การพัฒนาตัวแบบการพยากรณ์ผลผลิตลำไยรายแปลงโดยใช้เทคนิคการทำเหมืองข้อมูล

The Development of Longan Yield Forecasting Models By Using Data Mining Techniques

วินัย บังคมเนตร[1*]

Winai Bangkhomned[1*]

Received: October 17, 2023
Revised: April 11, 2024
Accepted: April 23, 2024

Abstract: This research presents a predictive model for longan production using data mining techniques. The model utilizes longan production data from farmers in the Chiang Mai, Lamphun and Phayao provinces, totaling 215 plots in the year 2021. The decision tree technique and relationship rules were employed in the analysis. Factors found to significantly impact longan production from the forecasting model include the month of flowering induction and the type of flowering-inducing agent. It was observed that the use of sodium chlorate as a flowering-inducing agent led to medium to high predicted yields based on confidence levels in both first and second-order relationship rules. Regarding the first rule, applying sodium chlorate in December resulted in medium to high production levels, with 53 transactions occurring simultaneously and a confidence level of 62%. In the second rule, the use of sodium chlorate also resulted in medium to high yields, with 68 transactions happening concurrently and a confidence level of 62%.

Keywords: forecast, data Mining, product, longan, model for forecasting

บทคัดย่อ: งานวิจัยนี้นำเสนอตัวแบบการพยากรณ์ปริมาณผลผลิตลำไยโดยใช้เทคนิคเหมืองข้อมูลใช้ข้อมูลการผลิตลำไยในฤดูของเกษตรกรผู้ผลิตลำไยในพื้นที่จังหวัดเชียงใหม่ จังหวัดลำพูนและจังหวัดพะเยา จำนวน 215 แปลง ปี พ.ศ.2564 ด้วยเทคนิคต้นไม้ตัดสินใจ และกฎความสัมพันธ์ พบปัจจัยที่ส่งผลกระทบต่อปริมาณผลผลิตลำไยจากโมเดลการพยากรณ์คือ เดือนที่ใส่สารกระตุ้นการออกดอกและประเภทของสารกระตุ้นการออกดอกซึ่งพบว่าเมื่อใช้สารโซเดียมคลอเรตจะมีค่าพยากรณ์ปริมาณผลผลิตในระดับปานกลางขึ้นไปพิจารณาจากกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นในลำดับที่ 1 และ 2 สอดคล้องกับกฎความสัมพันธ์ของข้อมูลในกฎที่ 1 แปลงที่มีการใช้สารโซเดียมคลอเรต ในเดือนธันวาคม จำนวน 85 แปลง จะได้ผลผลิตในระดับปานกลาง 53 แปลง ที่ระดับความเชื่อมั่น 62 เปอร์เซ็นต์ และกฎที่ 2 พบว่าแปลงที่มีการใช้สารโซเดียมคลอเรต จำนวน 110 แปลง จะได้ผลผลิตในระดับปานกลาง 68 แปลง ที่ระดับความเชื่อมั่น 62 เปอร์เซ็นต์

คำสำคัญ: พยากรณ์, เหมืองข้อมูล, ผลผลิต, ลำไย, แบบจำลองเพื่อการพยากรณ์

[1] สาขานวัตกรรมธุรกิจดิจิทัล คณะบริหารธุรกิจ มหาวิทยาลัยแม่โจ้ จ.เชียงใหม่ 50290

[1] Digital Business Innovation, Faculty of Business Administration, Maejo University, 50290

*Corresponding author: winai-b@mju.ac.th

คำนำ

ลำไยเป็นไม้ผลชนิดหนึ่งที่เกษตรกรนิยมปลูกมากในภาคเหนือของประเทศไทยซึ่งประเทศไทยสามารถผลิตลำไยออกสู่ตลาดทั้งในและต่างประเทศ เนื่องจากมีสายพันธุ์ที่ดี สภาพแวดล้อมที่เหมาะสมต่อการเจริญเติบโต ผลผลิตส่วนใหญ่จะออกสู่ตลาดมากในช่วงฤดูกาลปกติ คือ ปลายเดือนมิถุนายนถึงต้นเดือนกันยายน ส่วนผลผลิตนอกฤดูจะออกสู่ตลาดมากในเดือนมกราคมและเดือนกุมภาพันธ์ ซึ่งตรงกับเทศกาลปีใหม่ ตลาดหลักในการส่งออกได้แก่ ตลาดประเทศจีน อินโดนีเซีย และฮ่องกง การผลิตลำไยของประเทศไทยในปี พ.ศ. 2563 มีเนื้อที่ให้ผลผลิต 1,583,920 ไร่ ผลผลิต 1,181,652.53 ตัน พื้นที่ส่วนใหญ่อยู่บริเวณภาคเหนือ มีเนื้อที่ให้ผลผลิต 1,155,359 ไร่ ผลผลิต 817,020 ตัน จังหวัดเชียงใหม่ ลำพูนและพะเยามีเนื้อให้ผลผลิตรวม 806,958 ไร่ เป็นร้อยละ 69.85 ของภาคเหนือ (สำนักงานเศรษฐกิจการเกษตร, 2563) ภาพรวมปัญหาของเกษตรกรส่วนใหญ่คือการใช้องค์ความรู้ดั้งเดิมที่ถ่ายทอดจากรุ่นสู่รุ่น ขาดข้อมูลทางวิชาการและขาดแนวทางในการพยากรณ์ผลผลิตที่มีประสิทธิภาพ ทำให้การวางแผนด้านต้นทุนการผลิตเป็นไปได้ยาก อีกทั้งปริมาณผลผลิตลำไยไม่แน่นอนในแต่ละปี การไม่ทราบปริมาณผลผลิตส่งผลกระทบต่อการทำการตลาด จึงควรมีการพยากรณ์ปริมาณผลผลิตลำไยที่จะออกสู่ตลาดในแต่ละปี เพื่อจะได้เตรียมพร้อมในด้านการตลาด ซึ่งจะส่งผลกระทบต่อราคาและรายได้ของเกษตรกรผู้ปลูกลำไย เป็นต้น ดังนั้นจึงมีแนวคิดที่จะใช้เทคโนโลยีสารสนเทศมาเป็นเครื่องมือในการประมวลผลข้อมูลจากรูปแบบการผลิตในอดีตเพื่อพยากรณ์ปริมาณผลผลิตล่วงหน้าโดยเปรียบเทียบจากรูปแบบการผลิตในปัจจุบัน โดยอาศัยเทคนิคเหมืองข้อมูลในการจำแนกประเภท เพื่อให้ผู้ใช้สามารถทราบแนวโน้มของปริมาณผลผลิตล่วงหน้า โดยการพัฒนาตัวแบบการพยากรณ์ผลผลิตลำไยและศึกษาปัจจัยด้านการผลิตที่ส่งผลกระทบต่อปริมาณผลผลิตลำไยในแต่ละฤดูกาลผลิต

การตรวจเอกสาร

งานวิจัยนี้ศึกษาตัวแบบการพยากรณ์ผลผลิตลำไยและศึกษาปัจจัยด้านการผลิตที่ส่งผลกระทบต่อปริมาณผลผลิตลำไยในแต่ละฤดูกาลผลิต โดยอาศัยทฤษฎีการพยากรณ์ (forecasting) เพื่อประมาณการความต้องการวัตถุดิบหรือวางแผนกลยุทธ์ส่งเสริมการตลาด,การพยากรณ์ ปริมาณผลผลิต เพื่อวางแผนการตลาด เป็นต้นโดยการพยากรณ์จะทำการศึกษาแนวโน้มและรูปแบบการเกิดเหตุการณ์จากข้อมูลในอดีตหรือใช้ความรู้ ความสามารถ ประสบการณ์ และดุลยพินิจของผู้พยากรณ์ โดยอาศัยเทคนิคเหมืองข้อมูล (data mining) ค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลเพื่อสร้างโมเดลการพยากรณ์ ซึ่งต้องผ่านการรวบรวมและเตรียมข้อมูล (Fayyad *et al*., 1996) ประกอบด้วยขั้นตอนดังนี้

1) การคัดเลือกข้อมูล (data Selection) เป็นการระบุถึงแหล่งข้อมูลที่จะนำมาใช้ในการทำเหมืองข้อมูล รวมถึงการนำข้อมูลที่ต้องการออกมาจากฐานข้อมูลเพื่อการใช้งานในเบื้องต้น

2) การกรองข้อมูล (data cleaning) เป็นกระบวนการที่ทำให้เกิดความมั่นใจในคุณภาพของข้อมูลที่จะนำมาใช้ วิเคราะห์ข้อมูลที่มีความถูกต้อง โดยการลบข้อมูลที่ไม่ถูกต้องออกจากปัจจัยการพยากรณ์

3) การแปลงรูปแบบข้อมูล (data Transformation) เป็นการแปลงข้อมูลที่เลือกมาให้อยู่ในรูปแบบที่เหมาะสมสำหรับการนำไปใช้วิเคราะห์ตามอัลกอริทึม และแบบจำลองที่ใช้ในการทำเหมืองข้อมูลต่อไป การจัดข้อมูลให้เป็นรูปแบบมาตรฐานเดียวกัน (normalization)

โดยทฤษฎีดังกล่าวถูกนำมาประยุกต์ใช้วิจัยในด้านการพยากรณ์ข้อมูลในหลากหลายด้าน เช่น ทางด้านเศรษฐศาสตร์ (Chen *et al*.,2000; Sadeghi, 2015) พยากรณ์ราคามันสำปะหลังโดยใช้แบบจำลองอารีมา (Arima) และวิธีการปรับให้เรียบแบบเอกซ์โพเนนเชียลอย่างง่าย (simple exponential smoothing) ซึ่งในมิติทางด้านการเกษตรการ

พยากรณ์ข้อมูลในอนาคตจะช่วยให้เกษตรกรสามารถวางแผนเพื่อเตรียมปัจจัยการผลิตได้อย่างเหมาะสม เช่น การพยากรณ์ด้านราคาผลผลิตล่วงหน้า (Pattranurakyothin and Kumnungkit, 2012) การพยากรณ์น้ำในเขื่อน (วีรศักดิ์, 2560) การพยากรณ์ผลผลิตมันสำปะหลังด้วยเทคนิคการทำเหมืองข้อมูล (ปรีชา และคณะ, 2559) การพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์รายเดือนของประเทศไทย (ธรณินทร์, 2561) ซึ่งสอดคล้องกับงานวิจัยที่ต้องการพัฒนาตัวแบบการพยากรณ์ปริมาณผลผลิตไยจากข้อมูลรูปแบบการผลิตในระดับแปลงโดยเปรียบเทียบเทคนิค 3 กลุ่ม Technique Tree, Technique rules และ Technique Function

วิธีวิจัย

การศึกษาปัจจัยด้านการผลิตที่ส่งผลกระทบต่อปริมาณผลผลิตลำไยในแต่ละฤดูกาลผลิตและพัฒนาตัวแบบการพยากรณ์ผลผลิตลำไยรายแปลงเริ่มจากการรวบรวมข้อมูลที่เกี่ยวข้องกับการผลิตประกอบด้วย ข้อมูลรูปแบบการผลิต เช่น ความถี่ในการใส่ปุ๋ย สูตรปุ๋ยที่ใช้ การใช้สารกระตุ้นการออกดอก เป็นต้น และข้อมูลปริมาณผลผลิต ใน 1 รอบการผลิตปีเดียวกัน จากนั้นนำข้อมูลเข้าสู่กระบวนการเหมืองข้อมูล คือ การเลือกข้อมูล (data selection) การทำความสะอาดข้อมูล (data cleaning) และการแปลงรูปแบบข้อมูล (data Transformation) เพื่อจัดเก็บข้อมูลในอยู่ในรูปแบบฐานข้อมูล(database) เมื่อได้ข้อมูลในรูปแบบของฐานข้อมูลแล้วต่อไปเป็นขั้นตอนการในการสร้างตัวแบบพยากรณ์ โดยทำการ Training ข้อมูล ด้วยโปรแกรม Weka ตามหลักการเรียนรู้ของเครื่องอาศัย (machine learning) โดยเลือกอัลกอริทึม Technique Tree Technique Rules และ Technique Function เป็นตัวแบบในการเรียนรู้ซึ่งใช้ข้อมูลร้อยละ 80 ของข้อมูลทั้งหมด จากนั้นนำตัวแบบพยากรณ์ที่สร้างได้ทำการทดสอบประสิทธิภาพ โดยประเมินจากค่าความถูกต้อง (accuracy) ความความแม่นยำ (precision) เป็นต้น (Figure 1)

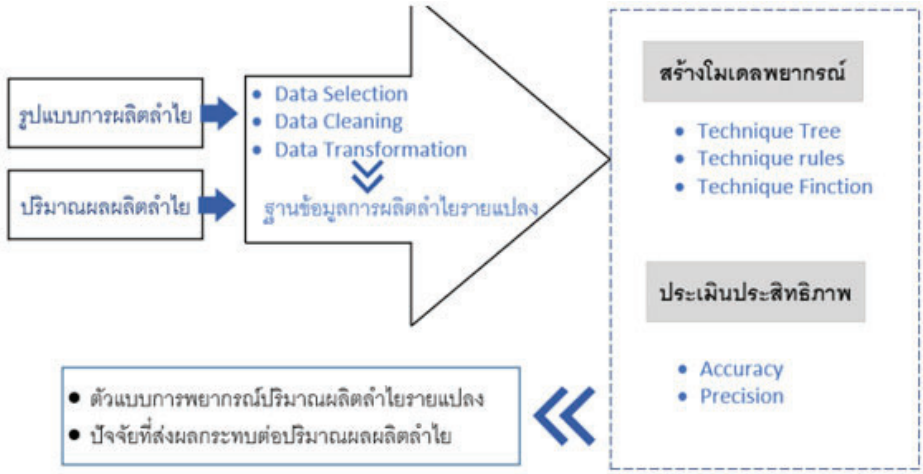

Figure 1 Research Framework.

โดยมีรายละเอียดวิธีวิจัย ทั้ง 3 ขั้นตอน ดังนี้

1) รวบรวมข้อมูลการผลิตลำไยโดยข้อมูลประชากรในการศึกษาวิจัย กำหนดเป็นเกษตรกรผู้ผลิตลำไยสายพันธุ์อีดอ ที่มีการปลูกมาแล้วไม่น้อยกว่า 5 ปี กำหนดขอบเขตประชากรโดยอาศัยขอบเขตการปกครอง โดยเลือกตามขอบเขตจังหวัดประกอบด้วย พะเยา เชียงใหม่และลำพูน จำนวน 215 แปลง กำหนดกลุ่มตัวอย่างโดยการเลือกกลุ่มตัวอย่างแบบไม่เป็นไปตามโอกาสทางสถิติ (non probability sampling) เครื่องมือที่ใช้ในการวิจัยประกอบด้วยแบบสอบถามเกี่ยวกับข้อมูลการผลิตลำไยของเกษตรกร โดยแบ่งข้อมูลเป็น 3 ส่วน คือ

ส่วนที่ 1 สภาพแวดล้อมของแปลงผลิตประกอบด้วย สภาพพื้นที่ (ที่ลุ่ม ที่ดอน) แหล่งน้ำ (คลอง น้ำฝน น้ำบาดาล อ่างเก็บน้ำ) ระบบน้ำ (ปริงเกอร์ น้ำฝน ร่องน้ำ สายยาง) ขนาดแปลง (จำนวนไร่)

อายุต้นลำไย (ปี) ความกว้างทรงพุ่ม (4x4 5x5 6x6 7x7 เมตร) ความสูงของต้นลำไย ระยะปลูก (เมตร)

ส่วนที่ 2 วิธีการผลิต ประกอบด้วย การใช้สารกระตุ้นการออกดอก (โพแทสเซียมคลอเรต โซเดียมคอลเรต) วิธีใส่ (ราดทางดิน พ่นทางใบ) เดือนที่ให้สารกระตุ้นการออกดอก

ส่วนที่ 3 เป็นข้อมูลปริมาณผลผลิต แบ่งเป็นระดับปริมาณผลผลิต เป็น 3 ระดับ

ระดับต่ำ หมายถึง ปริมาณผลผลิต ต่ำกว่า 500 กิโลกรัมต่อไร่

ระดับปานกลาง หมายถึง ประมาณผลผลิต ระหว่าง 501 – 1,000 กิโลกรัมต่อไร่

ระดับมาก หมายถึง ปริมาณผลผลิต มากกว่า 1,000 กิโลกรัมต่อไร่

2) ประมวลผลข้อมูลด้วยเทคนิคเหมืองข้อมูลประกอบด้วยการคัดเลือกข้อมูล (data selection) พิจารณาข้อมูลจากแบบสอบถามพบว่ามีข้อมูลบางส่วนที่ไม่เป็นประโยชน์และไม่มีความสัมพันธ์คำตอบที่ต้องการในประเด็นการพยากรณ์ปริมาณผลผลิต เช่น ชื่อเกษตรกร ราคาขาย และทำความสะอาดข้อมูล (data cleaning) พิจารณาจากความสมบูรณ์ของข้อมูลเป็นหลัก เนื่องจากความสมบูรณ์ของข้อมูลจะส่งผลโดยตรงต่อความแม่นยำในการพัฒนาตัวแบบ การแปลงข้อมูล (data transformation) เป็นการจัดรูปแบบข้อมูลดิบให้อยู่ในรูปแบบที่เหมาะสมสำหรับการทำเหมืองข้อมูล เช่น สภาพพื้นที่(area) ปรับรูปแบบข้อมูล เป็น 2 คำตอบ คือ 0 หมายถึง ที่ลุ่ม 1 หมายถึง ที่ดอน อายุต้นเฉลี่ย(age) ปรับรูปแบบข้อมูลจากตัวเลขอายุต้นลำไย เป็นช่วงอายุ โดยกำหนด เป็น 5 ช่วง 1 หมายถึง อายุ ต่ำกว่า 5 ปี 2 หมายถึง อายุมากกว่า 5 ปี แต่ไม่เกิน 10 ปี 3 หมายถึง อายุมากกว่า 10 ปี แต่ไม่เกิน 15 ปี 4 หมายถึง อายุมากกว่า 15 ปี แต่ไม่เกิน 20 ปี และปริมาณผลผลิต(product) ปรับรูปแบบเป็น 3 ช่วง คือ low (ผลผลิตต่ำกว่า 500 กิโลกรัมต่อไร่) moderate(ผลผลิต 500 - 1,000 กิโลกรัมต่อไร่) high (ผลผลิตมากกว่า 1,000 กิโลกรัมต่อไร่) ดัง (Figure 2)

area	age	distance	height	water	water system	chemical	month	method	product
1	5	6x6	5	3	4	1	Dec.	3	Low
1	2	6x6	1	1	4	2	Dec.	3	medium
1	3	6x6	1	2	1	2	Dec.	1	medium
1	3	6x6	1	2	4	3	Dec.	3	High
1	3	6x6	4	2	1	2	Dec.	3	High
1	3	5x5	1	3	4	2	Dec.	3	Low
1	4	6x6	2	3	4	2	Dec.	3	Low
1	3	8x8	1	3	4	2	Dec.	3	medium
1	4	6x6	4	2	1	2	Dec.	3	Low
1	4	6x6	4	2	4	2	Dec.	3	High
1	3	6x6	3	1	4	1	Dec.	3	medium
1	5	8x8	5	2	1	3	Dec.	3	Low
1	3	6x6	5	3	4	1	Dec.	3	medium
1	5	6x6	5	1	2	2	Dec.	3	medium
1	1	6x6	4	3	4	2	Dec.	3	High
1	1	6x6	5	3	4	2	Dec.	3	medium
1	3	6x6	5	3	4	2	Dec.	3	Low
1	4	6x6	5	3	4	2	Dec.	3	Low
1	4	4x4	5	2	1	2	Dec.	1	medium
1	2	6x6	3	3	4	1	Dec.	2	medium

Figure 2 Example data set.

3) วิเคราะห์ผลด้วยวิธีการจำแนกประเภทข้อมูล (classification) ด้วยโปรแกรม WEKA ด้วยเทคนิคต้นไม้ตัดสินใจ (decision tree) และการหากฎความสัมพันธ์ (association rules)

ผลการศึกษา

1) สภาพการผลิตลำไยของกลุ่มตัวอย่าง

รวบรวมข้อมูลการผลิตลำไยจากเกษตรกรจำนวน 215 แปลง สามารถวิเคราะห์ข้อมูลเบื้องต้นได้ใน 3 ประเด็น คือ ลักษณะทางกายภาพ พบว่า สภาพพื้นที่เป็นที่ลุ่ม ร้อยละ 54.42 และ ที่ดอน ร้อยละ 45.58 อายุต้นลำไยจะมีอายุระหว่าง 15 – 20 ปี ร้อยละ 28.84 อายุระหว่าง 5 – 10 ปี ร้อยละ 25.12 อายุระหว่าง 10 – 15 ปี ร้อยละ 24.19 อายุน้อยกว่า 5 ปี ร้อยละ 6.51 และอายุมากกว่า 20 ปี ร้อยละ 15.35 ตามลำดับ

ลักษณะระยะปลูกจะใช้ระยะห่างขนาด 8x8 เมตร ร้อยละ 39.07 ขนาด 6x6 ร้อยละ 35.35 และขนาด 4x4 ร้อยละ 12.09 ตามลำดับ ความสูงของต้นลำไยสูงระหว่าง 5.01 – 6.00 เมตร ร้อยละ 29.79 ความสูงระหว่าง 6.01 – 7.00 เมตร ร้อยละ 22.79 และความสูงน้อยกว่า 3 เมตร ร้อยละ 19.07 ตามลำดับ แหล่งน้ำใช้น้ำจากคลอง ร้อยละ 39.53 น้ำบาดาล ร้อยละ 26.51 อ่างเก็บน้ำ ร้อยละ 18.60 และ น้ำฝน ร้อยละ 13.35 ส่วนระบบน้ำใช้สายยางรดน้ำ ร้อยละ 40.93 ระบบสปริงเกอร์ ร้อยละ 41.86 ปล่อยตามธรรมชาติ ร้อยละ 13.02 และระบบปล่อยน้ำท่วมแปลง ร้อยละ 4.19 ดังแสดงใน (Table 1)

Table 1 The physical characteristics of longan cultivation plot.

Item	Number of plots	Percentage
Characteristics of the area		
1: Upland	117	54.42
2: Lowland	98	45.58
Total	**215**	**100.00**
Summary of data from early age		
1: Less Than 5 Year	14	6.51
2: 5 – 10 Year	54	25.12
3: 10 – 15 Year	52	24.19
4: 15 – 20 Year	62	28.84
5: More Than 20 Year	33	15.35
Total	**215**	**100.00**
Summary of data from distance		
10x10	13	6.05
4x4	26	12.09
5x5	9	4.19
6x6	76	35.35
7x7	7	3.26
8x8	84	39.07
Total	**215**	**100.00**
Summary of data from tree height		
1: Less Than 3 Meter	41	19.07
2 = 3.01 – 4.00 Meter	56	13.02
3 = 4.01 – 5.00 Meter	51	7.91
4 = 5.01 – 6.00 Meter	252	29.30
5 = 6.01 – 7.00 Meter	245	22.79
6 = More Than 7.01 Meter	102	7.91
Total	**215**	**100.00**

Table 1 (continued)

Item	Number of plots	Percentage
Summary of data from Water Source		
1: Canal	85	39.53%
2: Rain	33	15.35%
3: Ground water	57	26.51%
4: Reservoir	40	18.60%
Total	**215**	**100.00**
Summary of data from Water System		
1: Nature	28	13.02
2: Releasing flood waters	9	4.19
3: Sprinkler	90	41.86
4: Watering hose	88	40.93
Total	**215**	**100.00**

เทคนิคการใช้สารกระตุ้นออกดอก พบว่า ใช้สารโซเดียมคลอเรต ($NaClO_3$) ร้อยละ 51.16 ใช้สารโพแทสเซียมคลอเรต ($KClO_3$) ร้อยละ 44.19 และใช้ทั้งโซเดียมคลอเรต ($NaClO_3$) ผสมกับโพแทสเซียมคลอเรต ($KClO_3$) ร้อยละ 4.65 ตามลำดับ ช่วงเดือนที่ทำการให้สารกระตุ้นการออกดอก พบว่า ส่วนใหญ่ให้สารกระตุ้นในเดือนธันวาคม ร้อยละ 76.74 เดือน พฤศจิกายน ร้อยละ 12.56 และเดือนกันยายน ร้อยละ 5.12 ตามลำดับ ส่วนวิธีการให้สารกระตุ้นการออกดอก พบว่า ใช้การฉีดพ่นทางใบพร้อมกับราดทางดิน ร้อยละ 59.07 ฉีดพ่นทางใบอย่างเดียว ร้อยละ 39.07 และราดทางดินอย่างเดียว ร้อยละ 8.84 ตามลำดับ ดังข้อมูล (Table 2)

Table 2 Summary of flower-inducing agent application techniques.

Item	Number of plots	Percentage
Flowering stimulants		
1: Sodium chlorate	110	51.16
2: potassium chlorate	95	44.19
3: Both together	10	4.65
Total	**215**	**100.00**
Month of application of flowering stimulants		
September	11	5.12
October	2	0.93
December	165	76.74
November	27	12.56
January	10	4.65
Total	**215**	**100.00**

Table 2 (continued)

Item	Number of plots	Percentage
Methods of giving flowering stimulants		
1: Leaf	69	32.09
2: Soil	19	8.84
3: Leaf and Soil	127	59.07
Total	**215**	**100.00**

ปริมาณผลผลิต พบว่าเกือบครึ่งได้ผลผลิตในระดับ ปานกลาง (ผลผลิต มากกว่า 1,000 กิโลกรัมต่อไร่) ร้อยละ 47.91 รองลงมาได้ผลผลิตในระดับมาก (ผลผลิต มากกว่า 1,000 กิโลกรัมต่อไร่) ร้อยละ 32.56 และ ได้ผลผลิตในระดับต่ำ (ผลผลิต ต่ำกว่า 500 กิโลกรัมต่อไร่) ร้อยละ 19.53 ตามลำดับ ดังข้อมูล (Table 3)

Table 3 Summary of data from Productivity.

Item	Number of plots	Percentage
low	42	19.53
moderate	103	47.91
high	70	32.56
Total	**215**	**100.00**

2) ผลการสร้างแบบจำลองด้วยเทคนิค Technique Tree

จากการประมวลผลการสร้างแบบจำลองด้วยอัลกอริทึม J48 Hoeffding Tree และ Random พบว่าแบบจำลองที่สร้างด้วยอัลกอริทึม J48 ได้ค่าความถูกต้อง (correctly) และค่าความแม่นยำเฉลี่ยสูงที่สุด ร้อยละ 94.77 และ 94.80 รองลงมาเป็นอัลกอริทึม RandomTree ได้ค่าความถูกต้อง (correctly) และค่าความแม่นยำเฉลี่ย ร้อยละ 70.58 และ 29.41 สุดท้ายเป็นอัลกอริทึม HoeffdingTree ได้ค่าความถูกต้อง (correctly) และค่าความแม่นยำเฉลี่ย ร้อยละ 64.53 และ 35.36 ผู้วิจัยเลือกใช้อัลกอริทึม weka.classifiers.trees.J48 เนื่องจากได้ค่าความถูกต้องและความแม่นยำเฉลี่ยสูงที่สุด และพบปัจจัยที่ส่งผลต่อการพยากรณ์ปริมาณผลผลิต คือ ชนิดของสารกระตุ้นการออกดอก เดือนที่ใส่สารกระตุ้น และความสูงของต้นลำไย ดังข้อมูล (Table 4)

Table 4 Comparing the efficiency of data classification

Algorithm	Correctly Classified Instances (%)	Incorrectly Classified Instances (%)	precision
weka.classifiers.trees.J48	94.77	5.23	94.80
weka.classifiers.trees.HoeffdingTree	64.53	35.46	64.50
weka.classifiers.trees.RandomTree	70.58	29.41	71.4

3) ผลการสร้างโมเดลด้วยเทคนิค Technique rules

การหากฎความสัมพันธ์ (Association Rules) ด้วยวิธีใช้วิธี apriori ชุดข้อมูลประกอบด้วยปัจจัยดังนี้ สภาพพื้นที่ แหล่งน้ำ ระบบน้ำ ขนาดของแปลงผลิต อายุต้นลำไย ความกว้างทรงพุ่ม ความสูงของต้นลำไย ระยะปลูกของต้นลำไย การใช้สารกระตุ้น

การออกดอก วิธีการใช้ เดือนที่ให้สารกระตุ้นการออกดอก และปริมาณผลผลิตต่อไร่ ข้อมูลจำนวน 215 แปลง ด้วยการระบุคำตอบ กำหนดค่า minimum support >= 1 และ ค่า confidence >= 0.60 (Li et al., 2017) หาความสัมพันธ์ทั่วไปสามารถแจกแจงกฎได้ดังข้อมูล (Table 5)

Table 5 The Association Rule of the Apriori Algorithm weka.associations.Apriori.

	Rules	confidence	support
1	chemical=1 month=Dec. 85 ==> product=medium 53	0.62	53/215 = 0.24
2	chemical=1 110 ==> product=medium 68	0.62	68/215 = 0.31
3	water source=1 85 ==> product=medium 46	0.54	46/215 = 0.21
4	distance=8x8 84 ==> product=medium 45	0.54	45/215 = 0.20
5	month=Dec. 165 ==> product=medium 86	0.52	86/215 = 0.40

สามารถอธิบายกฎความสัมพันธ์ได้ดังนี้

กฎที่ 1 แปลงที่มีการใช้สารโซเดียมคลอเรต ในเดือนธันวาคม จำนวน 85 แปลง จะได้ผลผลิตในระดับปานกลาง 53 แปลง ที่ระดับความเชื่อมั่นร้อยละ 62

กฎที่ 2 แปลงที่มีการใช้สารโซเดียมคลอเรต จำนวน 110 แปลง จะได้ผลผลิตในระดับปานกลาง 68 แปลง ที่ระดับความเชื่อมั่น ร้อยละ 62

กฎที่ 3 , 4 และ 5 เป็นกฎที่มีระดับความเชื่อมั่นต่ำกว่า ร้อยละ 60 จึงไม่มีความน่าเชื่อถือ

อภิปรายผลการวิจัย

จากผลการวิจัยที่นำเสนอประสิทธิภาพของเทคนิคการจำแนกข้อมูลพบว่า Technique Tree อัลกอริทึม weka.classifiers.trees.J48 พบปัจจัยที่ส่งผลกระทบต่อปริมาณผลิตลำไยจากโมเดลการพยากรณ์ คือ เดือนที่ใส่สารกระตุ้นการออกดอก และประเภทของสารกระตุ้นการออกดอกซึ่งเมื่อใช้สารโซเดียมคลอเรตจะมีค่าพยากรณ์ปริมาณผลผลิตในระดับปานกลางขึ้นไปพิจารณาจากกฎความสัมพันธ์ที่มีค่าความเชื่อมั่นในลำดับที่ 1 และ 2 สอดคล้องกับกฎความสัมพันธ์ของข้อมูลในกฎที่ 1 พบว่าหากใส่สารโซเดียมคลอเรตในเดือนธันวาคมปริมาณผลผลิตจะอยู่ในระดับปานกลางขึ้นไปจำนวน 53 แปลงและมีความเชื่อมั่นที่ ร้อยละ 62 และกฎที่ 2 หากมีการใช้สารโซเดียมคลอเรตโดยไม่พิจารณาเดือนที่ใส่พบว่าได้ผลผลิตในระดับปานกลางเช่นกันจำนวน 68 แปลง และมีความเชื่อมั่นที่ ร้อยละ 62 ซึ่งสอดคล้องกับการศึกษาผลของสารกลุ่มคลอเรตต่อการชักนำการเปิดปิดของปากใบและคุณภาพของผลลำไยซึ่งพบว่าค่าการชักนำการเปิดปากใบของต้นลำไยที่ได้รับโพแทสเซียมคลอเรต ($KClO_3$) สูงกว่าต้นที่ไม่ได้ราดโพแทสเซียมคลอเรต ($KClO_3$) (วินัย, 2564) และการศึกษาผลของสารคลอเรตต่อการออกดอก ติดผล ของลำไยพันธุ์อีดอพบว่าต้นลำไยที่ให้สารโพแทสเซียมคลอเรตหรือโซเดียมคลอเรตในช่วงเดือนธันวาคม ใช้ระยะเวลาเฉลี่ยในการแทงช่อดอกนานที่สุดคือ 45 วันมากกว่าเดือนอื่น ๆ และอัตราการออกดอกมากที่สุด (บุญชาติ, 2551) สอดคล้องกับการศึกษาของ (Manochai *et al*., 2005) พบว่าการให้สารโพแทสเซียมคลอเรตในเดือนสิงหาคมซึ่งตรงกับฤดูฝนต้นลำไยออกดอกน้อยที่สุด

สรุป

จากผลการวิจัยสรุปได้ว่าตัวแบบการพยากรณ์ผลผลิตลำไยรายแปลงในพื้นที่จังหวัดเชียงใหม่ ลำพูนและพะเยา แบบจำลองที่สร้างด้วยอัลกอริทึม J48 Hoeffding Tree ได้ค่าความถูกต้อง (correctly) และค่าความแม่นยำเฉลี่ยสูงที่สุดร้อยละ 94.77 และ 94.80 และหากฎความสัมพันธ์ (Association Rules) ด้วยวิธีใช้วิธี apriori พบว่าปัจจัยด้านการผลิตที่ส่งผลกระทบต่อปริมาณผลผลิตลำไย คือ ประเภทของสารกระตุ้น การออกดอกและเดือนที่ทำการใส่สารกระตุ้นการออกดอก

ข้อเสนอแนะ

จากตัวแบบดังกล่าวสามารถนำตัวแบบไปพัฒนาระบบพยากรณ์ปริมาณผลผลิตล่วงหน้าระดับแปลงโดยเกษตรกรให้ข้อมูลตามตัวแปรที่กำหนด เช่น ลักษณะพื้นที่ ความสูงของต้นลำไย สารกระตุ้นการออกดอกที่ใช้ เดือนที่ใส่สารกระตุ้นการออกดอก เป็นต้น ระบบจะพยากรณ์ปริมาณระดับผลผลิต อย่างไรก็ตามตัวแบบที่ใช้ในการพยากรณ์อ้างอิงข้อมูลจากจังหวัดเชียงใหม่ ลำพูน และพะเยา ดังนั้นแปลงผลิตนอกเหนือจากจังหวัดดังกล่าวอาจส่งผลต่อความแม่นยำเนื่องจากมีปัจจัยแวดล้อมที่แตกต่างกันและข้อมูลการผลิตลำไยของกลุ่มตัวอย่างใช้เป็นเซตข้อมูลสำหรับพัฒนาโมเดลอาจต้องพิจารณาปัจจัยอื่น ๆ ที่ควบคุมไม่ได้ที่คาดว่ามีผลกระทบต่อปริมาณผลิต เช่น ปริมาณน้ำฝน หรือพิจารณาจากปัจจัยสนับสนุน เช่น สูตรปุ๋ย ระดับความสมบูรณ์ของต้นลำไย เป็นต้น และข้อมูลกลุ่มตัวอย่างของแต่ละจังหวัดมีความแตกต่างกันมากโดยเฉพาะปริมาณผลผลิตต่อไร่เฉลี่ยของแต่ละจังหวัดมีความแตกต่างกัน หากนำข้อมูลมารวมกันอาจส่งผลต่อประสิทธิภาพการพยากรณ์หรือสร้างโมเดลพยากรณ์เป็นรายจังหวัด

เอกสารอ้างอิง

บุญชาติ คติวัฒน์. 2551. ผลของสารโพแทสเซียมคลอเรตต่อการออกดอกติดผลของลำไยพันธุ์อีดอในรอบปี. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต. มหาวิทยาลัยแม่โจ้, เชียงใหม่. 90 หน้า.

วีรศักดิ์ ฟองเงิน. 2560. การพยากรณ์ปริมาณน้ำในเขื่อนโดยใช้เทคนิคเหมืองข้อมูล. วารสารวิชาการการจัดการเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม 4(1): 27–33.

ปรีชา ลิ้มตระกูล วิภา เจริญภัณฑารักษ์ และวิทยา พรพัชรพงศ์. 2559. การพัฒนาตัวแบบการพยากรณ์ผลผลิตมันสำปะหลังด้วยเทคนิคการทำเหมืองข้อมูล. วารสารวิชาการเวอริเดียน อี เจอร์นัล สาขาวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยศิลปากร 3(3): 15–36.

สำนักงานเศรษฐกิจการเกษตร. 2563. ข้อมูลการผลิตสินค้าเกษตร (ระบบออนไลน์). แหล่งข้อมูล: www.oae.go.th (30 มีนาคม 2567).

วินัย วิริยะอลงกรณ์. 2564. ผลของสารโพแทสเซียมคลอเรตต่อความเข้มข้นของคาร์บอนไดออกไซด์ในเขตรากพืชการชักนำการเปิดปิดของปากใบและคุณภาพของผลลำไย. วารสารวิจัยและส่งเสริมวิชาการเกษตร 38(1): 12–27.

ธรณินทร์ สัจวิริยทรัพย์. 2561. ตัวแบบโครงข่ายประสาทเทียมสำหรับการพยากรณ์ราคาข้าวโพดเลี้ยงสัตว์รายเดือนของประเทศไทย. Srinakharinwirot Science Journal 34(1): 91–108.

Chen, F., Drezner, Z., Ryan, J. K., and Simchi-Levi, D. 2000. Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times and Information. Management Science 46(3): 436–443.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. 1996. From Data Mining to Knowledge Discovery in Databases. AI Magazine 17(3): 37–37.

Li, L., Shrestha, S., and Hu, G. 2017. Analysis of road traffic fatal accidents using data mining techniques. 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA): 363–370.

Manochai, P., P. Sruamsiri., W. Wiriya-alongkorn., D. Naphrom., M. Hegele and F. Bangerth. 2005. Year around off season flower induction in longan (*Dimocarpus longan*, Lour.) trees. by $KClO_3$ applications: potentials and problems. Scientia Horticulturae 104: 379-390.

Pattranurakyothin, T., & Kumnungkit, K. 2012. Forecasting Model for Para Rubber's Export Sales. Current Applied Science and Technology 12(2): 198–202.

Sadeghi, A. 2015. Providing a measure for bullwhip effect in a two-product supply chain with exponential smoothing forecasts. International Journal of Production Economics 169: 44–54.