

การออกแบบแบบจำลองการพยากรณ์ผลผลิตมันสำปะหลังในเขตพื้นที่
จังหวัดนครราชสีมาด้วยเทคนิคเหมืองข้อมูล
Applied Data Mining for forecasting cassava yield in Nakhon Ratchasima.

ชนิษฐา กุลนาวิน* วันเพ็ญ โพธิ์เกษม และ ศันสนีย์ เลี้ยงพานิชย์
Kanitta Kulnawin* Wanpen Pokasem and Sansanee Liangpanit
สาขาวิชาวิทยาการสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา
Informatics Program, Faculty of Sciences and Technology,
Nakhon Ratchasima Rajabhat University
E-Mail: nonone007@yahoo.com

บทคัดย่อ

บทความนี้มีวัตถุประสงค์เพื่อออกแบบแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูลในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมา ซึ่งเป็นข้อมูลย้อนหลังตั้งแต่ พ.ศ. 2555-2559 ประกอบด้วยคุณลักษณะ (attribute) 8 คุณลักษณะมาสร้างแบบจำลองการพยากรณ์ โดยผ่านกระบวนการคัดเลือกปัจจัยที่มีอิทธิพลต่อการพยากรณ์ (Feature selection) ซึ่งใช้วิธีการวิเคราะห์องค์ประกอบ จากนั้นนำไปพัฒนาแบบจำลองเพื่อการพยากรณ์ โดยใช้เทคนิคโครงสร้างต้นไม้ เทคนิคโครงข่ายประสาทเทียม เทคนิคเครือข่ายเบย์ และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน พบว่าประสิทธิภาพของแบบจำลองโดยใช้เทคนิคเครือข่ายเบย์ให้ค่าความถูกต้อง เมื่อทดสอบด้วยวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ ได้แสดงค่าความแม่นยำ ค่าความระลึกได้ และค่าการวัดประสิทธิภาพในภาพรวมที่ดีที่สุด มีค่าดังนี้ 0.673, 0.606, 0.685 และ 0.629 ตามลำดับ เมื่อเทียบกับค่าความผิดพลาดของรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และยังวัดค่าความแม่นยำ โดยใช้เกณฑ์ในการเปรียบเทียบจากรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย มีค่าดังนี้ 0.303 และ 0.201 ตามลำดับ แสดงให้เห็นว่าค่าความคลาดเคลื่อนน้อยแบบจำลองที่ได้ยังมีความแม่นยำของการพยากรณ์ดี

คำสำคัญ: พยากรณ์ ผลผลิตมันสำปะหลัง นครราชสีมา

ABSTRACT

The purpose of this research was to study forecasting models about cassava yield in Nakhon Ratchasima from the data during year 2012-2016. The data set

consisted 8 features of attributes to construct the forecasting Model. Then the features were selected by Factor Analysis. The researcher has chosen four models: Decision Tree, Artificial neural network, Naïve Bayes and Support vector machine for the comparison, with the purpose to identify the most efficiency model for the forecast of cassava yield. Finally, the model was compared by the Performance Test with a Percentage split. These results showed that Naïve Bayes model was most efficiency used as a forecasting model. Its efficiency rates for forecasting the cassava yield in Nakhon Ratchasima. Overall, the best accuracy of model is 0.673, precision is 0.606, recall is 0.685, and so on F-measure is 0.629. Moreover, performance of model was measure by using the basic of comparison: Root Mean Square Error and Mean Absolute Error. According to the performance, Root mean square Error is 0.303 and mean absolute error 0.201. That mean the smaller error is higher accurate and efficient in prediction outcome.

Keywords: forecast, cassava yield, Nakhon Ratchasima

บทนำ

จังหวัดนครราชสีมาเป็นจังหวัดที่มีแหล่งปลูกมันสำปะหลังมากที่สุดในประเทศไทยประมาณ 2 ล้านไร่ การปลูกมันสำปะหลังในจังหวัดนครราชสีมา ส่วนใหญ่จะอาศัยน้ำฝนเป็นหลัก โดยจะมีการปลูกในช่วงต้นฤดูฝน ทำให้ผลผลิตหัวสดที่ได้จะสูงกว่าการปลูกในช่วงอื่น ๆ แต่การอยู่รอดของต้นมันสำปะหลังมีน้อย ซึ่งหากปลูกในช่วงฤดูแล้งจะให้ผลผลิตสูงที่สุด โดยคุณสมบัติของดินจะมีลักษณะเนื้อดินค่อนข้างหยาบ (กรมส่งเสริมการเกษตร, 2548) ดังนั้นในการตัดสินใจเลือกช่วงการปลูกมันสำปะหลังที่เหมาะสม จึงต้องพิจารณาทั้งปริมาณน้ำฝนและลักษณะของดิน แนวโน้มของการผลิตมันสำปะหลังลดลงอาจเนื่องมาจากความผันผวนของสภาพแวดล้อมที่เปลี่ยนแปลง และสภาพภูมิอากาศที่แปรปรวน ซึ่งเป็นปัจจัยที่ไม่สามารถควบคุมได้ ทำให้บางช่วงเวลา บางพื้นที่มีสภาพแห้งแล้งและหนาวเย็นผิดปกติ ส่งผลกระทบทำให้ผลผลิตมันสำปะหลังลดลง เนื่องจากไม่สามารถคาดเดาสถานการณ์ของการปลูกมันสำปะหลังที่ดีได้ นอกจากนี้จังหวัดนครราชสีมายังไม่มีเครื่องมือที่ช่วยสนับสนุน ให้คำแนะนำ ตัดสินใจ หรือคาดการณ์ผลผลิตที่จะเกิดขึ้นในอนาคตที่ดี เพื่อให้ผู้ที่เกี่ยวข้องสามารถนำไปใช้เพื่อส่งเสริมการเกษตร และการวางแผนการผลิต เพื่อให้เกิดผลผลิตที่ดีได้

จากปัญหาดังกล่าว งานวิจัยจึงได้ทำการศึกษาและออกแบบแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูลในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมา โดยมีการ

นำเสนอแบบจำลองเพื่อการพยากรณ์โดยใช้ 4 เทคนิคดังนี้ คือโครงสร้างต้นไม้ โครงข่ายประสาทเทียม เครือข่ายเบย์ และซัพพอร์ตเวกเตอร์ แมชชีน แล้วทำการทดสอบความแม่นยำในการพยากรณ์ แล้วทำการคัดเลือกเทคนิคที่ให้ความแม่นยำดีที่สุดไปใช้ในการออกแบบแบบจำลองการพยากรณ์ โดยแบบจำลองพยากรณ์ที่จะมีความแม่นยำ เพื่อช่วยให้เกษตรกรหรือผู้เกี่ยวข้องสามารถนำไปใช้ประกอบการตัดสินใจในการบริหารจัดการผลผลิตได้อย่างมีประสิทธิภาพ โดยช่วยการพยากรณ์เพื่อใช้วางแผนการปลูกมันสำปะหลัง ซึ่งจะส่งเสริมให้ผลผลิตมันสำปะหลังสูงขึ้นโดยใช้พื้นที่เพาะปลูกเท่าเดิม ส่งผลให้ลดต้นทุนการผลิตและมีผลผลิตที่สูงขึ้น ทำให้ความเป็นอยู่ของเกษตรกรดีขึ้น

1. วัตถุประสงค์การวิจัย

เพื่อศึกษาและออกแบบแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูลในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมา

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

ปิยะพร แซ่ลิ้ม และคณะ (2562) ได้ทำการวิจัยเพื่อพยากรณ์ผลผลิตข้าวหอมมะลิจังหวัดนครราชสีมา เพื่อหาวิธีการพยากรณ์ที่เหมาะสมสำหรับพยากรณ์ผลผลิตข้าวหอมมะลิ จังหวัดนครราชสีมา เทคนิคการพยากรณ์เชิงสถิติที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ วิธีการวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) โดยวิธีการถดถอยแบบลำดับขั้น (Stepwise Regressions) และวิธีการวิเคราะห์อนุกรมเวลา (Time Series Analysis) โดยวิธีบ็อกซ์-เจนกินส์ (Box-Jenkins method) และใช้ค่าเฉลี่ยของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์เป็นเกณฑ์ในการเปรียบเทียบวิธีการพยากรณ์ทั้ง 2 วิธี ในการวิจัยครั้งนี้ ข้อมูลที่นำมาศึกษานั้น เป็นข้อมูลทุติยภูมิรายปีจากแหล่งข้อมูล 2 แหล่งคือกรมเศรษฐกิจการเกษตร และกรมอุตุนิยมวิทยา ผลการเปรียบเทียบตัวแบบพยากรณ์ที่ได้จากเทคนิคการพยากรณ์เชิงสถิติทั้ง 2 วิธี พิจารณาจากค่าเฉลี่ยของเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ พบว่าตัวแบบการถดถอยพหุคูณ (Multiple Regression Analysis) โดยวิธีการถดถอยแบบลำดับขั้น (Stepwise Regressions) เป็นตัวแบบที่เหมาะสมกับการพยากรณ์ผลผลิตข้าวหอมมะลิจังหวัดนครราชสีมา และเมื่อนำตัวแบบมาพยากรณ์ผลผลิตข้าวหอมมะลิ จังหวัดนครราชสีมา ล่วงหน้าอีก 1 คาบเวลา คือปี 2560 จากผลการพยากรณ์ สรุปได้ว่า ปริมาณผลผลิตข้าวหอมมะลิจังหวัดนครราชสีมา จะมีแนวโน้มเพิ่มสูงขึ้นจากปี 2559 ประมาณ 7.429%

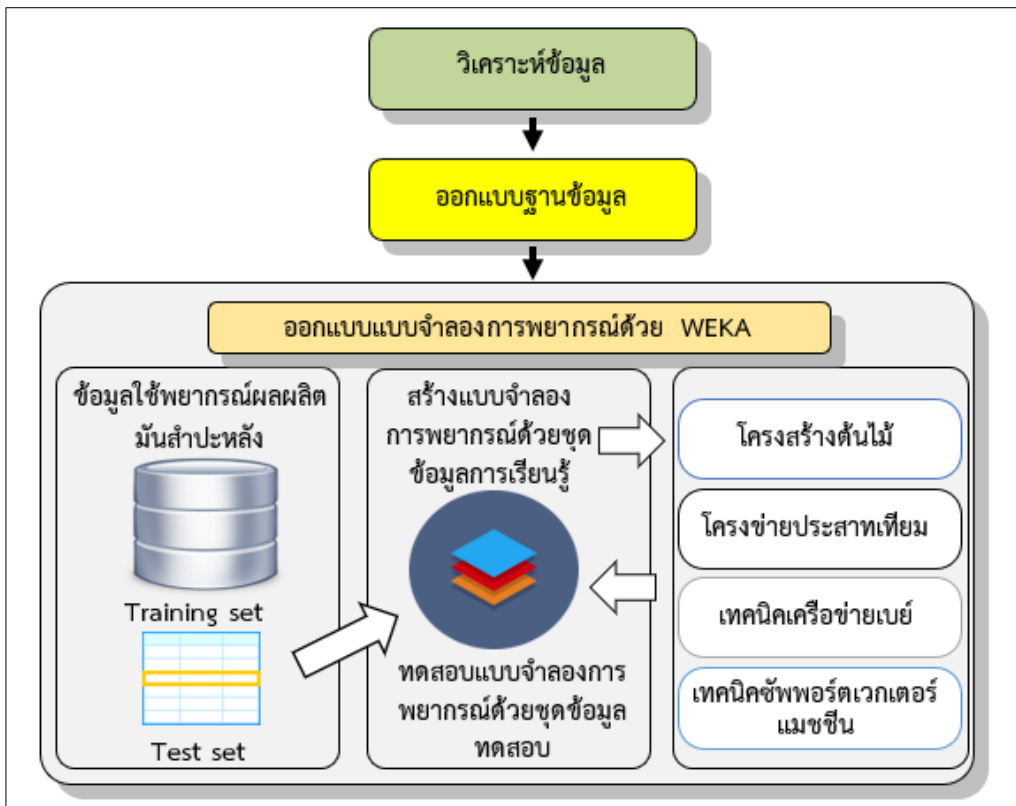
อรนงค์ บุเกตุ และพฤษดี ศิริแสงตระกูล (2556) ได้ทำการวิจัยเรื่อง การพยากรณ์ปริมาณผลผลิตอ้อยของภาคตะวันออกเฉียงเหนือ โดยการประยุกต์ใช้โครงข่ายประสาทเทียมที่มีกระบวนการเรียนรู้แบบแพร่ย้อนกลับ ในการศึกษาผู้วิจัยได้ทำการรวบรวมปัจจัยต่าง ๆ ที่มีความสัมพันธ์กับผลผลิตอ้อยจากงานวิจัยที่เกี่ยวข้องซึ่งประกอบด้วย ข้อมูลจังหวัด ปีการผลิต ข้อมูลปริมาณน้ำฝนเฉลี่ย อุณหภูมิเฉลี่ย ผลผลิตอ้อยเฉลี่ยต่อไร่ และพื้นที่ปลูกอ้อยเป็นข้อมูลนำเข้าของตัวแบบ ANN1 เพื่อยืนยันความสัมพันธ์ของปัจจัยกับปริมาณผลผลิตอ้อย ผู้วิจัยจึงได้ทำวิเคราะห์ความสัมพันธ์ของปัจจัยโดยใช้

Correlation และ Regression ผลการศึกษาพบว่าปัจจัยที่มีความสัมพันธ์ได้แก่ ข้อมูลจังหวัด ปีการผลิต ข้อมูลปริมาณน้ำฝนเฉลี่ย ผลผลิตอ้อยเฉลี่ยต่อไร่ และพื้นที่ปลูกอ้อย ซึ่งจะเป็นข้อมูลนำเข้าตัวแบบ ANN2 และจากสมมติฐานด้านพืชแข่งขันทางเศรษฐกิจของอ้อย โดยราคาพืชแข่งขันทางเศรษฐกิจ น่าจะมีผลต่อปริมาณการปลูกอ้อยในเขตพื้นที่ภาคตะวันออกเฉียงเหนือ ดังนั้นผู้วิจัยจึงได้เพิ่มปัจจัยด้านราคาพืชแข่งขันที่สำคัญ 3 ปัจจัยคือ ราคาอ้อยเฉลี่ย ราคายางพาราเฉลี่ย และราคามันสำปะหลังเฉลี่ย เมื่อรวมกับปัจจัยที่มีความสัมพันธ์ที่ได้จากการวิเคราะห์ด้วยวิธีทางสถิติ ได้แก่ ข้อมูลจังหวัด ปีการผลิต พื้นที่ปลูกอ้อย ผลผลิตอ้อยเฉลี่ยต่อไร่ และข้อมูลปริมาณน้ำฝนเฉลี่ยจะเป็นข้อมูลนำเข้าของตัวแบบ ANN3 ในการศึกษา ผู้วิจัยใช้ข้อมูลระหว่างปี พ.ศ. 2547-2552 เพื่อสร้างและเปรียบเทียบ โมเดลโครงข่ายประสาทเทียมจากข้อมูลปัจจัยนำเข้าทั้ง 3 ชุด ซึ่งผลการศึกษาพบว่าตัวแบบ ANN3 ที่มีสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบ 8:4:1 ให้ผลพยากรณ์แม่นยำสูงที่สุด มีค่าสัมประสิทธิ์สหสัมพันธ์ 0.9973 เมื่อเทียบกับ ANN1 ที่มีสถาปัตยกรรม 5:10:1 และ ANN2 ที่มีสถาปัตยกรรม 6:8:1 โดยมีค่าสัมประสิทธิ์สหสัมพันธ์ 0.9794, 0.9822 ตามลำดับ

ชนิษฐา กุลนาวิน และคณะ (2557) การศึกษาเปรียบเทียบเทคนิคเหมืองข้อมูลใช้ในการพยากรณ์ผลผลิตพืชไร่ทางการเกษตร: กรณีศึกษาข้าวไทย ได้ศึกษาเปรียบเทียบเทคนิคเหมืองเพื่อใช้ในการพยากรณ์ผลผลิตข้าวไทยที่มีความแม่นยำสูง โดยเริ่มจากการศึกษาเปรียบเทียบเพื่อหาปัจจัยที่เหมาะสมในการพยากรณ์ผลผลิตข้าว แล้วค่อยศึกษาเทคนิคของเหมืองข้อมูลเพื่อหาเทคนิคที่พยากรณ์ผลแม่นยำสูง ผลของการศึกษานี้พบว่าปัจจัยที่เหมาะสมในการพยากรณ์ ได้แก่ ข้อมูลผลผลิตทางการเกษตร ปริมาณน้ำฝน อุณหภูมิสูงสุด อุณหภูมิต่ำสุด จำนวนวันที่ฝนตก และความชื้นสัมพัทธ์ พื้นที่เพาะปลูก พื้นที่เก็บเกี่ยวโดยเป็นข้อมูลทุกจังหวัดตามชนิดของพันธุ์ข้าวและฤดูการปลูก ข้าวย้อนหลังเป็นเวลา 10 ปีตั้งแต่ ปี 2543-2552 ข้อมูลได้มาจากสำนักงานเศรษฐกิจการเกษตรและกรมอุตุนิยมวิทยา ในส่วนของการทำการเปรียบเทียบเทคนิคเหมืองข้อมูลได้พิจารณา 4 อัลกอริทึม ดังนี้ โครงข่ายประสาทเทียม การถดถอยเชิงเส้น ซัพพอร์ตเวกเตอร์แมชชีน และต้นไม้การตัดสินใจ ผลของการศึกษานี้ พบว่าเทคนิคการถดถอยเชิงเส้นเป็นวิธีที่มีประสิทธิภาพเพื่อใช้พยากรณ์ผลผลิตข้าว ซึ่งมีค่าเฉลี่ยสัมประสิทธิ์สหสัมพันธ์โดยรวมดีที่สุด ผลที่ได้จากการพยากรณ์สามารถนำมาใช้เพื่อกำหนดนโยบายการวางแผนการเพาะปลูกและการพัฒนาของการเพาะปลูกข้าวที่เหมาะสมเพื่อเพิ่มศักยภาพการทำเกษตรกรรม

วิธีดำเนินการวิจัย

ผู้วิจัยได้ทำการศึกษาและออกแบบแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูล สามารถแสดงการศึกษาและวิธีดำเนินการวิจัยได้ดังภาพประกอบ 1



ภาพประกอบ 1 ขั้นตอนการวิเคราะห์และออกแบบแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูล

1. วิเคราะห์ข้อมูล

เริ่มจากการเก็บรวบรวมข้อมูลผลผลิตมันสำปะหลัง เนื้อที่เพาะปลูก เนื้อที่เก็บเกี่ยว ของจังหวัดนครราชสีมา ส่วนปริมาณน้ำฝน จำนวนวันที่ฝนตก ความชื้นสัมพัทธ์ อุณหภูมิมีการเก็บข้อมูลของจังหวัดนครราชสีมาเช่นเดียวกัน ข้อมูลที่ทำการรวบรวมมาจาก 2 แหล่งข้อมูล คือ ข้อมูลที่ได้จากสำนักงานเศรษฐกิจการเกษตร คือผลผลิตมันสำปะหลัง เนื้อที่เพาะปลูก เนื้อที่เก็บเกี่ยว ส่วนอีกแหล่งข้อมูล คือข้อมูลที่ได้มาจากกรมอุตุนิยมวิทยา คือ จำนวนวันที่ฝนตก ปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิสูงสุดและต่ำสุดของทุกอำเภอในจังหวัดนครราชสีมา ข้อมูลทั้งหมดที่ได้มาเป็นข้อมูลย้อนหลัง 5 ปี (พ.ศ. 2555-2559) และใช้ข้อมูลปี 2559 ทำการวิเคราะห์ปัจจัยที่มีอิทธิพลต่อการพยากรณ์ผลผลิตมันสำปะหลังในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมาเพื่อทำการศึกษหาความสัมพันธ์ของโครงสร้างของตัวแปรโดยใช้วิธีการวิเคราะห์องค์ประกอบเชิงสำรวจ (Exploratory Factor Analysis : EFA) ถ้าตัวแปรมีความสัมพันธ์กันมาก หรือมีความสัมพันธ์กันอย่างมีนัยสำคัญ

(เพชรน้อย, 2549) ซึ่งถึงค่า KMO ที่คำนวณได้มีค่าน้อยกว่า 0.5 และเข้าใกล้ 0.5 ดังนั้นตัวแปรนำมาใช้ได้แก่ ผลผลิตมันสำปะหลังรายอำเภอ เนื้อที่เพาะปลูกมันสำปะหลังรายอำเภอ เนื้อที่เก็บเกี่ยวมันสำปะหลังรายอำเภอ ปริมาณน้ำฝนสูงสุดรายอำเภอ ค่าความชื้นสัมพัทธ์เฉลี่ยรายอำเภอ ค่าอุณหภูมิสูงสุดรายอำเภอ ค่าอุณหภูมิต่ำสุดรายอำเภอ และค่าความเข้มแสงเฉลี่ยรายอำเภอ

2. ออกแบบฐานข้อมูล

สามารถแสดงขั้นตอนการออกแบบฐานข้อมูล ดังนี้

1. สร้างระบบฐานข้อมูลด้วยระบบจัดการฐานข้อมูล เพื่อจัดเก็บข้อมูลที่ได้จากการสำรวจให้เป็นระเบียบ เป็นมาตรฐานเดียวกัน

2. การเตรียมความพร้อมสำหรับทำเหมืองข้อมูล (Data Mining) โดยมีขั้นตอนดังนี้

2.1 ตรวจสอบความถูกต้องของข้อมูลที่รวบรวมมาได้

2.2 ทำความเข้าใจกับข้อมูล

2.3 จัดทำคลังข้อมูล (Data Warehouse) เพื่อเตรียมพร้อมสำหรับการเหมืองข้อมูล ข้อมูลที่จัดเก็บมาจาก 2 ส่วนคือข้อมูลจากสถาบันสารสนเทศทรัพยากรน้ำและการเกษตร และส่วนปฏิบัติการข้อมูลการเกษตร ศูนย์สารสนเทศการเกษตร สำนักงานเศรษฐกิจการเกษตร โดยมีการกำหนดลักษณะของข้อมูลที่จะทำการศึกษาเป็นข้อมูลย้อนหลังของอำเภอต่าง ๆ ในจังหวัดนครราชสีมาเป็นเวลา 5 ปีย้อนหลัง ตั้งแต่ปี 2555-2559 ลักษณะของข้อมูลที่ใช้มีดังนี้

2.3.1 ผลผลิตมันสำปะหลังรายอำเภอ

2.3.2 เนื้อที่เพาะปลูกมันสำปะหลังรายอำเภอ

2.3.3 เนื้อที่เก็บเกี่ยวมันสำปะหลังรายอำเภอ

2.3.4 ปริมาณน้ำฝนสูงสุดรายอำเภอ

2.3.5 ค่าความชื้นสัมพัทธ์เฉลี่ยรายอำเภอ

2.3.6 ค่าอุณหภูมิสูงสุดรายอำเภอ

2.3.7 ค่าอุณหภูมิต่ำสุดรายอำเภอ

2.3.8 ค่าความเข้มแสงเฉลี่ยรายอำเภอ

2.4 ทำความสะอาดข้อมูลและเนื่องจากข้อมูลที่ได้มามีความขาดหายหรือเมื่อนำมารวมกันแล้วเกิดความไม่สมบูรณ์ของข้อมูล

2.5 ปรับเปลี่ยนข้อมูลให้มีรูปแบบที่เหมาะสมสอดคล้องกับวิธีการเรียนรู้ของอัลกอริทึมแบบใดแบบต่าง ๆ ซึ่งผลลัพธ์ที่ได้จะแสดงออกมาเป็นค่าระดับ หรือค่าที่เป็นช่วง เช่น การเปลี่ยนแปลงตัวเลขอุณหภูมิให้มีค่าเป็นช่วง ดังตารางที่ 1

ตารางที่ 1 แสดงตัวอย่างการปรับเปลี่ยนข้อมูลอุณหภูมิที่เป็นตัวเลขให้มีค่าเป็นช่วง

ข้อมูลอุณหภูมิเฉลี่ยสูงสุด	ช่วงผลผลิต	แปลงเป็น	ความหมาย
องศาเซลเซียส	มากกว่า 44.5	High	อุณหภูมิสูงเฉลี่ยสูง
	ต่ำกว่า 44.5	Low	อุณหภูมิสูงเฉลี่ยต่ำ

3. ออกแบบแบบจำลองการพยากรณ์ด้วย WEKA

ผู้วิจัยได้ใช้ WEKA ในการออกแบบและสร้างแบบจำลองการพยากรณ์ โดยมีขั้นตอนดังนี้

3.1 เลือกอัลกอริทึมที่เหมาะสม (Select modeling technique) อัลกอริทึมที่เลือกใช้ทดลอง (Cios et al., 2007; Han et al., 2006; Larose, 2005) มีดังนี้

3.1.1 เทคนิคโครงสร้างต้นไม้ เป็นเทคนิคที่มีการตัดสินใจโดยที่จุดเริ่มต้นของต้นไม้เรียกว่าโหนดราก แต่ละกิ่งแสดงค่าของคุณลักษณะของแต่ละโหนด และสีโหนดแสดงกลุ่ม (Class) ซึ่งเป็นผลลัพธ์ที่สามารถแยกแยะได้

3.1.2 เทคนิคโครงข่ายประสาทเทียม เป็นเทคนิคที่มาจากงานวิจัยด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) เพื่อใช้ในการคำนวณค่าฟังก์ชันจากกลุ่มข้อมูล วิธีการนี้เป็นการให้เครื่องเรียนรู้จากตัวอย่างต้นแบบ แล้วฝึก (Train) ให้ระบบได้รู้จักที่จะคิดแก้ปัญหาที่กว้างขึ้นได้ ภายในโครงข่ายมีโหนดและจุดเชื่อมต่อโครงข่าย เมื่อป้อนข้อมูลอินพุตและเอาต์พุตโครงข่ายประสาทเทียม จะทำการคำนวณและปรับค่าตัวเลขน้ำหนักเองโดยใช้กฎเกณฑ์ต่าง ๆ เข้าช่วยจนผลลัพธ์ที่ได้ออกมาถูกต้องแม่นยำอยู่ในเกณฑ์ที่น่าพอใจ

3.1.3 เทคนิคเครือข่ายเบย์ หรือเรียกโดยย่อว่า (Bayes net) เป็นวิธีที่สามารถอธิบายความสัมพันธ์ระหว่างตัวแปรและแสดงออกมา ในรูปแบบของแบบจำลองโดยอาศัยฐานความรู้ก่อนหน้า (Prior knowledge) ในการที่จะอธิบายและสร้างข่ายงานเบย์ และสามารถอธิบายความไม่ขึ้นต่อกันอย่างมีเงื่อนไข (Condition independent) ระหว่างตัวแปรบริบทของข่ายงานความเชื่อเบย์นิยมใช้คำว่า “ตัวแปร” (Variable) แทนคำว่า “คุณสมบัติ” เพื่อให้กระบวนการเรียนรู้มีประสิทธิภาพ โดยเราสามารถใส่ความรู้ก่อนในข่ายงานความเชื่อเบย์ให้อยู่ในรูปของโครงสร้างข่ายงานและตารางความน่าจะเป็นมีเงื่อนไข โดยนิยามความไม่ขึ้นต่อกันอย่างมีเงื่อนไขได้

3.1.4 เทคนิคซัพพอร์ตเวกเตอร์แมชชีน แบบจำลองแบบไม่เป็นเชิงเส้น ที่มีการเรียนรู้ตามหลักการซัพพอร์ตเวกเตอร์แมชชีนที่หาคะนาบการตัดสินใจในการแบ่งข้อมูลตามสมการเส้นตรง เพื่อแบ่งเขตข้อมูลเป็นสองกลุ่มออกจากกัน โดยจะพยายามสร้างเส้นแบ่งตรงกึ่งกลางระหว่างกลุ่มให้มีระยะห่างระหว่างขอบเขตของทั้งสองกลุ่มมากที่สุด เป็นการจำลองลักษณะของเซลล์ประสาทด้วยการใช้ Kernal function ที่เรียกว่าตัวแปรในการตัดสินใจว่าคุณสมบัติและตัวแปรที่เปลี่ยนแปลงใช้ในการกำหนดระนาบหลายมิติ เรียกว่า คุณลักษณะ (Feature) ส่วนการเลือกที่มีความเหมาะสมที่สุด เรียกว่า การคัดเลือกคุณลักษณะ (Feature selection) จำนวนเซตของคุณลักษณะที่ใช้อธิบายในกรณีหนึ่ง

(เช่น แกนของการค่าคาดการณ์) เรียกว่า เวกเตอร์ (Vector) ดังนั้นจุดมุ่งหมายของตัวแบบ svm คือ การประโยชน์สูงสุดจากระนาบหลายมิติที่แบ่งแยกกลุ่มของเวกเตอร์ ในกรณีนี้ด้วยหนึ่งกลุ่มของตัวแปร เป้าหมายที่อยู่ข้างหนึ่งของระนาบ และกรณีของกลุ่มอื่นที่อยู่ทางระนาบต่างกัน ซึ่งเวกเตอร์ที่อยู่ข้าง ระนาบหลายมิติทั้งหมดเรียกว่า ซัพพอร์ตเวกเตอร์ (Support vectors) SVM เป็นวิธีการที่สามารถ นำมาใช้ในการจำแนกรูปแบบหรือกลุ่มของข้อมูลได้ โดยจะอาศัยระนาบ มาใช้ในการแบ่งเขตของข้อมูล ออกเป็นสองฝั่ง และ Support vector machines นี้จะมีคุณลักษณะแบบ Inner-product ระหว่างตัว Support vector และ Input vector

3.2 ทำการทดลองด้วยข้อมูลที่จัดเตรียมและวิธีการที่ออกแบบ แล้วบันทึกผลการทดลองแต่ละ ครั้งอย่างเป็นระบบ

3.3 กำหนดรูปแบบการทดสอบ (Generate test design) ซึ่งทำการทดสอบด้วยวิธีวิธีการแบ่ง ข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ แล้วสร้างแบบจำลอง (Build a model) มีการทดสอบประสิทธิภาพ และความแม่นยำของแบบจำลองการพยากรณ์ เพื่อตรวจสอบค่าความผิดพลาดหรือหาค่าความถูกต้อง ในการคาดการณ์ของแบบจำลอง มีรายละเอียดดังนี้

3.3.1 ค่าความถูกต้อง (Accuracy) คือ ค่าการวัดจากการทดสอบหาค่าความถูกต้องการ พยากรณ์ของข้อมูลว่ามีความถูกต้องมากน้อยเพียงใดโดยคิดเป็นค่าร้อยละ (%)

3.3.2 ค่าความแม่นยำ (Precision) ซึ่งจะเป็นการวัดความแม่นยำของการพยากรณ์ คำนวนได้จากการหาอัตราส่วนของการพยากรณ์ว่าถูกต้องเมื่อเทียบกับข้อมูลที่ถูกต้อง

3.3.4 ค่าความระลึกได้ (Recall) เป็นการวัดจากความถูกต้องของการพยากรณ์

3.3.5 วัดประสิทธิภาพโดยรวม (F-measure) วิธีการคำนวณค่าวัดประสิทธิภาพซึ่งเกิด จากการรวมเอาค่าความแม่นยำ และค่าความระลึกได้มาคำนวณ

3.3.6 รากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root mean square error: RMSE) เป็นค่าที่ใช้ในการวัดขนาดของความคลาดเคลื่อนของการพยากรณ์ หาก RMSE มีค่าน้อย แสดงว่าแบบจำลองสามารถประมาณค่าประมาณได้ใกล้เคียงกับค่าจริง ดังนั้นหากนี้มีค่าเท่ากับศูนย์ แล้วจะหมายความว่าไม่เกิดความคลาดเคลื่อนในแบบจำลองนี้

3.3.7 ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean absolute error: MAE) เป็นค่าเฉลี่ย ของความแตกต่างสัมบูรณ์ระหว่างค่าพยากรณ์และค่าจริง หากค่า MAE มีค่าน้อย แสดงว่าแบบจำลอง สามารถประมาณค่าประมาณได้ใกล้เคียงกับค่าจริง

3.4 ดำเนินการทดลองเพื่อประเมินประสิทธิภาพของแบบจำลองด้วยเทคนิคทั้ง 4 ได้แก่ เทคนิค โครงสร้างต้นไม้ เทคนิคโครงข่ายประสาทเทียม เทคนิคเครือข่ายเบย์ และเทคนิคซัพพอร์ตเวกเตอร์ แมชชีน

3.5 เลือกเทคนิคที่เหมาะสมเพื่อพยากรณ์ผลผลิตมันสำปะหลัง

ผลการวิจัย

จากการดำเนินงานวิจัย สามารถแสดงผลการวิจัย แบ่งเป็น 2 ส่วน ดังนี้

1. ผลการประเมินประสิทธิภาพของแบบจำลองการพยากรณ์

จากการใช้เทคนิคโครงสร้างต้นไม้ เทคนิคโครงข่ายประสาทเทียม เทคนิคเครือข่ายเบย์ และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน แสดงผลได้ดังนี้

1.1 ผลการทดลองการพยากรณ์ด้วยโครงสร้างต้นไม้ ตัวแบบการพยากรณ์ที่ใช้เทคนิคนี้คือ Id3 โดยการทดลองวัดผลโดยการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิค เพื่อหาค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก ค่าการวัดประสิทธิภาพโดยรวม รากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย โดยรวมแสดงดังตารางที่ 2

ตารางที่ 2 ผลการทดลองการพยากรณ์ด้วยโครงสร้างต้นไม้

ความเที่ยงตรงของ แบบจำลอง	ค่าความ ถูกต้อง	ค่าความ แม่นยำ	ค่าความ ระลึกได้	วัดประสิทธิภาพ โดยรวม	RMSE	MAE
วิธีการใช้ชุดเรียนรู้และ ทดสอบชุดเดียวกัน	0.764	0.784	0.769	0.765	0.242	0.117
วิธีการตรวจสอบไขว้	0.658	0.644	0.658	0.644	0.312	0.153
วิธีการแบ่งข้อมูลแบบสุ่ม ด้วยการแบ่งร้อยละ	0.630	0.652	0.630	0.623	0.276	0.137

จากตารางที่ 2 จะเห็นได้ว่าความเที่ยงตรงของแบบจำลองการพยากรณ์ด้วยโครงสร้างต้นไม้เมื่อนำมาเปรียบเทียบหาความเที่ยงตรงของแบบจำลองด้วยวิธีการใช้ชุดเรียนรู้และทดสอบชุดเดียวกัน วิธีการตรวจสอบไขว้ และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ

1.2 ผลการทดลองการพยากรณ์ด้วยโครงข่ายประสาทเทียม ตัวแบบการพยากรณ์ที่ใช้เทคนิคนี้คือ Multilayer Perceptron โดยการทดลองวัดผลโดยการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิคเพื่อหาค่าความถูกต้องค่าความแม่นยำ ค่าความระลึก ค่าการวัดประสิทธิภาพโดยรวม รากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยโดยรวมแสดงดังตารางที่ 3

ตารางที่ 3 ผลการทดลองการพยากรณ์ด้วยโครงข่ายประสาทเทียม

ความเที่ยงตรงของ แบบจำลอง	ค่าความ ถูกต้อง	ค่าความ แม่นยำ	ค่าความ ระลึกได้	วัดประสิทธิภาพ โดยรวม	RMSE	MAE
วิธีการใช้ชุดเรียนรู้และ ทดสอบชุดเดียวกัน	0.769	0.770	0.769	0.760	0.247	0.129
วิธีการตรวจสอบไขว้	0.644	0.640	0.644	0.636	0.314	0.164
วิธีการแบ่งข้อมูลแบบ สุ่มด้วยการแบ่งร้อยละ	0.582	0.58	0.593	0.567	0.317	0.170

จากตารางที่ 3 จะเห็นได้ว่าความเที่ยงตรงของแบบจำลองการพยากรณ์ด้วยโครงข่ายประสาทเทียมเมื่อนำมาเปรียบเทียบกับหาความเที่ยงตรงของแบบจำลองด้วยวิธีการใช้ชุดเรียนรู้และทดสอบชุดเดียวกัน วิธีการตรวจสอบไขว้ และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ

1.3 ผลการทดลองการพยากรณ์ด้วยเทคนิคเครือข่ายเบย์ ตัวแบบการพยากรณ์ที่ใช้เทคนิคนี้คือ Naïve Bayes โดยการทดลองวัดผลโดยการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิคเพื่อหาค่าความถูกต้องค่าความแม่นยำ ค่าความระลึก ค่าการวัดประสิทธิภาพโดยรวม รากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย โดยรวมแสดงดังตารางที่ 4

ตารางที่ 4 ผลการทดลองการพยากรณ์ด้วยเทคนิคเครือข่ายเบย์

ความเที่ยงตรงของ แบบจำลอง	ค่าความ ถูกต้อง	ค่าความ แม่นยำ	ค่าความ ระลึกได้	วัดประสิทธิภาพ โดยรวม	RMSE	MAE
วิธีการใช้ชุดเรียนรู้และ ทดสอบชุดเดียวกัน	0.698	0.690	0.694	0.652	0.288	0.180
วิธีการตรวจสอบไขว้	0.658	0.617	0.663	0.592	0.307	0.195
วิธีการแบ่งข้อมูลแบบสุ่ม ด้วยการแบ่งร้อยละ	0.673	0.606	0.685	0.629	0.303	0.201

จากตารางที่ 4 จะเห็นได้ว่าความเที่ยงตรงของแบบจำลองการพยากรณ์ด้วยเครือข่ายเบย์เมื่อนำมาเปรียบเทียบกับหาความเที่ยงตรงของแบบจำลองด้วยวิธีการใช้ชุดเรียนรู้และทดสอบชุดเดียวกัน วิธีการตรวจสอบไขว้ และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ

1.4 ผลการทดลองการพยากรณ์ด้วยเทคนิคซัพพอร์ตเวกเตอร์แมชชีน ตัวแบบการพยากรณ์ที่ใช้เทคนิคนี้คือ SMO โดยการทดลองวัดผลโดยการเปรียบเทียบประสิทธิภาพของแต่ละเทคนิคเพื่อหา

ค่าความถูกต้องค่าความแม่นยำ ค่าความระลึก ค่าการวัดประสิทธิภาพโดยรวม รากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย ผลโดยรวมแสดงได้ตารางที่ 5

ตารางที่ 5 ผลการทดลองการพยากรณ์ด้วยเทคนิคซ์พอร์ตเวกเตอร์แมชชีน

ความเที่ยงตรงของ แบบจำลอง	ค่าความ ถูกต้อง	ค่าความ แม่นยำ	ค่าความ ระลึกได้	วัดประสิทธิภาพ โดยรวม	RMSE	MAE
วิธีการใช้ชุดเรียนรู้และ ทดสอบชุดเดียวกัน	0.688	0.591	0.688	0.598	0.340	0.256
วิธีการตรวจสอบไขว้	0.669	0.506	0.669	0.565	0.342	0.258
วิธีการแบ่งข้อมูลแบบสุ่ม ด้วยการแบ่งร้อยละ	0.722	0.616	0.722	0.646	0.335	0.253

จากตารางที่ 5 จะเห็นได้ว่าความเที่ยงตรงของแบบจำลองการพยากรณ์ด้วยเทคนิคซ์พอร์ตเวกเตอร์แมชชีนเมื่อนำมาเปรียบเทียบบหาความเที่ยงตรงของแบบจำลองด้วยวิธีการใช้ชุดเรียนรู้และทดสอบชุดเดียวกัน วิธีการตรวจสอบไขว้ และวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ

2. ผลการพยากรณ์ผลผลิตมันสำปะหลัง

จากผลการทดลองการพยากรณ์ ผู้วิจัยจึงพิจารณาใช้เทคนิคเครือข่ายเบย์ในการนำมาพัฒนาแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลัง ด้วยเทคนิคเหมืองข้อมูลในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมา เพื่อใช้ในการพยากรณ์ที่จะเกิดขึ้นในอนาคต โดยมีผลของการพยากรณ์เมื่อใช้ข้อมูลปี 2559 แล้วทดสอบความเที่ยงตรงของแบบจำลองด้วยวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ สามารถทำนายผลได้ดังนี้ ค่าความถูกต้อง คือ 0.417 ค่าความแม่นยำ คือ 0.506 ค่าความระลึกได้ คือ 0.455 วัดประสิทธิภาพโดยรวม คือ 0.399 RMSE คือ 0.393 MAE คือ 0.263

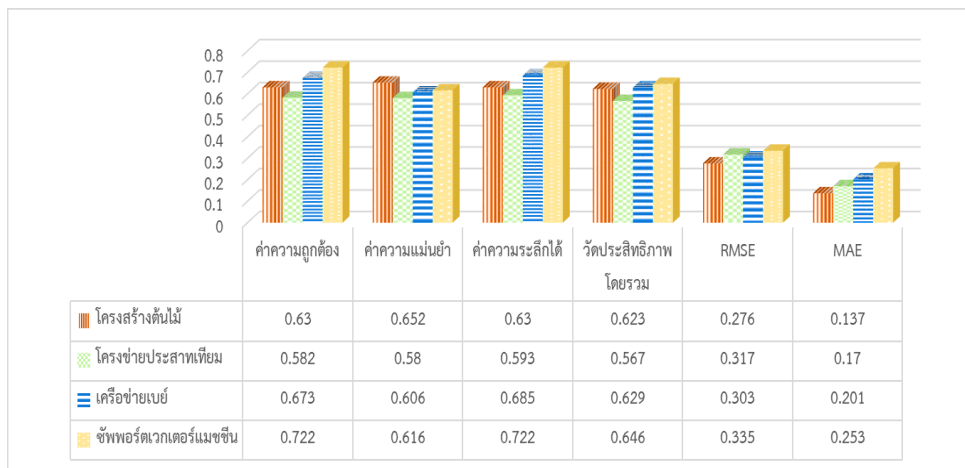
สรุปผลการวิจัย

จากการเปรียบเทียบประสิทธิภาพของแบบจำลองแต่ละเทคนิค ในภาพรวมเทคนิคซ์พอร์ตเวกเตอร์แมชชีนแสดงค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกได้ และค่าการวัดประสิทธิภาพโดยรวมมีค่าสูงสุด มีค่าดังนี้ 0.722, 0.616, 0.722 และ 0.646 ตามลำดับ แต่ความผิดพลาดของรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยก็มีค่าสูงเช่นกัน ซึ่งมีค่าดังนี้ 0.335 และ 0.253 ตามลำดับ แต่ในส่วนของเทคนิคเครือข่ายเบย์ซึ่งมีค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกได้ และค่าการวัดประสิทธิภาพโดยรวมมีค่ารองลงมา มีค่าดังนี้ 0.673, 0.606, 0.685

และ 0.629 ตามลำดับ แต่ความผิดพลาดของรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยมีค่าไม่สูงมากนัก ซึ่งมีค่าดังนี้ 0.303 และ 0.201 ตามลำดับ ดังนั้นผู้วิจัยจึงพิจารณาใช้เทคนิคเครือข่ายเบย์ ในการนำมาพัฒนาแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูลในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมา เพื่อใช้ในการพยากรณ์ที่จะเกิดขึ้นในอนาคต โดยมีผลของการพยากรณ์เมื่อใช้ข้อมูลปี 2559 แล้วทดสอบความเที่ยงตรงของแบบจำลองด้วยวิธีการแบ่งข้อมูลแบบสุ่มด้วยการแบ่งร้อยละ สามารถทำนายผลได้ดังนี้ ค่าความถูกต้องคือ 0.417 ค่าความแม่นยำ คือ 0.506 ค่าความระลึกได้ คือ 0.455 วัดประสิทธิภาพโดยรวม คือ 0.399 RMSE คือ 0.393 MAE คือ 0.263

อภิปรายผลการวิจัย

จากการวิจัยได้นำผลของข้อมูลมาวิเคราะห์ถึงการพัฒนาตัวต้นแบบในการพยากรณ์ จากข้อมูลที่ทำกรรวบรวมมาจาก 2 แหล่งข้อมูล คือ ข้อมูลที่ได้จากสำนักงานเศรษฐกิจการเกษตร คือผลผลิตมันสำปะหลัง เนื้อที่เพาะปลูก เนื้อที่เก็บเกี่ยว ส่วนอีกแหล่งข้อมูล คือข้อมูลที่ได้มาจากกรมอุตุนิยมวิทยา คือ จำนวนวันที่ฝนตก ปริมาณน้ำฝน ความชื้นสัมพัทธ์ อุณหภูมิสูงสุดและต่ำสุดของทุกอำเภอในจังหวัดนครราชสีมา ข้อมูลทั้งหมดที่ได้มาเป็นข้อมูลย้อนหลัง 5 ปี (พ.ศ. 2555-2559) โดยข้อมูลที่ใช้ทดลองมีเพียง 160 ระเบียบข้อมูลโดยทำการทดลองด้วยเทคนิคเทคนิคโครงสร้างต้นไม้ เทคนิคโครงข่ายประสาทเทียม เทคนิคเครือข่ายเบย์ และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน และมีการทดสอบเพื่อตรวจสอบค่าความผิดพลาดหรือหาค่าความถูกต้องด้วยวิธีการค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกได้ วัดประสิทธิภาพโดยรวม รากที่สองของค่าความคลาดเคลื่อนเฉลี่ย กำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย สามารถแสดงผลการทดลองได้ดังภาพประกอบ 2



ภาพประกอบ 2 แสดงผลการเปรียบเทียบการทดลองด้วยเทคนิคโครงสร้างต้นไม้ เทคนิคโครงข่ายประสาทเทียม เทคนิคเครือข่ายเบย์ และเทคนิคซัพพอร์ตเวกเตอร์แมชชีน

จากภาพประกอบ 2 แสดงผลการทดลองที่ได้จากการเปรียบเทียบประสิทธิภาพของแบบจำลองแต่ละเทคนิค ในภาพรวมเทคนิคซัพพอร์ตเวกเตอร์แมชชีนแสดงค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกได้ และค่าการวัดประสิทธิภาพโดยรวมมีค่าสูงสุด มีค่าดังนี้ 0.722, 0.616, 0.722 และ 0.646 ตามลำดับ แต่ความผิดพลาดของรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยก็มีค่าสูงเช่นกัน ซึ่งมีค่าดังนี้ 0.335 และ 0.253 ตามลำดับ แต่ในส่วนของเทคนิคเครือข่ายเบย์ซึ่งมีค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกได้ และค่าการวัดประสิทธิภาพโดยรวมมีค่ารองลงมา มีค่าดังนี้ 0.673, 0.606, 0.685 และ 0.629 ตามลำดับ แต่ความผิดพลาดของรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยมีค่าไม่สูงมากนัก ซึ่งมีค่าดังนี้ 0.303 และ 0.201 ตามลำดับ ในที่นี้จึงพิจารณาใช้เทคนิคเครือข่ายเบย์ในการนำมาพัฒนาแบบจำลองการพยากรณ์ของผลผลิตมันสำปะหลังด้วยเทคนิคเหมืองข้อมูลในเขตพื้นที่เพาะปลูกจังหวัดนครราชสีมา และเมื่อใช้ข้อมูลปี 2559 เพื่อทำการพยากรณ์ ผลของการพยากรณ์ไม่ได้แตกต่างแสดงให้เห็นว่าวิธีการนี้สามารถทำนายได้แม่นยำ

จากผลการทดลองแสดงให้เห็นว่าผลการทดลองสอดคล้องวิจัยของชุดิมา อุตมะมุณี และประสงค์ ประณีตพลกรัง (2553) ที่พัฒนาระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา นำผลลัพธ์ที่ดีที่สุดจากเทคนิคเครือข่ายเบย์มาสร้างแบบจำลองเพื่อใช้ในการพยากรณ์และยังสอดคล้องกับงานวิจัยของวิชา วิรัชกุล (2558) เพื่อการศึกษาการพัฒนาการสอบมาตรฐานด้านเทคโนโลยีสารสนเทศบัณฑิตของมหาวิทยาลัยราชภัฏนครปฐม โดยการวัดประสิทธิภาพการพยากรณ์จากการทำเหมืองข้อมูล (Data Mining) เทคนิคการจำแนก Decision Tree C4.5 , k-NN และ Naive Bayes เพื่อใช้ในการพยากรณ์ผลการสอบของนักศึกษา

โดยผลจากการศึกษาทำให้ทราบว่า การพยากรณ์ด้วยเทคนิค Naive Bayes ให้ค่าความถูกต้องในการพยากรณ์มากที่สุด

สรุปได้ว่าเทคนิคเครือข่ายเบย์ให้ค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกได้ และค่าการวัดประสิทธิภาพในภาพรวมดีที่สุด มีค่าดังนี้ 0.673, 0.606, 0.685 และ 0.629 ตามลำดับ เมื่อเทียบกับค่าความผิดพลาดของรากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง และค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยที่มีค่าไม่สูงมากนัก ซึ่งมีค่าดังนี้ 0.303 และ 0.201 ตามลำดับ

ข้อเสนอแนะ

1. ความรู้ที่ได้จากงานวิจัยนี้สามารถนำผลการวิจัยไปใช้ในการพยากรณ์เพื่อใช้ในการวางแผนการเพาะปลูกในเชิงพื้นที่ได้จะมีประโยชน์ต่อเกษตรกร และผู้ที่เกี่ยวข้อง เพื่อใช้เป็นแนวทางในการพยากรณ์ผลผลิต และในอนาคตสามารถปรับแบบจำลองการพยากรณ์สามารถนำไปใช้ในการพยากรณ์ผลผลิตมันสำปะหลังในเขตพื้นที่เพาะปลูกจังหวัดอื่นได้
2. การเลือกเทคนิคเพื่อใช้ในการสร้างแบบจำลองเพื่อการพยากรณ์นั้นควรทดสอบด้วยเทคนิคที่หลากหลาย เพื่อให้ได้แบบจำลองที่มีความเหมาะสมมากที่สุด และควรใช้ข้อมูลที่ละเอียดมากขึ้นในการสร้างแบบจำลอง

เอกสารอ้างอิง

- กรมส่งเสริมการเกษตร. (2548). *การปลูกมันสำปะหลัง*. กรุงเทพฯ : โรงพิมพ์สำนักพิมพ์การถ่ายทอดเทคโนโลยี กรมส่งเสริมการเกษตร.
- ชนิษฐา กุลนาวิน, ชัยพร เจริญพร, ปานจิตร หลงประดิษฐ์ และ ภัทรสินี ภัทรโกศลและคณะ. (2557). การศึกษาเปรียบเทียบเทคนิคเหมืองข้อมูลใช้ในการทำนายผลผลิตพืชไร่ทางการเกษตร: กรณีศึกษาข้าวไทย. *วารสารวิจัย มช.*, 19(1), 31-43.
- ชุดิมา อุดมะมุณี และ ประสงค์ ปราณีตพลกรัง. (2553). การพัฒนาตัวแบบระบบสนับสนุนการตัดสินใจแบบอัตโนมัติออนไลน์สำหรับการเลือกสาขาวิชาเรียนของนักศึกษาระดับอุดมศึกษา. *Journal of Information Science and Technology*, 1(2), 39-48.
- ปิยะพร แซ่ลิ่ม, วิวรรณ กาญจนวจิ, พัทนา สุวรรณแสน และ ณพฐ์ โสภีพันธ์ (2562). การพยากรณ์ผลผลิตข้าวหอมมะลิ ในจังหวัดนครราชสีมา. *วารสารวิจัยวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครราชสีมา*, 4(2), 25-37.
- เพชรน้อย สิงห์ช่างชัย. (2549). *หลักการและการใช้สถิติการวิเคราะห์ตัวแปรหลายตัว สำหรับการวิจัยทางการพยาบาล*. พิมพ์ครั้งที่ 3. สงขลา: ชานเมืองการพิมพ์.

- วิรัช วัชรกุล และ ธวัชชัย งามสันติวงศ์. (2558). การเปรียบเทียบประสิทธิภาพระหว่างเทคนิค Decision Tree C4.5 , k-NN และ Naive Bayes เพื่อใช้พยากรณ์ผลสอบมาตรฐานด้านเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏนครปฐม. (วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ) กรุงเทพฯ : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ)
- อรนงค์ บุเกตุ และ พุทธิศ ศิริแสงตระกูล (2556). แบบจำลองการพยากรณ์ปริมาณอ้อยของภาคตะวันออกเฉียงเหนือ โดยการใช้โครงข่ายประสาทเทียม. *วารสารวิทยาศาสตร์ มช.*, 41(1), 213-225.
- Cios, K., Pedrycz, W., Swiniarski R., and Kurgan, L. (2007). **Data Mining: A Knowledge Discovery Approach**. New York: Springer.
- Han, J., & Kamber, M. (2006). **Data Mining Concepts and Techniques**. 2nd ed. California: Morgan Kaufmann.
- Larose, D.T. (2005). **Discovery Knowledge in Data : An Introduction to Data Mining**. New York: John Wiley.