในการวิจัยด้านการเพิ่มประสิทธิภาพในการทำนายผลผลิตเพื่อช่วยเกษตรกรหรือผู้ที่เกี่ยวข้องในอุตสาหกรรมเกษตรและการผลิตมันสำปะหลังได้รับข้อมูลและข้อมูลที่ถูกต้องเพื่อการวางแผนการผลิตที่มีประสิทธิภาพสูงขึ้น ซึ่งเป็นประโยชน์อย่างมากต่อการพัฒนาเศรษฐกิจและการเกษตรในพื้นที่ดังกล่าว

ปิยะพร แซ่ลิ้ม และคณะ (2562) บทความวิจัยเกี่ยวกับการพยากรณ์ผลผลิตข้าวหอมมะลิในจังหวัดนครราชสีมาที่มีความสำคัญเนื่องจากข้าวหอมมะลิเป็นพืชเศรษฐกิจที่สำคัญของประเทศไทย โดยทั้งการบริโภคภายในประเทศและการส่งออกไปยังตลาดโลกมีความสำคัญมาก การพยากรณ์ผลผลิตจะช่วยให้ผู้ที่เกี่ยวข้องสามารถวางแผนการผลิตได้อย่างเหมาะสมตามแนวโน้มของสถานการณ์ทางเศรษฐกิจและอุตสาหกรรมเกษตร เช่น การวางแผนการเพาะปลูกและการตลาดสินค้าข้าวได้ให้ผลผลิตที่มีประสิทธิภาพและการดำเนินงานที่มีประสิทธิผลมากขึ้น ในการวิจัยดังกล่าว ได้ใช้เทคนิคการพยากรณ์เชิงสถิติด้วยวิธีการถดถอยพหุคูณ (Multiple Regression Analysis) และวิธีการวิเคราะห์อนุกรมเวลา (Time Series Analysis) โดยใช้วิธีบอกซ์-เจนกินส์ (Box-Jenkins method) เพื่อพยากรณ์ผลผลิตข้าวหอมมะลิในจังหวัดนครราชสีมา จากการเปรียบเทียบผลลัพธ์จากวิธีการพยากรณ์ทั้งสองวิธี พบว่า การใช้วิธีการถดถอยพหุคูณ (Multiple Regression Analysis) โดยใช้วิธีการถดถอยแบบลำดับขั้น (Stepwise Regressions) เป็นวิธีที่เหมาะสมที่สุดสำหรับการพยากรณ์ผลผลิตข้าวหอมมะลิ

ในจังหวัดนครราชสีมา ผลลัพธ์ของการพยากรณ์ในปี 2560 แสดงให้เห็นถึงแนวโน้มของการเพิ่มขึ้นของปริมาณผลผลิตข้าวหอมมะลิในจังหวัดนครราชสีมา ประมาณ 7.429% จากปี 2559 ซึ่งเป็นข้อมูลที่มีประโยชน์สำหรับการวางแผนการผลิตข้าวในอนาคตและการบริหารจัดการในภาคการเกษตร โดยการพยากรณ์นี้จะช่วยให้เกษตรกรหรือผู้ที่เกี่ยวข้องสามารถดำเนินการตลาดข้าวได้อย่างมีประสิทธิภาพและทันสถานการณ์ทางเศรษฐกิจอย่างเหมาะสม และช่วยให้ผู้บริหารภาคการเกษตรสามารถตัดสินใจเกี่ยวกับนโยบายและการวางแผนการดำเนินงานในอนาคตได้อย่างมีความมั่นใจ

คมกริช พรหมหากุล และคณะ (2562) รูปแบบการเปลี่ยนแปลงและคาดการณ์ผลผลิตอ้อย ด้วยข้อมูลภาพถ่ายจากดาวเทียม Landsat และภาพถ่ายจากอากาศยานไร้คนขับรายงานนี้ใช้วิธีการแจกแจงของมอราแน (Moran's I distribution method) เพื่อวิเคราะห์รูปแบบการเปลี่ยนแปลงของพื้นที่และโมเดลการคาดการณ์แบบถดถอยอย่างง่าย (simple regression model) เพื่อการคาดการณ์ผลผลิตของอ้อย โดยเฉพาะ: วิธีการแจกแจงของมอราแน (Moran's I Distribution Method): เทคนิคนี้ถูกใช้เพื่อประเมินความสัมพันธ์พื้นที่ซึ่งระบบการจำแนกพื้นที่ทำให้สามารถรู้จักได้ว่าค่าของตัวแปรหนึ่งๆ มีการจัดกลุ่มรวมกันอยู่หรือกระจายอย่างไรในพื้นที่ มันช่วยในการเข้าใจรูปแบบและความสัมพันธ์ทางพื้นที่ภายในตัวแปรที่ศึกษา โมเดลการคาดการณ์แบบถดถอยอย่างง่าย (Simple Regression Model): ใช้โมเดลการถดถอยอย่างง่ายเพื่อประมาณค่าผลผลิตของอ้อย โดยใช้ความสัมพันธ์ระหว่างดัชนีสีเขียวเกินจำเป็น (Excessive Green Index - ExG) และจำนวนจริงของต้นอ้อยและผลผลิตที่เกี่ยวข้อง โมเดลการถดถอยอย่างง่ายนี้ถูกนำมาใช้เพื่อการคาดการณ์ค่าผลผลิตของอ้อย การวิเคราะห์ดังกล่าวนี้มีผลทำให้สามารถเข้าใจรูปแบบการเปลี่ยนแปลงของพื้นที่และการคาดการณ์ผลผลิตของอ้อย ซึ่งเป็นปัจจัยสำคัญที่ช่วยในการวางแผนการใช้ประโยชน์ที่ดินและการพัฒนาอุตสาหกรรมน้ำตาลอ้อยให้เข้าสู่ระดับพาณิชย์ได้อย่างมีประสิทธิภาพ

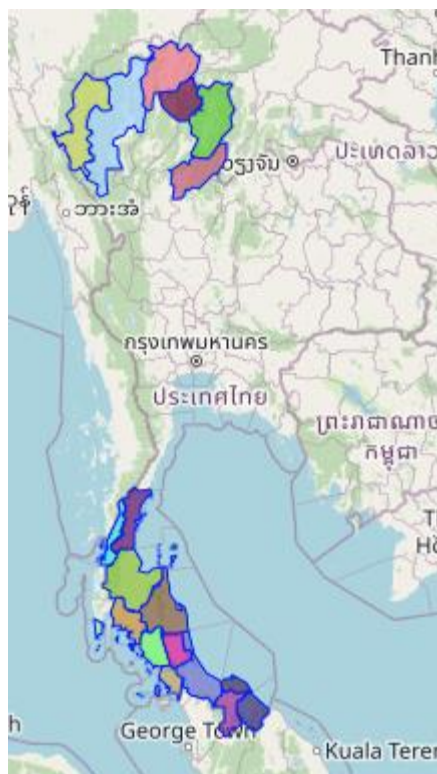
กรสิริณัฐ โรจนวรรณ และคณะ (2565) การพยากรณ์ข้อมูลอนาคตของข้อมูลอนุกรมเวลาด้วยเครื่องจักรการเรียนรู้ในวิจัยนี้ใช้โมเดลการเรียนรู้ของเครื่อง (Machine Learning) เพื่อทำการพยากรณ์ (Forecasting) ข้อมูลอนุกรมเวลา (Time Series Data) โดยมีวัตถุประสงค์เพื่อสร้างรูปแบบการพยากรณ์ที่มีประสิทธิภาพสูงสุด โดยใช้ข้อมูลอนุกรมเวลาหลายตัวแปรในการสอน (Multivariable Time Series Data for Training) โมเดลของเครื่องจักรการเรียนรู้ (Machine Learning Model) เพื่อพยากรณ์ข้อมูลในอนาคต โดยตัววิจัยได้คัดเลือกใช้เทคนิคของ

เครื่องจักรการเรียนรู้ที่แตกต่างกันอย่างไรร้อยครั้ง ซึ่งได้แก่: Moving Average (MAV): เทคนิคที่ใช้ค่าเฉลี่ยของค่าก่อนหน้าเพื่อทำการพยากรณ์ค่าใหม่ K-Nearest Neighbor (KNN): เทคนิคที่ใช้ค่าเฉลี่ยของค่าใกล้เคียงที่มีระยะห่างต่ำสุดเพื่อทำการพยากรณ์ค่าใหม่ Adaptive Neuro Fuzzy Inference Systems (ANFIS): โมเดลที่ใช้ระบบผสมระหว่างเครือข่ายประสาทเทียมและระบบความรู้ทางระบบฟัซซีลอจิกเพื่อทำการพยากรณ์ Multilayer Perceptron (MLP): โมเดลเชิงลึกที่ใช้เครือข่ายประสาทเทียมหลายชั้นเพื่อทำการพยากรณ์ ตัววิจัยในงานนี้ได้ใช้ข้อมูลน้ำตามระดับน้ำของโครงการพระราชดำริลุ่มน้ำปากพนัง ในช่วงเวลา 7 ปี โดยมีตัวแปรทั้งหมด 12 ตัวแปร โดยใช้ตัวแปร 1 ถึง 11 เป็นข้อมูลนำเข้า (input) ในการพยากรณ์ตัวแปรที่ 12 ซึ่งเป็นผลลัพธ์ของการวิจัยที่เป็นไปในทิศทางของการพยากรณ์และการเรียนรู้ของเครื่องจักรการเรียนรู้

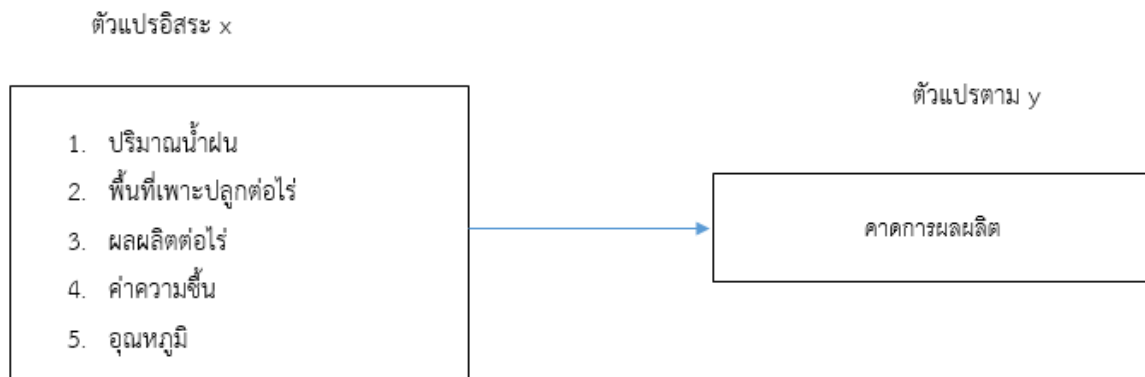
นรวัฒน์ เหลืองทอง และคณะ (2559) วิจัยนี้มุ่งเน้นการพยากรณ์ผลผลิตการเกษตรของพืช 4 ชนิดที่มีผลผลิตสูงสุดในประเทศ ได้แก่ ข้าวนาปี ข้าวนาปรัง มันสำปะหลัง และสับปะรด โดยการเปรียบเทียบวิธีการพยากรณ์เชิงสาเหตุ 3 วิธีคือ: วิธีการวิเคราะห์การถดถอยเชิงเส้น (Linear Regression Analysis): วิธีที่ใช้สร้างโมเดลเชิงเส้นเพื่อสร้างความเข้าใจในความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามเพื่อทำนายผลผลิต ระบบผสมของขั้นตอนวิธีทางพันธุกรรมและการวิเคราะห์การถดถอยเชิงเส้น (Genetic Algorithm-Linear Regression Analysis Hybrid System): วิธีการที่ใช้การค้นหาตัวแปรและพารามิเตอร์ที่เหมาะสมด้วยขั้นตอนวิธีทางพันธุกรรมร่วมกับการวิเคราะห์การถดถอยเชิงเส้น วิธีโครงข่ายประสาทเทียม (Artificial Neural Network): วิธีที่ใช้โครงข่ายประสาทเทียมเพื่อจำลองความซับซ้อนและความสัมพันธ์ที่ซับซ้อนระหว่างข้อมูลของตัวแปรอิสระและตัวแปรตามเพื่อทำนายผลผลิต ผลการศึกษาพบว่าวิธีการโครงข่ายประสาทเทียมมีค่าร้อยละความคลาดเคลื่อนสัมบูรณ์เฉลี่ยต่ำสุดในทุกชนิดของพืชที่ศึกษา ซึ่งค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ยของข้าวนาปีอยู่ในช่วง 4.07 ถึง 8.56 เปอร์เซ็นต์ ข้าวนาปรัง 6.51 ถึง 9.30 เปอร์เซ็นต์ มันสำปะหลัง 4.46 ถึง 9.57 เปอร์เซ็นต์ และสับปะรด 6.91 ถึง 8.67 เปอร์เซ็นต์ ซึ่งสะท้อนถึงความแม่นยำของการพยากรณ์ที่สูงสุดในงานวิจัยนี้ การเลือกใช้โครงข่ายประสาทเทียมเป็นตัวเลือกการพยากรณ์ที่เหมาะสมที่สุดสำหรับการพยากรณ์ผลผลิตการเกษตรในสถานการณ์นี้

กรอบแนวคิดการวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเพื่อศึกษาการพยากรณ์ ผลผลิตเงาะภายในภาคใต้ และ ภาคเหนือ ภายในประเทศไทยทั้งหมด 19 จังหวัดได้แก่ เชียงใหม่ อุดรดิตถ์ น่าน พะเยา เชียงรายแม่ฮ่องสอน นครศรีธรรมราช กระบี่ ภูเก็ต สุราษฎร์ธานี ระนอง ชุมพร สงขลา สตูล ตรัง พัทลุง ปัตตานี ยะลา รยละเอียดดังรูปที่ 1 พื้นที่การศึกษา เพื่อพยากรณ์ผลผลิตเงาะ ในเขตพื้นที่ ผู้วิจัยได้ทำการออกแบบกรอบแนวคิด ที่ใช้ในการ การวิจัยครั้งนี้โดยกำหนดตัวแปรเป็นตัวแปร อิสระ และตัวแปรตาม กำหนดความสัมพันธ์แสดงรายละเอียดดังรูปที่ 2 กรอบแนวคิดในการวิจัย



รูปที่ 1 พื้นที่การศึกษาทั้ง 19 จังหวัด



รูปที่ 2 กรอบแนวคิดในการวิจัย

เปรียบเทียบประสิทธิภาพด้านความแม่นยำของโมเดลการคาดการณ์ปริมาณผลผลิตเงาะโดยใช้เทคนิคดังต่อไปนี้

1. การวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis)
2. ขั้นตอนวิธีเพื่อนบ้านใกล้สุด (K-Nearest Neighbors)

วิธีการดำเนินการวิจัย

การเตรียมความพร้อมสำหรับทำเหมืองข้อมูล (Data Mining) โดยมีขั้นตอนดังนี้

- 1 ตรวจสอบความถูกต้องของข้อมูลที่รวบรวมมาได้
- 2 ทำความเข้าใจกับข้อมูล
- 3 จัดทำคลังข้อมูล (Data Warehouse) เพื่อเตรียมพร้อมสำหรับการเหมืองข้อมูล ข้อมูลที่จัดเก็บมาจากหลายส่วน สถาบันนวัตกรรมและรรรมาภิบาลข้อมูล สำนักงานเศรษฐกิจการเกษตร กรมอุทยานแห่งชาติ สัตว์ป่า และพันธุ์พืช สถาบันสารสนเทศทรัพยากรน้ำ (องค์การมหาชน) โดยข้อมูลที่นำมาทั้งหมดเป็นข้อมูลที่มีความเกี่ยวข้องกับปี 2561-2562(2018-2019)

- | | |
|--------------------------------------|-----------------------------|
| 1. ข้อมูลปริมาณน้ำฝนทั้งประเทศไทย | 4. ค่าความชื้นทั่วประเทศไทย |
| 2. เนื้อที่เพาะปลูกเงาะทั่วประเทศไทย | 5. ค่าอุณหภูมิทั่วประเทศไทย |
| 3. เนื้อที่เก็บเกี่ยวทั่วประเทศไทย | 6. ผลผลิตเงาะทั่วประเทศไทย |

4.ทำความสะอาดข้อมูลและเนื่องจากข้อมูลที่ได้อาจมีความขาดหายหรือเมื่อ นำมา รวมกันแล้วเกิดความไม่สมบูรณ์ของข้อมูล

การวิเคราะห์ข้อมูล

1.Model ที่ใช้ในการดำเนินการวิจัยครั้งนี้จะใช้โปรแกรมสำเร็จรูป ทางสถิติในการประมวลผล จะแบ่งออกเป็น 2 ข้อใหญ่ๆ คือ การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) โดยวิธีการถดถอยแบบลำดับขั้น และการ KNN หรือ K-Nearest Neighbors ซึ่งมีรายละเอียดดังต่อไปนี้

1.1.การวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis) โดยวิธีการถดถอยแบบลำดับขั้น ในการวิเคราะห์ข้อมูลสำหรับหาตัวแบบถดถอยนั้น ในการ วิจัยครั้งนี้จะใช้วิธีการวิเคราะห์ถดถอยพหุคูณ ซึ่งเป็นการ วิเคราะห์รูปแบบความสัมพันธ์ระหว่างตัวแปรตาม บน ตัวแปรอิสระมากกว่าหนึ่งตัวแปร เราเรียกว่า “ตัวแบบ ถดถอยพหุคูณ” ซึ่งมีรูปแบบทั่วไปดังนี้

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_kx_k$$

โดยที่ตัวแปรตาม แทนด้วย y = ปริมาณผลผลิตเงาะ และ x_k คือตัวแปรอิสระซึ่งแทนปัจจัยต่างๆ ที่เกี่ยวข้อง ซึ่งในการวิจัยครั้งนี้มีปัจจัยที่นำมาศึกษาดังนี้

x_1 = ปริมาณน้ำฝนโดยเฉลี่ย

x_4 = อุณหภูมิโดยเฉลี่ย

x_2 = พื้นที่เพาะปลูก

x_5 = ความชื้นสัมพัทธ์

x_3 = ผลผลิตเก็บเกี่ยวต่อไร่

1.2. เพื่อนบ้านใกล้สุด (k Nearest Neighbor: KNN) เป็นเทคนิคที่ได้รับความนิยมเป็นอย่างมากในกลุ่มของการจำแนกประเภทเนื่องจากเป็นวิธีที่สามารถเข้าใจได้ง่ายไม่ ซับซ้อนและมีประสิทธิภาพสูง ไม่เพียงแต่การจำแนกประเภทเท่านั้นที่เทคนิคนี้สามารถใช้ได้ แต่ ยังรวมถึงถึงกลุ่มของการถดถอยเช่นกัน โดยปกติ ในงานของการจำแนกประเภทเทคนิคนี้จะทำการ จัดกลุ่มของข้อมูลใหม่ให้อยู่ในกลุ่มเดียวกับข้อมูลที่ ใกล้เคียง ค่า k จะเป็นค่าที่กำหนดจำนวนชุดข้อมูล ใกล้เคียงที่ต้องการใช้ในการจำแนกประเภท แต่ใน ส่วนของการถดถอยนั้นจะเป็นไปตามสมการ Euclidean distance

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_k - q_k)^2}$$

เริ่มต้นด้วยการกำหนด Feature ทั้งหมดจากนั้นทำการ Normalize ข้อมูลหาค่าสูงสุดและต่ำสุดในแต่ละ Feature เพื่อนำมาคำนวณค่าหาค่า k ที่ดีที่สุดในการนำมาคำนวณหาค่า Prediction ภายใน Model

2.ทำการทดลองด้วยข้อมูลที่จัดเตรียมและวิธีการที่ออกแบบ แล้วบันทึกผลการทดลองแต่ละครั้งอย่างเป็นระบบ

การประเมินโมเดล

ขั้นตอนถัดมาหลังจากที่ได้ทำการสร้าง โมเดลเป็นที่เรียบร้อยแล้วคือการประเมินโมเดล ซึ่งก็คือ ประเมินความสามารถและความแม่นยำในการ ทำงานของโมเดล เพื่อทำการตัดสินใจว่าโมเดลที่ได้ ทำการ สร้างขึ้นมานั้นเหมาะสมที่จะนำไปใช้ หรือไม่ หรือควรทำการปรับแต่งให้มีความ เหมาะสมก่อน เนื่องจาก ผู้วิจัยเลือกใช้ อัลกอริทึม การเรียนรู้ของเครื่องแบบมีผู้สอนที่สามารถจำแนก ได้เป็นสองประเภทคือ การ ถดถอยและการจำแนก ดังนั้นการประเมินโมเดลจึงมีความแตกต่างกันตาม ประเภทของการนำไปใช้ในงานวิจัยฉบับนี้ต้องการ คาดการณ์ ผลผลิตของเงาะ อัลกอริทึมประเภทของการถดถอยและจะให้ค่า ผลลัพธ์ในลักษณะของค่า ตัวเลข ดังนั้นการ ประเมินโมเดลจะการใช้การวัดระยะที่คลาดเคลื่อน (Error) ด้วยวิธีดังต่อไปนี้

1.สัมประสิทธิ์การกำหนด (Coefficient of Determination: R²)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

2.ความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

3.ค่าเฉลี่ยของผลต่างสัมบูรณ์ (Mean Absolute Error: MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

y คือค่าจริง (Actual Value)

\bar{y} คือค่าเฉลี่ยของค่าจริง

\hat{y} คือค่าทำนาย (Predicted Value)

n คือจำนวนข้อมูล

การเปรียบเทียบผลลัพธ์

โมเดลที่ถูกสร้างขึ้น จากสองเทคนิคได้แก่ การวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis) และ เพื่อนบ้านใกล้สุด (k Nearest Neighbor: KNN) จะถูกนำมาเปรียบเทียบประสิทธิภาพกันด้วย สัมประสิทธิ์การกำหนด (Coefficient of Determination: R^2) ความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE) ค่าเฉลี่ยของผลต่างสัมบูรณ์ (Mean Absolute Error: MAE) โดยจะทำการแสดงผลในรูปแบบของตารางแสดงการเปรียบเทียบ

ผลการวิจัย

จากการดำเนินงานวิจัย สามารถแสดงผลการวิจัย แบ่งเป็น 2 ส่วน ดังนี้

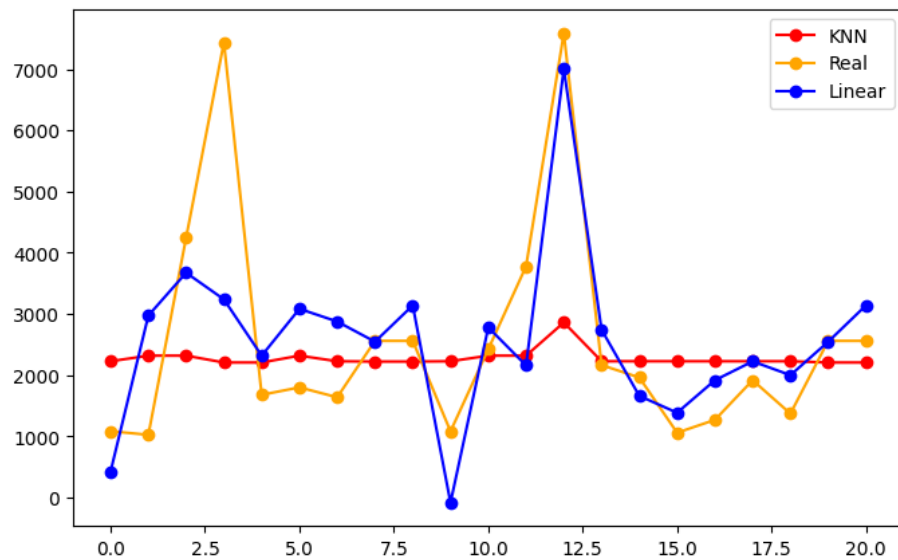
1. ผลการประเมินประสิทธิภาพของแบบจำลองการพยากรณ์ การวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis) และ เพื่อนบ้านใกล้สุด (k Nearest Neighbor: KNN)

1.1 ผลการทดลองการวิเคราะห์การถดถอยเชิงเส้น (linear regression analysis) และ เพื่อนบ้านใกล้สุด (k Nearest Neighbor: KNN) โดยการทดลองวัดผลโดยการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิค เพื่อหาความถูกต้องจาก สัมประสิทธิ์การกำหนด (Coefficient of Determination: R^2) ความผิดพลาดกำลังสองเฉลี่ย (Mean Square Error: MSE) และ ค่าเฉลี่ยของผลต่างสัมบูรณ์ (Mean Absolute Error: MAE) โดยรวมแสดงดังตารางที่ 1

ตารางที่ 1 ผลการเปรียบเทียบของทั้งสองโมเดล

Model	R^2	MSE	MAE
Linear Regression	0.33	1175894.73	950.67
KNN	-1193.35	1870353.46	1268.45

จากตารางที่ 1 เป็นการแสดงความผิดพลาดของ เทคนิคต่าง ๆ ที่นำไปใช้ในโมเดลการพยากรณ์ในการ คาดเดาผลผลิตเงาะ จะเห็นว่าจะเห็นว่า Linear Regression และ KNN มีค่า MSE และ MAE ที่แตกต่างกันซึ่ง Linear Regression มีค่า MSE และ MAE ที่น้อยกว่า KNN ทำให้เห็นได้ว่า Linear Regression มีประสิทธิภาพ ในการจำลอง Model มากกว่า KNN เปรียบเทียบดังรูป MAE จากรูปที่ 3 การเปรียบเทียบความแตกต่าง MAE ทั้งสอง Model และ ค่าจริง



รูปที่ 3 การเปรียบเทียบความแตกต่าง MAE ทั้งสอง Model และ ค่าจริง

ตารางที่ 2 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างผลผลิตเงาะและปัจจัยอื่น ๆ

ตัวแปร	ผลผลิต เงาะ	ปริมาณน้ำฝน โดยเฉลี่ย	พื้นที่เพาะปลูก	พื้นที่เก็บเกี่ยวต่อไร่	ค่าความชื้น	อุณหภูมิ
ผลผลิตเงาะ	1	0.16	0.68	-0.089	0.0072	0.19
ปริมาณน้ำฝนโดยเฉลี่ย	0.16	1	-0.077	0.003	0.49	0.6
พื้นที่เพาะปลูก	0.68	-0.077	1	-0.37	0.0036	0.043
พื้นที่เก็บเกี่ยวต่อไร่	-0.089	0.003	-0.37	1	-0.13	-0.051
ค่าความชื้น	0.0072	0.49	0.00036	-0.13	1	-0.031
อุณหภูมิ	0.19	0.6	0.043	-0.051	-0.031	1

* มีนัยสำคัญทางสถิติที่ระดับ 0.05

จากตารางที่ 2 แสดงผลลัพธ์ค่าสัมประสิทธิ์สหสัมพันธ์(Correlations) ระหว่างตัวแปรทำนายต่าง ๆ ที่ไม่มีการควบคุมตัวแปรทำนายอื่นๆ พบว่า ตัวแปรทำนาย คือ พื้นที่เพาะปลูก มีความสัมพันธ์กับตัวแปรตาม (y) ผลผลิตเงาะ สูงที่สุด ดูที่ค่า $r = 0.68$ รองลงมาคือ อุณหภูมิ $r = 0.19$ ส่วน พื้นที่เก็บเกี่ยวต่อไร่ ค่าความขึ้นปริมาณน้ำฝนโดยเฉลี่ย มีความสัมพันธ์เพียง -0.089 0.0072 และ 0.16 ตามลำดับ

2. ผลการพยากรณ์ผลผลิตเงาะ จากผลการทดลองการพยากรณ์ ผู้วิจัยจึงพิจารณาใช้เทคนิค Linear Regression ในการนำมาพัฒนา แบบจำลองการพยากรณ์ของผลผลิตเงาะ ด้วยเทคนิคเหมือนข้อมูลในเขตพื้นที่เพาะปลูก จังหวัดทั้งหมด 19 จังหวัด เพื่อใช้ในการพยากรณ์ที่จะเกิดขึ้นในอนาคต โดยมีผลของการพยากรณ์เมื่อใช้ ข้อมูลปี 2561-2562 แล้วทดสอบความเที่ยงตรงของแบบจำลองด้วยวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่ง ร้อยละ 80 ต่อร้อยละ 20สามารถทำนายผลได้ดังนี้ สัมประสิทธิ์การกำหนด = 0.33 ความผิดพลาดกำลังสองเฉลี่ย = 117589 4.73 และค่าเฉลี่ยของผลต่างสมบูรณ์ = 950.67

สรุปผลการวิจัย

จากผลการทดลองการพยากรณ์ผลผลิตเงาะด้วยเทคนิค Linear Regression ในพื้นที่เพาะปลูกของ 19 จังหวัดทั้งหมดในประเทศ พบว่าโมเดลที่พัฒนาขึ้นมีความสามารถในการทำนายผลผลิตได้ดังนี้: สัมประสิทธิ์การกำหนด (Coefficient of Determination) มีค่าเท่ากับ 0.33 ซึ่งหมายความว่าโมเดลสามารถอธิบายข้อมูลได้ประมาณ 33.% ของความเปลี่ยนแปลงในผลผลิตของเงาะ ความผิดพลาดกำลังสองเฉลี่ย (Mean Squared Error) มีค่าเท่ากับ 1,175,894.73 ซึ่งหมายความว่าค่าผลคำนวณแต่ละค่าที่ทำนายจากโมเดลกับค่าจริงมีความแตกต่างที่มาก ค่าเฉลี่ยของผลต่างสมบูรณ์ (Mean Absolute Error) มีค่าเท่ากับ 950.67 ซึ่งหมายความว่าค่าผลคำนวณแต่ละค่าที่ทำนายจากโมเดลกับค่าจริงมีความแตกต่างเฉลี่ยอยู่ที่ประมาณ 950.67 จากผลลัพธ์ดังกล่าว โมเดลการพยากรณ์ผลผลิตเงาะที่ใช้ Linear Regression ยังมีความไม่แม่นยำอยู่ในระดับที่พึงพอใจ เนื่องจากค่าสัมประสิทธิ์การกำหนดต่ำและค่าความผิดพลาดสูง อาจจำเป็นต้องพัฒนาโมเดลหรือใช้วิธีการพยากรณ์อื่นที่มีประสิทธิภาพมากขึ้นในการทำนายผลผลิตของเงาะในอนาคต

ข้อเสนอแนะ

หลังจากการวิเคราะห์ผลการพยากรณ์ผลผลิตของเงาะด้วยโมเดล Linear Regression และการประเมินผลลัพธ์ที่ได้ ข้อเสนอแนะที่สามารถเสนอได้เพื่อปรับปรุงความแม่นยำของการพยากรณ์ได้แก่: การใช้ข้อมูลเพิ่มเติม: ควรพิจารณาเพิ่มเติมหรือปรับปรุงข้อมูลที่ใช้ในการพยากรณ์ เช่น การรวบรวมข้อมูลสภาพอากาศที่

สมบูรณ์และระเบียบ เพื่อเพิ่มความแม่นยำในการทำนายผลผลิต การเลือกใช้โมเดลที่เหมาะสม: หาก Linear Regression ไม่ให้ผลลัพธ์ที่น่าพอใจ ควรพิจารณาการใช้โมเดลอื่นที่มีประสิทธิภาพมากขึ้น เช่น Random Forest, Gradient Boosting เป็นต้น ซึ่งอาจช่วยเพิ่มความแม่นยำในการพยากรณ์ได้

การปรับปรุงข้อมูล: อาจจำเป็นต้องนำเข้าข้อมูลเพิ่มเติมหรือปรับปรุงข้อมูลที่มีอยู่ เพื่อให้มีความสอดคล้องและเหมาะสมกับโมเดลที่ใช้

การวิเคราะห์ข้อผิดพลาด: ควรทบทวนและวิเคราะห์ข้อผิดพลาดจากการพยากรณ์เพื่อทำความเข้าใจถึงสาเหตุของความไม่แม่นยำ และนำข้อคิดนั้นมาปรับปรุงโมเดลต่อไป

การทดสอบและประเมินอย่างต่อเนื่อง: ควรทำการทดสอบและประเมินโมเดลอย่างต่อเนื่อง เพื่อตรวจสอบประสิทธิภาพและความแม่นยำของโมเดลในระยะยาว

เอกสารอ้างอิง

kulnawin, kanitta, โพธิ์เกษม ว. และ เลี้ยงพานิชย ศ. 2022. การออกแบบแบบจำลองการพยากรณ์ผลผลิตมันสำปะหลังในเขตพื้นที่ จังหวัดนครราชสีมาด้วยเทคนิคเหมืองข้อมูล. วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏบุรีรัมย์ (ออนไลน์). 6, 1 (พ.ค. 2022), 1-16.

นรวัฒน์ เหลืองทอง และนันทชัย กานตานันทะ* (2016) การเลือกตัวแบบพยากรณ์ผลผลิตการเกษตรที่เหมาะสม *Physical Sciences vol.24 No.3 (2016)*

กิตติบรรกุล ว., นนท์ศิริ ศ. และ คำอินทร์ พ. 2022. การพยากรณ์ความเสียหายของชิ้นส่วนเครื่องยนต์อากาศยานด้วยเทคนิคการเรียนรู้ของเครื่อง. วารสารวิชาการ ฉบับวิทยาศาสตร์และเทคโนโลยีสมาคมสถาบันอุดมศึกษาเอกชน แห่งประเทศไทย. 11, 1 (ม.ย. 2022), 1-14.

คมกริช พรหมหากุล และคณะ (2020) รูปแบบการเปลี่ยนแปลงและคาดการณ์ผลผลิตอ้อยด้วยข้อมูลภาพถ่ายจากดาวเทียม Landsat และภาพถ่ายจากอากาศยานไร้คนขับ. วารสารวิทยาศาสตร์บูรพา ฉบับ vol.25 No. 1 (2563)

ปิยะพร แซ่ลิ่ม (2019) การพยากรณ์ผลผลิตข้าวหอมมะลิ ในจังหวัดนครราชสีมา วารสารวิจัย
วิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา ปีที่ 4 ฉบับที่ 2 (กรกฎาคม - ธันวาคม 2562)

Kornsirinut Rothjanawan (2022) Forecasting the Future of Time Series Data Using
Machine Learning *Technical Education Journal KMUTNB Vol 13, No 2 (2022)*

ไพฑูรย์ ศรีนิล (2018) ระบบพยากรณ์ปริมาณผลผลิตทุเรียนพันธุ์หมอนทองบนพื้นฐานปัจจัยเด่นของ
ข้อมูลทางสภาพภูมิอากาศ มหาวิทยาลัยบูรพา วิทยาเขตจันทบุรี. คณะวิทยาศาสตร์และศิลปศาสตร์

ยุทธศาสตร์ อนุรักษิพันธุ์ และคณะ (2023) การคาดการณ์ผลผลิตอ้อยในประเทศไทยด้วยแบบจำลอง
AquaCrop วารสารวิทยาศาสตร์เกษตร ปีที่ 54 ฉบับที่ 3 (2023): กันยายน – ธันวาคม

สุวรรณแสน พ., มากมี ภ. และ ลาเต้ อ. 2020. การทำนายผลผลิตข้าวหอมมะลิ ภาค
ตะวันออกเฉียงเหนือ กรณีที่มีการสูญหายของข้อมูลโดยการประมาณค่าสูญหายจากการปรับแก้วิธีเคเนียร์เรส
เนเบอร์โดยใช้ค่าเฉลี่ยเดไซล์. วารสารวิชาการมหาวิทยาลัยอีสเทิร์นเอเชีย ฉบับวิทยาศาสตร์และเทคโนโลยี
(Online). 14, 3 (ธ.ค. 2020), 72–85.

จารินี ศานติจรรยาพร และคณะ (2017) การเลือกลักษณะสำหรับการแทนค่าข้อมูลสูญหายในการวัด
ประสิทธิภาพการผลิตปาล์มในจังหวัด สุพรรณบุรี Vol. 15 No. 2 (2560): วารสารวิทยาการวิจัยและ
วิทยาการปัญญา