



การทำนายความหวานผลเมล่อนแบบรวดเร็วโดยใช้เทคนิคการประมวลผลภาพร่วมกับแบบจำลองอัลกอริทึม

Rapid Prediction of Melon Sweetness Using Image Processing Techniques and Algorithmic Models

นพพร รัตนช่วง*

Nopporn Rattanachoung*

ภาควิชาฟิสิกส์ คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตกำแพงแสน จ.นครปฐม 73140
Department of Physics, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, THAILAND

*Corresponding author e-mail: faasnor@ku.ac.th

ARTICLE INFO

Article history:

Received: March 2, 2023

Revised: April 19, 2023

Accepted: April 24, 2023

Available online: April 28, 2023

DOI: 10.14456/jarst.2023.11

Keywords: Image processing, melon fruit, Python language, Random Forest (RF), eXtreme Gradient Boosting (XGB)

ABSTRACT

Rapid prediction of melon sweetness using image processing techniques and algorithmic models is studied. The sweetness of melon is usually greater than 12.5 Brix, which can be measured with a digital sweet meter where most consumers do not know such a problem. This work proposed an image processing technology combined with an algorithmic model to determine the sweetness of the melon. The melon sweetness measurement method collects 100 samples of full-leaf melon photographs from 4 sides, top, bottom, left and right, consisting of weight, circumference, stalk size, and small circle axis size. The size of the core to the edge of the large circle net is height, light intensity, and relative humidity, color temperature and intensity using image processing techniques. All of this data is learned using Random Forest (RF) and eXtreme Gradient Boosting (XGB) techniques, one of the machine learning algorithms used to predict melon sweetness. The test results showed that the R-squared value of model learning in the XGB algorithm was 99%, while that of the RF algorithm was 95%. It could be concluded that the XGB algorithm was effective in model learning-better than the RF algorithm, with the weight of the melon being the most prominent. The melon color value was the least significant in interpreting the melon sweetness value. For the actual test, a

model of the XGB algorithm was used to predict the sweetness value of 50 melons. The prediction accuracy is up to 95%, utilized for checking melon results.

บทคัดย่อ

การทำนายความหวานผลเมล่อนแบบรวดเร็ว โดยใช้เทคนิคการประมวลผลภาพร่วมกับแบบจำลองอัลกอริทึม ค่าความหวานของเมล่อนปกติจะมากกว่า 12.5 Brix สามารถตรวจวัดได้จากเครื่องวัดความหวานแบบดิจิตอล ซึ่งผู้บริโภคส่วนใหญ่ไม่ทราบปัญหาดังกล่าวผู้วิจัยได้เสนอเทคนิคการประมวลผลภาพร่วมกับแบบจำลองอัลกอริทึมเพื่อหาค่าความหวานของเมล่อน วิธีการตรวจวัดความหวานของเมล่อนจะดำเนินการเก็บข้อมูลภาพถ่ายผลเมล่อนเต็มใบ 4 ด้าน คือ ด้านบน ด้านล่าง ด้านซ้ายและด้านขวา ทั้งหมด 100 ตัวอย่าง ประกอบด้วย ค่าน้ำหนัก เส้นรอบวง ขนาดก้าน ขนาดแกนกลาง วงกลมเล็ก ขนาดแกนกลางถึงขอบตาข่ายวงกลมใหญ่ ความสูง ความเข้มแสง ความชื้นสัมพัทธ์ อุณหภูมิและความชื้นสี โดยใช้เทคนิคการประมวลผลภาพ ข้อมูลทั้งหมดนี้จะถูกนำมาเรียนรู้ด้วยเทคนิค Random Forest (RF) และ eXtreme Gradient Boosting (XGB) เป็นอัลกอริทึมอันหนึ่งในกระบวนการเรียนรู้เครื่องจักรที่ถูกเลือกใช้ในการทำนายค่าความหวานของเมล่อน ผลการทดสอบพบว่าค่า R-squared ของการเรียนรู้โมเดลในอัลกอริทึม XGB คือ 99% ในขณะที่อัลกอริทึม RF คือ 95% จึงสรุปได้ว่าอัลกอริทึม XGB มีประสิทธิภาพในการเรียนรู้โมเดลได้ดีกว่าอัลกอริทึม RF โดยค่าน้ำหนักของเมล่อนมีข้อเด่นชัดสำคัญที่สุด และค่าสีของเมล่อนมีข้อเด่นชัดสำคัญน้อยที่สุดในการแปลผลกับค่าความหวานของเมล่อน สำหรับการทดสอบจริงได้นำแบบจำลองของอัลกอริทึม XGB มาทำนายค่าความหวานของเมล่อนจำนวน 50 ตัวอย่าง พบว่ามีค่าความแม่นยำในการทำนายมากถึง 95% สามารถนำไปใช้ประโยชน์ในการตรวจสอบผลเมล่อน

คำสำคัญ: การประมวลผลภาพ ผลเมล่อน ภาษาไพทอน แบบจำลองอัลกอริทึม RF แบบจำลองอัลกอริทึม XGB

บทนำ

เมล่อนเป็นพืชในตระกูลแตงที่นิยมปลูกเพื่อการค้าชนิดหนึ่งที่มีราคาต่อผลสูง เนื่องจากเป็นแตงที่มีรสหอมหวาน มีคุณค่าทางโภชนาการ ประกอบด้วยวิตามินเอและวิตามินซีสูง มีถิ่นกำเนิดในต่างประเทศในแถบทวีปแอฟริกาใช้รับประทานผลสุกมีกลิ่นหอมรสหวาน เจริญได้ดีในสภาพร้อนแห้ง แสงแดดจัด ปัจจุบันมีการผลิตออกมามากหลายสายพันธุ์ที่เจริญเติบโตได้ดีในสภาพภูมิประเทศของประเทศไทย [1] การผลิตเมล่อนที่มีคุณภาพสูงจำเป็นต้องมีการดูแลเอาใจใส่ในการเพาะปลูกและการจัดการแปลงผลผลิต นอกจากนี้เมล่อนอุดมไปด้วยวิตามินอี ที่มีชื่อเสียงมากในวงการบำรุงผิวพรรณ มันให้ความชุ่มชื้น ลดการอักเสบของผิว รวมทั้งช่วยผลิตเซลล์ผิว ทำให้ผิวเนียน เรียบขาวกระจ่างใส นอกจากนี้ยังมีวิตามินซี ที่ถือว่าเป็นสารต้านอนุมูลอิสระโดยธรรมชาติที่ทรงประสิทธิภาพ ช่วยลดการเกิดความเสื่อมของเซลล์ต่าง ๆ ในร่างกายและมีประสิทธิภาพสูงในการเสริมสร้างคอลลาเจน [1] โดยเฉพาะคอลลาเจนไตรเปปไทด์และที่สำคัญผู้บริโภคยังไม่ทราบค่าความหวานของผลเมล่อนที่เลือกซื้อก่อนหน้านี้ไปบริโภค

จากคุณประโยชน์ข้างต้นผู้วิจัยจึงต้องการทำนายความหวานของผลเมล่อน โดยเลือกพันธุ์เมล่อนที่ต้องการพร้อมทั้งนำผลเมล่อนมาศึกษาโดยใช้วิธีการตรวจสอบและทำนายความหวานผลเมล่อนแบบรวดเร็ว โดยใช้เทคนิคการประมวลผลภาพ [2, 3] ร่วมกับแบบจำลองอัลกอริทึม ศึกษาลักษณะทางกายภาพภายนอกของเมล่อน เก็บข้อมูลซึ่งประกอบด้วย ค่าน้ำหนัก (x1) เส้นรอบวง (x2) ขนาดก้าน (x3) ขนาดแกนกลางวงกลมเล็ก (x4) ขนาดแกนกลางถึงขอบตาข่ายวงกลมใหญ่ (x5) ความสูง (x6) ความเข้มแสง (x7) ความชื้นสัมพัทธ์ (x8) อุณหภูมิ (x9) และความชื้นสี (x10) เก็บข้อมูลภาพถ่ายผลเมล่อนเต็มใบ ด้านบน ด้านล่าง ด้านซ้ายและด้านขวา ทั้งนี้ดำเนินการวัดค่าความหวานของผลเมล่อน โดยใช้เครื่องวัดความหวาน (Brix) แบบดิจิตอล จำนวนด้านละ 4 ตำแหน่ง ทำการ

ประมวลผลภาพถ่ายผลไม้โดยนำภาพถ่ายของผลไม้มาประมวลผลภาพโดยใช้ภาษาไพทอนด้วยวิธีภาพถูกทำให้เป็นความเข้ม/จางของสีเทาที่มีพิสัย (Grayscale) แล้วตัดเอาเฉพาะส่วนตรงกลางขนาด 200x200 PIXEL เพื่อนำมาคำนวณหาค่าความเข้มของสีและคำนวณค่าเฉลี่ยความเข้มของสี จะได้ข้อมูลความเข้มสีจำนวนทั้งหมด 40,000 ค่า สามารถคำนวณค่าของสีได้ ในงานวิจัยนี้ทางผู้วิจัยนำการเรียนรู้เครื่องจักร [4] (Machine Learning) เป็นการสร้างชุดข้อมูล (Data Set) ได้เก็บตัวอย่างของชุดข้อมูลจำนวน 100 ชุด ประกอบด้วยข้อมูลขาเข้า 10 อย่าง และข้อมูลขาออก 1 อย่าง นำชุดข้อมูลมาสร้างค่าสถิติของชุดข้อมูลและนำมาหาความสัมพันธ์ระหว่างข้อมูลขาเข้าทั้งหมดและข้อมูลขาออก เพื่อนำไปสู่การเลือกอัลกอริทึม [5] ดำเนินการทดลองการทำงานของโค้ด Seaborn Heatmap โดยใช้โปรแกรม PyCharm ด้วยภาษาไพทอนพร้อมนำไปทดสอบตัวอย่าง เพื่อนำไปใช้ประโยชน์ในการตรวจสอบผลแม่นยำแบบรวดเร็วที่มีประสิทธิภาพต่อไป

เทคนิคการเรียนรู้เครื่องจักร [6] มีตัวชี้วัดที่ใช้วัดความถูกต้องของแบบจำลองหลายอย่าง สำหรับตัวชี้วัดที่ใช้ในแบบจำลอง Regression [7] ด้วยวิธีการต่อไปนี้

1. ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Squared Error, MSE) เป็นการวัดค่าความคลาดเคลื่อนของแบบจำลอง โดยค่าที่ได้ยิ่งมีค่าน้อยก็ยิ่งแสดงให้เห็นว่าแบบจำลองที่ใช้มีความแม่นยำมาก โดยค่า MSE จะมีวิธีคำนวณได้จากการนำค่าความคลาดเคลื่อนมายกกำลัง จากนั้นนำไปหาค่าเฉลี่ย ดังสมการที่ (2.1)

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad (2.1)$$

โดยที่ n คือ จำนวนข้อมูลที่ใช้

Y_t คือ ค่าจริงที่ t ใด ๆ

$\hat{Y}_t$ คือ ค่าทำนายที่ t ใด ๆ

2. ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสอง (Root Mean Square Error, RMSE) เป็นวิธีการวัดค่าความคลาดเคลื่อนแบบมาตรฐาน ที่นิยมใช้กันอย่างแพร่หลาย โดยค่าที่ได้ยิ่งมีค่าน้อยก็ยิ่งแสดงว่า

แบบจำลองที่ได้มีความแม่นยำมาก โดยค่า RMSE จะมีวิธีการคำนวณคือการนำค่า MSE ที่คำนวณได้ แล้วมาหารากที่สอง ดังสมการที่ (2.2)

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (2.2)$$

3. ค่าสัมประสิทธิ์การตัดสินใจ (Coefficient of determination, R-Squared) เป็นค่าที่ใช้พิจารณาว่าแบบจำลองที่ใช้มีความเหมาะสมมากน้อยเพียงใด โดยมีค่า 0-1 ซึ่งค่า R-Squared ยิ่งเข้าใกล้ 1 ยิ่งดี โดยทั่วไปควรมีค่า R-Squared มากกว่า 0.6 จะถือว่าแบบจำลองที่ใช้เป็นแบบจำลองที่มีประสิทธิภาพ ดังสมการที่ (2.3)

$$R^2 = 1 - \left(\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \right) \quad (2.3)$$

โดยที่ $\bar{Y}$ คือ ค่าเฉลี่ยของ Y_i

วิธีดำเนินการวิจัย

1. เครื่องมือ

1.1 เครื่องชั่งน้ำหนักแบบดิจิตอล รุ่น CS 2000 ยี่ห้อ OHAUS ดังรูปที่ 1 เก็บข้อมูลน้ำหนักของผลไม้



รูปที่ 1 เครื่องชั่งน้ำหนักแบบดิจิตอล รุ่น CS 2000 ยี่ห้อ OHAUS

1.2 เครื่องวัดความยาวอย่างละเอียด เวอร์เนียร์แคลิเปอร์ วัดขนาดเส้นผ่านศูนย์กลางและความสูง ดังรูปที่ 2 เก็บข้อมูลปริมาตรของผลไม้เพื่อนำไปคำนวณหาปริมาตร



รูปที่ 2 เครื่องวัดความยาวอย่างละเอียดเวอร์เนียร์ แคลิเปอร์ ยี่ห้อ Mitutoyo ความละเอียด 0.05 มิลลิเมตร

1.3 เครื่องวัดความหวาน (Brix) แบบดิจิตอล Handheld Sugar Meter Model RF019 Measuring Range:Brix ช่วง 0.0–32.0% ดังรูปที่ 3 ใช้เก็บข้อมูลความหวานของผลเมล่อน



รูปที่ 3 เครื่องวัดความหวาน (Brix) แบบดิจิตอลแบบ Handheld Sugar Meter Model RF019 Measuring Range : Brix ช่วง 0.0–32.0%

1.4 ชุดอุปกรณ์และเครื่องมือซึ่งประกอบด้วยฐานรองรับซึ่งเป็นแท่นโต๊ะแม่เหล็กที่มีส่วนต่าง ๆ ประกอบด้วย แกนเหล็ก ที่ยึดจับโทรศัพท์มือถือและอุปกรณ์ปรับความละเอียดของระยะที่ต้องการทดลอง ดังรูปที่ 4 ใช้เก็บข้อมูลภาพผลเมล่อน



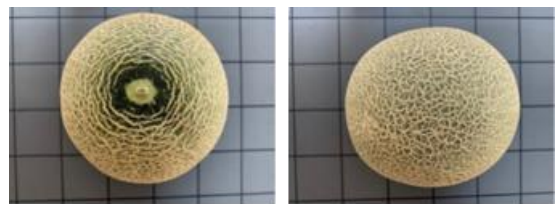
รูปที่ 4 แสดงชุดอุปกรณ์ถ่ายภาพ ทำการทดลองที่ระยะความสูง 220.85 มิลลิเมตร

1.5 เครื่องมือวัด Weatherstation ยี่ห้อ PHYWE ใช้วัดและเก็บข้อมูลอุณหภูมิ ความชื้นสัมพัทธ์ และความเข้มแสง ดังรูปที่ 5



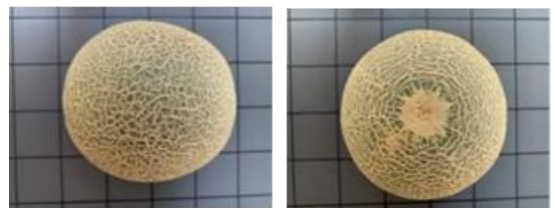
รูปที่ 5 เครื่องวัด Weatherstation ยี่ห้อ PHYWE ใช้สำหรับวัดอุณหภูมิ ความชื้นสัมพัทธ์และความเข้มแสง

1.6 ผลเมล่อนญี่ปุ่นสี่เหลี่ยมที่ใช้สำหรับการตรวจสอบความหวานของผลเมล่อนลูกที่ 1



(ก)

(ข)



(ค)

(ง)

รูปที่ 6 แสดงภาพถ่ายเต็มลูกของผลเมล่อนลูกที่ 1 (ก) ด้านบน (ข) ด้านซ้าย (ค) ด้านขวา และ (ง) ด้านล่าง



รูปที่ 7 แสดงชุดถ่ายภาพเมล่อน ที่ระยะความสูง 220.80 มิลลิเมตร

2. การดำเนินการวิจัย

2.1 วิธีการประมวลผลภาพ

เขียนโค้ดแปลงภาพเป็นตัวเลขโดยใช้รูปที่เก็บได้จากการเก็บข้อมูลมาทำเป็นภาพขาวดำ และครอบภาพตรงกลางขนาด 200x200 พิกเซลมาทำการแปลงตัวเลขเฉลี่ยด้วยภาษาไพทอน

2.2 การเรียนรู้เครื่องจักร

2.2.1 ชุดข้อมูล ผู้วิจัยได้เก็บตัวอย่างของชุดข้อมูลจำนวน 100 ชุด ประกอบด้วยค่าน้ำหนัก (x1) เส้นรอบวง (x2) ขนาดก้าน (x3) ขนาดแกนกลางวงกลมเล็ก (x4) ขนาดแกนกลางถึงขอบตาข่ายวงกลมใหญ่ (x5) ความสูง (x6) ความเข้มแสง (x7) ความชื้นสัมพัทธ์ (x8) อุณหภูมิ (x9) และความชื้นสี (x10) ข้อมูลขาเข้า 10 อย่าง และค่าความหวานของผลเมล่อน (y1) ข้อมูลขาออก 1 อย่าง

2.2.2 นำชุดข้อมูลเหล่านี้มาสร้างค่าสถิติของชุดข้อมูล นำมาหาค่าความสัมพันธ์ระหว่างข้อมูลขาเข้าทั้งหมด และข้อมูลขาออก โดยใช้แผนผังแถบสีของข้อมูล (Heatmap)

2.2.3 พารามิเตอร์ที่ดีที่สุด ในงานวิจัยนี้ทางผู้วิจัยเลือกอัลกอริทึมไว้ 2 อย่าง ได้แก่ อัลกอริทึม XGB และอัลกอริทึม RF เพื่อค้นหาค่าของพารามิเตอร์ที่ดีที่สุดของอัลกอริทึม

- อัลกอริทึม eXtreme Gradient Boosting (XGB) เป็นอัลกอริทึมที่แนวคิดการเรียนรู้ให้สำเร็จโดยการสร้าง [8] และรวมผู้เรียนหลายๆ คนเข้าด้วยกันเป็นรูปแบบการเรียนรู้แบบบูรณาการผ่านการวนซ้ำอย่างต่อเนื่อง และความแม่นยำขึ้นด้วยการเพิ่มขึ้นของเวลาการทำซ้ำ [9]

- อัลกอริทึม Random Forest (RF) เป็นอัลกอริทึมการเรียนรู้ตามแผนผังการตัดสินใจ โดยใช้เทคโนโลยีบูตสเตรปเพื่อแยกตัวอย่างแบบสุ่มจากตัวอย่างดั้งเดิม เพื่อสร้างแผนผังการตัดสินใจเดี่ยว ที่แต่ละโหมดของแผนผังการตัดสินใจจะใช้ซับซ้อน ย่อยของพีเจอร์แบบสุ่มเพื่อเลือกจุดจัดเรียง [10]

- การเปรียบเทียบอัลกอริทึม ในการเปรียบเทียบอัลกอริทึมจะตัดสินจากค่า Training R-squared ที่ได้จากผลการทำงานของโค้ดอัลกอริทึม XGB และอัลกอริทึม RF

2.3 ความสำคัญของคุณลักษณะ

ในการประเมินผลกระทบของคุณลักษณะอินพุต และความแม่นยำของการทำนายสำหรับอัลกอริทึม XGB และอัลกอริทึม RF

ผลการศึกษาและอภิปรายผล

1. การประมวลผลภาพถ่ายเมล่อน

การประมวลผลภาพถ่ายผลเมล่อน โดยนำภาพถ่ายของเมล่อนมาประมวลผลภาพโดยใช้ภาษาไพทอนด้วยวิธีภาพถูกทำให้เป็นความเข้ม/จางของสีเทาที่มีพิสัย (Grayscale) แล้วตัดเอาเฉพาะส่วนตรงกลางขนาด 200x200 PIXEL เพื่อนำมาคำนวณหาค่าความเข้มของสี และจากการประมวลผลภาพถ่ายของเมล่อนจะได้ข้อมูลความเข้มสีจำนวนทั้งหมด 40,000 ค่าและสามารถคำนวณค่าของสีได้

[[118 117 117 ... 118 119 118] [118 118 119 ... 118 119 118] [117 119 118 ... 117 118 118] ... [115 115 117 ... 115 114 115] [115 115 117 ... 116 116 118] [115 113 115 ... 114 114 118]]	[[71 69 70 ... 69 68 66] [70 69 70 ... 71 71 71] [67 68 68 ... 68 67 69] ... [68 68 68 ... 70 69 68] [69 69 68 ... 69 70 69] [69 69 68 ... 69 71 70]]
ขนาด (200, 200) 40000 ข้อมูล	ขนาด (200, 200) 40000 ข้อมูล

(ก)

(ข)

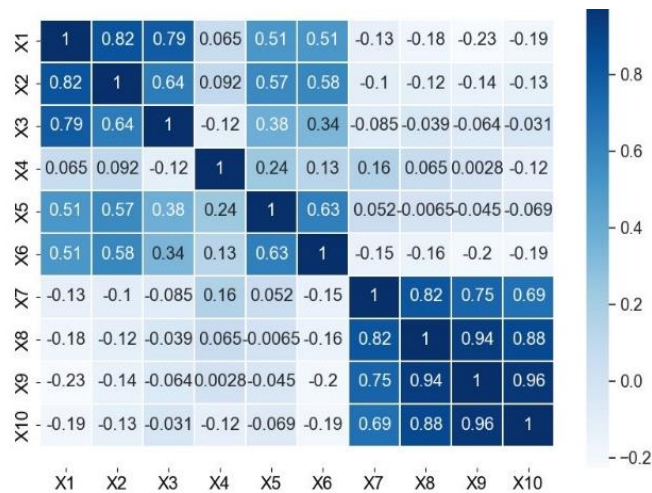
รูปที่ 8 (ก) ผลการประมวลผลตัวอย่างที่ 17 ความเข้มของสีเท่ากับ 67.040225, (ข) ผลการประมวลผลเมล่อนตัวอย่างที่ 24 ความเข้มของสีเท่ากับ 68.692450

จากรูปที่ 8 ผลการประมวลผลตัวอย่างสีของผลเมล่อนที่ 17 มีค่าความเข้มที่คำนวณได้อยู่ที่ 67.04 ในขณะที่ตัวอย่างที่ 24 มีค่าความเข้มที่คำนวณได้อยู่ที่ 68.69 เห็นได้ชัดเจนว่าตัวเลขไม่มีความแตกต่างกัน แสดงว่าความเข้มสีของเมล่อนไม่มีผลต่อความหวาน

2. การเรียนรู้เครื่องจักร

2.1 ชุดข้อมูล

ในงานวิจัยนี้ผู้วิจัยได้เก็บตัวอย่างของชุดข้อมูลจำนวน 100 ชุด ข้อมูลขาเข้า 10 อย่าง และ ข้อมูลขาออก 1 อย่าง นำชุดข้อมูลเหล่านี้มาสร้างค่าสถิติของชุดข้อมูลได้ ดังตารางที่ 1 นำมาหาค่าความสัมพันธ์ระหว่างข้อมูลขาเข้าทั้งหมด และข้อมูลขาออก โดยใช้แผนผังแถบสีของข้อมูล (Heatmap) โดยทดลองการทำงานของโค้ด Seaborn heatmap โดยใช้โปรแกรม PyCharm ด้วยภาษาไพทอน ดังรูปที่ 9



รูปที่ 9 ความสัมพันธ์ระหว่างคู่ข้อมูลนำเข้าในแต่ละคู่ในรูปแบบสีและตัวเลข

จากรูปที่ 9 จะกำหนดจุดสนใจที่ข้อมูลระหว่างข้อมูลนำเข้า และข้อมูลออกว่าข้อมูลไหนมีค่ามากจะสามารถเลือกมาใช้สร้าง Model โดยข้อมูลขาออกคือ DTX ผลที่ได้คือค่าความสัมพันธ์ระหว่างข้อมูลนำเข้าทั้งหมด และข้อมูลขาออกอยู่ในลักษณะที่สามารถนำไปใช้งานได้

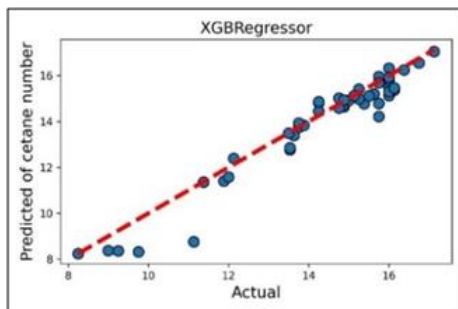
2.2 พารามิเตอร์ที่ดีที่สุด

การทดลองการทำงานของโค้ดพารามิเตอร์ที่ดี

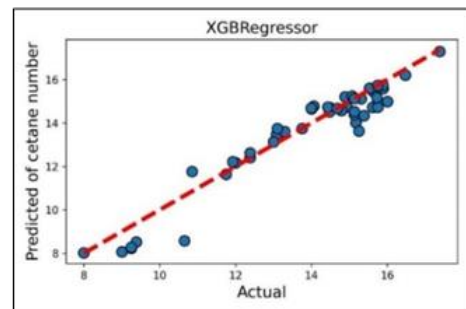
ที่สุดของอัลกอริทึม XGB และอัลกอริทึม RF โดยใช้โปรแกรม PyCharm ด้วยภาษาไพทอน ดังตารางที่ 2 จะนำค่าพารามิเตอร์ที่ดีที่สุดที่ได้จากการรันพารามิเตอร์ที่ดีที่สุด (Hyperparameter) ไปใช้กับอัลกอริทึม XGB และอัลกอริทึม RF

2.2.1 ข้อมูลอัลกอริทึม XGB

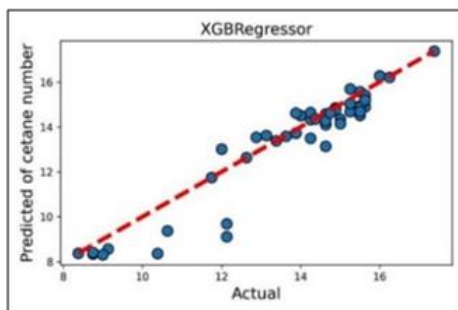
ข้อมูลผลการทำงานของโค้ดอัลกอริทึม XGB Regressor แสดงดังรูปที่ 10 และ ตารางที่ 3



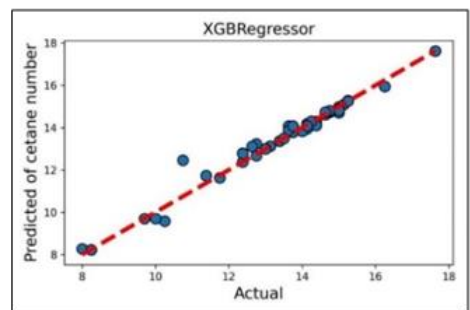
(ก) ด้านบน



(ข) ด้านล่าง



(ค) ด้านซ้าย



(ง) ด้านขวา

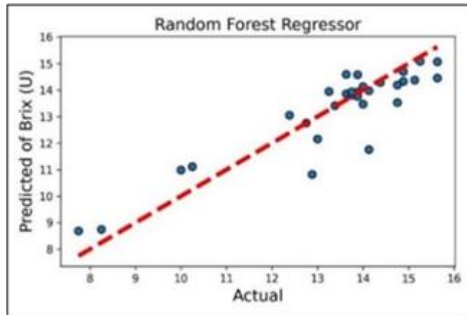
รูปที่ 10 ข้อมูลผลการทำงานของโค้ดอัลกอริทึม XGB Regressor ที่ตำแหน่ง (ก) ด้านบน (ข) ด้านล่าง (ค) ด้านซ้าย และ (ง) ด้านขวา

2.2.2 ข้อมูลอัลกอริทึม RF

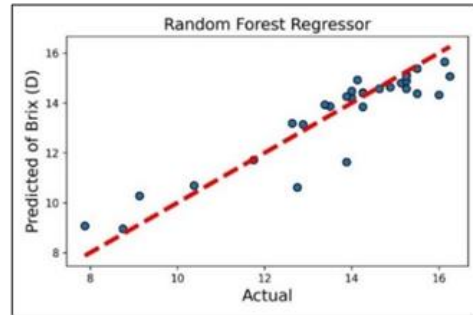
ข้อมูลผลการทำงานของโค้ดอัลกอริทึม RF แสดงดังรูปที่ 11 และ ตารางที่ 4

2.2.3 การเปรียบเทียบอัลกอริทึม

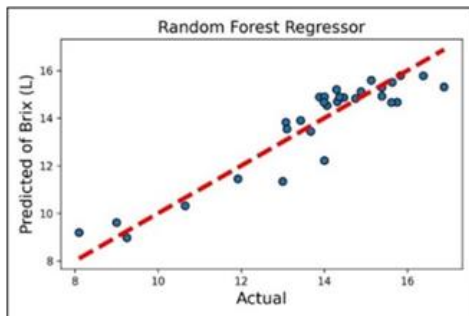
ในการเปรียบเทียบอัลกอริทึมจะตัดสินจากค่า Training R-squared ที่ได้จากผลการทำงานของโค้ดอัลกอริทึม XGB และอัลกอริทึม RF โดยใช้โปรแกรม PyCharm ด้วยภาษาไพทอน ดังรูปที่ 12 และ ตารางที่ 5



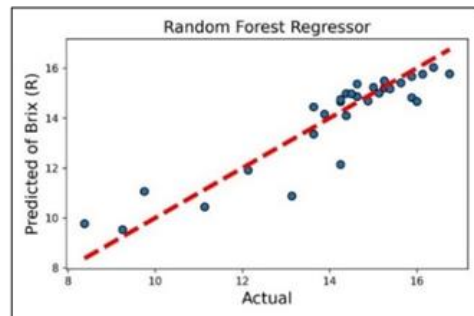
(ก) ด้านบน



(ข) ด้านล่าง

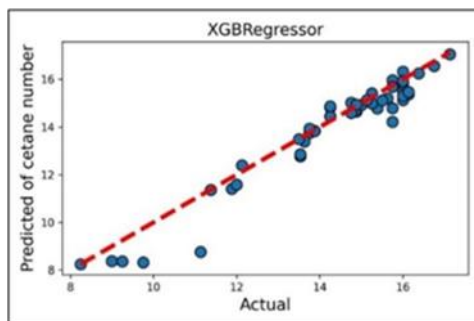


(ค) ด้านซ้าย

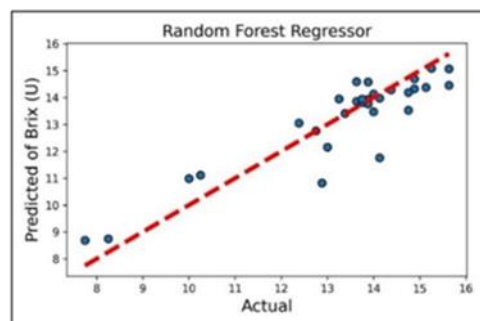


(ง) ด้านขวา

รูปที่ 11 ข้อมูลผลการทำงานของโค้ดอัลกอริทึม RF ที่ตำแหน่ง (ก) ด้านบน (ข) ด้านล่าง (ค) ด้านซ้าย และ (ง) ด้านขวา



(ก)



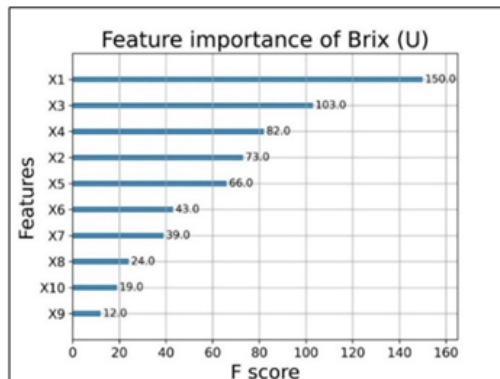
(ข)

รูปที่ 12 ผลการทำงานของโค้ดอัลกอริทึม (ก) XGB และ (ข) อัลกอริทึม RF

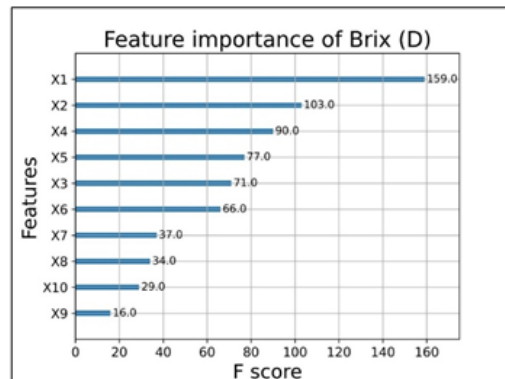
จากรูปที่ 12 (ก) มีการกระจายตัวของข้อมูลค่อนข้างน้อย แต่ในขณะที่ (ข) มีการกระจายตัวของข้อมูลมากกว่า หากพิจารณาจากรูปภาพอัลกอริทึม XGB จะเป็นอัลกอริทึมที่ดีที่สุด

3. ความสำคัญของคุณลักษณะ

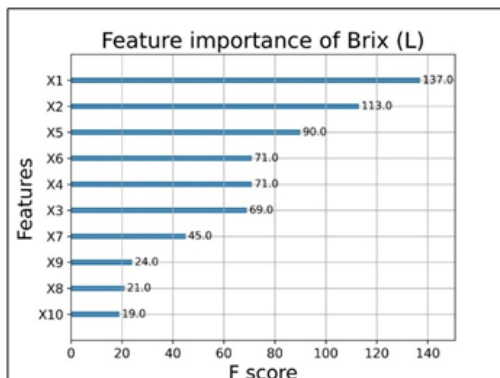
จากผลสรุปของอัลกอริทึม XGB เป็นอัลกอริทึมที่ดีที่สุดจึงได้นำอัลกอริทึม XGB มาใช้ในการเขียนโค้ดความสำคัญของคุณลักษณะ ดังรูปที่ 13



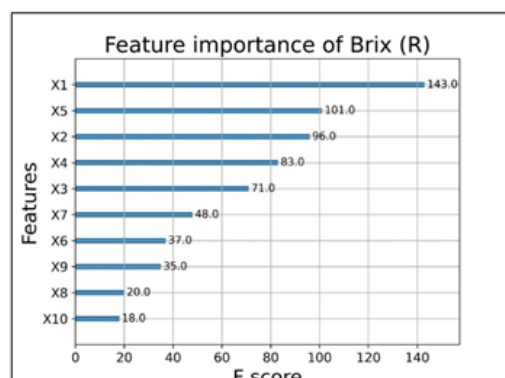
(ก) ด้านบน



(ข) ด้านล่าง



(ค) ด้านซ้าย

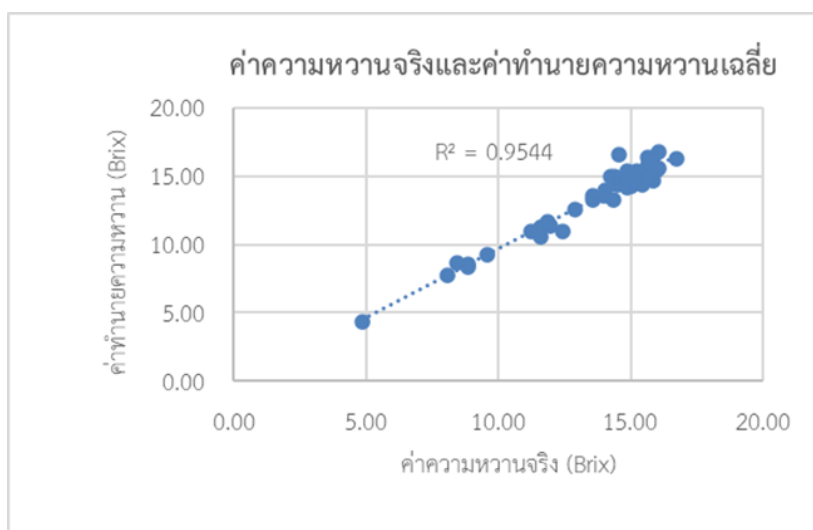


(ง) ด้านขวา

รูปที่ 13 คะแนนความสัมพันธ์ F-score ของข้อมูลขาเข้าและค่าความหวานของผลเมลอน (ก) ด้านบน (ข) ด้านล่าง (ค) ด้านซ้าย และ (ง) ด้านขวา

จากรูปที่ 13 คะแนนความสัมพันธ์ F-score ของข้อมูลขาเข้าและค่าความหวานด้านบน ด้านล่าง ด้านซ้าย และด้านขวาของผลเมลอน พบว่าน้ำหนักมี

คะแนนเฉลี่ยมากที่สุด 147.25 Brix รองลงมาคือเส้นรอบวงเฉลี่ย 96.25 Brix และค่าความเข้มข้นของผลเมลอนเฉลี่ยน้อยที่สุด 21.25 Brix



รูปที่ 14 กราฟความสัมพันธ์ระหว่างค่าความหวานจริงกับค่าทำนายความหวานเฉลี่ย

4. ผลการทำนายค่าความหวานของเมล่อน

ผลการทดสอบพบว่า ค่า R-squared ของการเรียนรู้โมเดล [10] ในอัลกอริทึม XGB คือ 99% ในขณะที่อัลกอริทึม RF คือ 95% จึงสรุปได้ว่าอัลกอริทึม XGB มีประสิทธิภาพในการเรียนรู้โมเดลได้ดีกว่าอัลกอริทึม RF โดยค่าน้ำหนักของเมล่อนมีข้อเด่นชัดสำคัญที่สุด และค่าสีของเมล่อนมีข้อเด่นชัดสำคัญน้อยที่สุด ในการแปลผลกับ

ค่าความหวานของเมล่อน สำหรับการทดสอบจริงได้นำแบบจำลองของอัลกอริทึม XGB มาทำนายค่าความหวานของเมล่อนจำนวน 50 ตัวอย่าง เมื่อเทียบกับค่าความหวานจริง ค่าความคลาดเคลื่อนอยู่ในช่วงระหว่าง 0.14–28.36 % และมีความคลาดเคลื่อนเฉลี่ยอยู่ที่ 4.90 % จึงถือว่าอยู่ในเกณฑ์ที่สามารถยอมรับได้ โดยเกณฑ์พิจารณาค่า R-Squared ระหว่างค่าจริงและค่าทำนายเท่ากับ 95.44

ตารางที่ 1 ค่าสถิติของชุดข้อมูล

ค่าสถิติการกระจาย ของข้อมูล	ค่าน้ำหนัก (กรัม)	เส้นรอบวง (มิลลิเมตร)	ขนาดก้าน (มิลลิเมตร)	ขนาดแกนกลาง (วงกลมเล็ก) (มิลลิเมตร)	ขนาดแกนกลางถึง ขอบตาข่าย (วงกลมใหญ่) (มิลลิเมตร)
ค่าต่ำสุด	856	112.10	3.65	13.10	15.75
ค่าสูงสุด	2570	165.65	11.10	30.15	36.25
ค่าเฉลี่ย	1612.92	137.00	6.74	20.44	26.66
ส่วนเบี่ยงเบนมาตรฐาน	355.07	37.05	1.87	4.68	4.56

ตารางที่ 2 เปรียบเทียบค่าพารามิเตอร์ที่ดีที่สุดระหว่างอัลกอริทึม XGB และอัลกอริทึม RF

Parameter	XGB		RF	
	ขอบเขต	ค่าที่ดีที่สุด	ขอบเขต	ขอบเขต
eta	0.05-0.5	0.4	eta	0.05-0.5
max_depth	4-10	6	max_depth	4-10
subsample	0.5-1	1	subsample	0.5-1
colsample_bytree	0.5-1	1	colsample_bytree	0.5-1
eval_metric	rmse	rmse	eval_metric	rmse
seed	42	42	seed	42
n_estimators	10-1001	120	n_estimators	10-1001
max_features	-	-	max_features	-
min_samples_split	-	-	min_samples_split	-

ตารางที่ 3 ผลของอัลกอริทึมที่ใช้ XGB

ML Methods	RF			
	ด้านบน	ด้านล่าง	ด้านซ้าย	ด้านขวา
Training R-squared	0.99	0.99	0.99	0.99
Test R-squared	0.92	0.91	0.86	0.97
MSE	0.39	0.43	0.68	0.11
RMSE of test	0.62	0.65	0.82	0.33
RMSE of train	0.00	0.00	0.00	0.00
Mean accuracy	0.85	0.77	0.86	0.91

ตารางที่ 4 ผลของอัลกอริทึมที่ใช้ RF

ML Methods	RF			
	ด้านบน	ด้านล่าง	ด้านซ้าย	ด้านขวา
Training R-squared	0.95	0.95	0.95	0.95
Test R-squared	0.81	0.81	0.81	0.81
MSE	0.69	0.69	0.69	0.69
RMSE of test	0.62	0.62	0.62	0.62
RMSE of train	0.59	0.59	0.59	0.59
Mean accuracy	0.61	0.61	0.61	0.61

ตารางที่ 5 เปรียบเทียบอัลกอริทึมที่ใช้กับความหวานของผลเมล่อน

ML Methods	XGB	RF
Training R-squared	0.99	0.95
Test R-squared	0.92	0.81
MSE	0.40	0.69
RMSE of test	0.61	0.62
RMSE of train	0.00	0.59
Mean accuracy	0.85	0.61

จากตารางที่ 5 ผลการทำงานของโค้ด Model ค่า Training R-squared ของอัลกอริทึม XGB มีค่าเท่ากับ 99% ขณะที่ค่า Training R-squared ของอัลกอริทึม RF มีค่าเท่ากับ 95% แสดงว่าอัลกอริทึม XGB เป็นอัลกอริทึมที่ดีที่สุด

ทำนายที่แม่นยำมากขึ้นและควรเก็บข้อมูลผลการทดลองในห้องปฏิบัติการที่ได้มาตรฐานเพิ่มขึ้นเพื่อควบคุมตัวแปรต่าง ๆ ให้สมบูรณ์ มีความคลาดเคลื่อนน้อยที่สุด เพื่อให้การตรวจสอบผลเมล่อนมีประสิทธิภาพมากที่สุด

สรุปผล

การตรวจสอบและทำนายความหวานผลเมล่อนแบบรวดเร็วโดยใช้เทคนิคการประมวลผลภาพร่วมกับแบบจำลองอัลกอริทึม เพื่อหาค่าความหวานของเมล่อน ซึ่งค่าความหวานของผลเมล่อนถือว่าอยู่ในเกณฑ์ที่สามารถยอมรับได้ ผลการทำงานของโค้ด Model ค่า Training R-squared ของอัลกอริทึม XGB มีค่าเท่ากับ 99% ขณะที่ค่า Training R-squared ของอัลกอริทึม RF มีค่าเท่ากับ 95% แสดงว่าอัลกอริทึม XGB เป็นอัลกอริทึมที่ดีที่สุด โดยเกณฑ์พิจารณาค่า R-Squared ระหว่างค่าจริงและค่าทำนายเท่ากับ 95.44% ทั้งนี้สามารถพัฒนาและนำไปใช้ประโยชน์ในการตรวจสอบผลเมล่อนแบบรวดเร็วให้มีประสิทธิภาพสูงขึ้นจะต้องดำเนินการดังนี้ ชุดข้อมูลต้องได้รับการเก็บข้อมูลเพิ่มเติมเพื่อให้มีผลการ

เอกสารอ้างอิง

1. Allkaset. Melon [Internet]. [Cited 2022 Jan 11]. Availability from: <https://shorturl.asia/nZwJa>
2. Eiamopas K. Digital image processing. introduction and principles. Kasetsart Engineering Journal (Thailand). 1992;6(18): 30-42. Thai.
3. Utanon W, Phongsuphap S, Sawangarom W, Singkleewan N. Non-destructive Image Processing Technique and Rule - based Classification in Estimating the Quality of Fruit. J Inform. 2020;19:129-38. Thai.

4. Visitsora-at P. Basic Metrics for Measuring the Performance of Machine Learning Models. [Internet]. [Cited 2019 Oct 29]. Availability from: <https://shorturl.asia/nrKXT>.
5. Van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, et al. Scikit-image: image processing in Python. *PeerJ*. 2014;2:453.
6. Lyer LS. AI enabled applications towards intelligent transportation. *Transport Eng*. 2021;5(1):100083.
7. Golberg M, Cho H. Introduction to Regression Analysis. 2010.
8. He T, Li S, Wu S, Li K. Small-signal Stability analysis for power system frequency regulation with renewable energy participation. *Math Probl Eng*. 2021;2021:5556062.
9. Yang Y, Zhang X, Yang L. Data-driven power system small-signal stability assessment and correction control model based on XGBoost. *Energy Reports*. 2022;8:710-17.
10. Zhu L, Zhou X, Zhang C. Rapid identification of high-quality marine shale gas reservoirs based on the oversampling method and random forest algorithm. *Artificial Intelligence in Geosciences*. 2021;2:76-81.