การพัฒนาตัวแบบวินิจฉัยโรคใบสับปะรดด้วยการสกัดคุณลักษณะภาพร่วมกับเหมืองข้อมูล

กิตติธร จันทนะ[1*] และ ภรัณยา ปาลวิสุทธิ์[1]

[1]สาขาวิชาวิทยาการข้อมูล คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏนครปฐม
*combo7238@gmail.com

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและพัฒนาตัวแบบวินิจฉัยโรคในสับปะรด ด้วยการสกัดคุณลักษณะภาพร่วมกับเทคนิคเหมืองข้อมูล โดยเก็บข้อมูลโรคในใบสับปะรด จำนวนทั้งหมด 150 รูป โดยแบ่งเป็น โรคเหี่ยว 50 รูป โรครากเน่า 50 รูป และใบปกติ 50 รูป โดยใช้อัลกอริทึม Color layout filter , Simple color histogram filter ร่วมกับ Multilayer Perceptron , Naïve Bayes , LMT , Random Forest การดำเนินงานได้ทำการเปรียบเทียบระหว่างการสกัดคุณลักษณะภาพโดยใช้ค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก และค่าความเหวี่ยง เพื่อวัดประสิทธิภาพของตัวแบบได้ ผลการดำเนินงานพบว่าอัลกอริทึม Simple Color Histogram Filter ร่วมกับ Naive Bayes 84.66 % ดีที่สุด รองลงมาคือ อัลกอริทึม Simple Color Histogram Filter ร่วมกับ LMT 84.66 % และ อัลกอริทึม Simple Color Histogram Filter ร่วมกับ Random Forest 84.66 % ตามลำดับ

คำสำคัญ : เหมืองข้อมูล โรคใบสับปะรด ฮีสโตแกรมสี

Development of a pineapple leaf disease diagnostic model by extracting visual features together with data mining.

Kittithorn Chantana[1*] and Paranya Palvisut[1]

[1]Data Science Faculty of Science and Technology Nakhon Pathom Rajabhat University
*combo7238@gmail.com

Abstract

The purpose of this research was to study and develop a diagnostic model in pineapple using extracting image features together and data mining techniques. The disease of pineapple leaves data was collected in 150 leaves, categorized as 50 wilts, 50 root rot, and 50 normal leaves. Using algorithm Color layout filter, Simple color histogram filter, together with Multilayer Perceptron, Naive Bayes, LMT and Random Forest. The operations were compared between image feature extraction. using the accuracy, precision, recall, and F-measure to measure the performance of the model. Results showed that Simple Color Histogram Filter with Naive Bayes 84.66% was the best, followed by the Simple Color Histogram Filter algorithm with LMT 84.66%, and the Simple Color Histogram Filter algorithm with the Random Forest 84.66%, respectively.

Keywords: data mining , pineapple disease , color histogram.

1.บทนำ

ปัจจุบันประเทศไทยเป็นประเทศเกษตรกรรม และการเกษตรกรรมถือเป็นอาชีพหลักของคนไทย ทำให้ประเทศไทยมีสินค้าทางเกษตรและผลผลิตทางเกษตรอยู่หลายอย่าง และก็มีสินค้าทางการเกษตรที่เป็นสินค้าส่งออก อาชีพเกษตรกรเป็นอาชีพหลักของประชากรในประเทศไทยทั้งการปลูก พืช เลี้ยงสัตว์และการประมง การปลูกพืชที่เกษตรกรนิยมปลูกมาก โดยเฉพาะในพื้นที่ภูมิภาคกลางคือ สับปะรด ซึ่งสับปะรดเป็นพืชที่มีความสำคัญ ต่อเศรษฐกิจชนิดหนึ่ง ที่มีผู้นิยมบริโภคและยังเป็นที่ต้องการของตลาด รวมทั้งยังเป็นพืชที่ส่งออก ไปขายต่างประเทศที่สำคัญในรูปแบบการบรรจุใส่กระป๋อง ในการปลูกสับปะรด พบว่าเกษตรกรจะประสบปัญหาต่าง ๆ ในการปลูก ไม่ว่าจะเป็นภัยธรรมชาติที่ไม่สามารถหลีกเลี่ยงได้ วัชพืช และศัตรูพืช และโรคที่ เกิดจากธรรมชาติ ที่เป็นปัญหาหลักของเกษตรกร ดังนั้นเกษตรกรจึงจำเป็นต้องใช้สารเคมีช่วยในการบำรุงดูแลรักษาเพื่อให้ได้ผลผลิตตามความต้องการของตลาด ในปัจจุบันวิธีการเลือกใช้สารเคมีของ เกษตรกรจะใช้วิธีที่เกษตรกรเป็นผู้บอกเล่าอาการและลักษณะที่เกิดขึ้นแก่เจ้าของร้านหรือพนักงาน ขายในร้านเคมีการเกษตร ดั้งนั้นเจ้าของร้านหรือพนักงานขายจะเป็นผู้จัดผลิตภัณฑ์เคมี จึงไม่สามารถและวินิจฉัยโรคได้อย่างมีประสิทธิภาพ อาจทำให้เกิดข้อผิดพลาดเนื่องจากการบอกอาการไม่ ถูกต้อง หรือพนักงานไม่มีความเชี่ยวชาญในการวินิจฉัยโรค ซึ่งสาเหตุเหล่านี้อาจจะส่งผลต่อการนำ

สารเคมี เหล่านี้ไปใช้ไม่ได้ผล หรือบางครั้งจะต้องใช้เวลาในการรักษานานขึ้นเนื่องจากสารเคมีที่นำไปใช้ไม่สามารถรักษาโรคได้ อาจทำให้เกิดความสิ้นเปลืองและอาจทำให้เกิดความเสียหายกับพืชได้

ดังนั้นงานวิจัยนี้จึงได้พัฒนาตัวแบบวินิจฉัยโรคที่เกิดกับใบสับปะรดด้วยการสกัดคุณลักษณะภาพร่วมกับเหมืองข้อมูล ซึ่งเทคนิคนี้จะใช้วินิจฉัยว่า อาการผิดปกติที่เกิดขึ้นกับสับปะรดนั้นเป็นโรคอะไร เพื่อเกษตรกรจะสามารถทราบชนิดโรคก่อนที่จะไปขอคำปรึกษาจาก เจ้าของร้านหรือพนักงานร้านเคมีการเกษตร หรือผู้เชี่ยวชาญในการเลือกซื้อสารเคมี

2.วัตถุประสงค์

2.1 เพื่อศึกษาและพัฒนาตัวแบบวินิจฉัยโรคในใบสับปะรด ด้วยการสกัดคุณลักษณะภาพร่วมกับเทคนิคเหมืองข้อมูล

2.2 เพื่อเปรียบเทียบประสิทธิภาพตัวแบบวินิจฉัยโรคในใบสับปะรด ด้วยการสกัดคุณลักษณะภาพร่วมกับเทคนิคเหมืองข้อมูล

3.ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

3.1 การทำเหมืองข้อมูล [1] เป็นกระบวนการในการค้นหารูปแบบในชุดข้อมูลขนาดใหญ่ โดยใช้วิธีการของการเรียนรู้ของเครื่อง สถิติ และระบบฐานข้อมูล การทำเหมืองข้อมูลเป็นขั้นตอนวิธีการในการ การค้นหาความรู้ในฐานข้อมูล การทำเหมืองข้อมูลเป็นเทคนิคเพื่อค้นหารูปแบบ ของจากข้อมูลจำนวนมหาศาลโดยอัตโนมัติ โดยใช้ขั้นตอนวิธีจากวิชาสถิติ การเรียนรู้ของเครื่อง และ การรู้จำแบบ หรือในอีกนิยามหนึ่งการทำเหมืองข้อมูล คือ กระบวนการที่กระทำกับข้อมูล (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้น โดยอาศัยหลักสถิติ การรู้จำ การเรียนรู้ของเครื่อง และหลักคณิตศาสตร์

3.2 มัลติเลเยอร์ เปอร์เซปทรอน (**Multilayer Perceptron**) [2] เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่มีโครงสร้างเป็นแบบหลายๆชั้น ใช้สำหรับงานที่มีความซับซ้อนได้ผลเป็นอย่างดี โดยมีกระบวนการฝึกฝนเป็นแบบมีผู้สอนและใช้ขั้นตอนการส่งค่าย้อนกลับ สำหรับการฝึกฝนกระบวนการส่งค่าย้อนกลับ ประกอบด้วย 2 ส่วนย่อยคือ การส่งผ่านไปข้างหน้า การส่งผ่านย้อนกลับ

3.3 เนอีฟเบย์ (Naive Bayes) [3] เป็นเทคนิคง่ายๆ ในการสร้างตัวแยกประเภท: โมเดลที่กำหนด class label ให้กับอินสแตนซ์ที่มีปัญหา แสดงเป็นเวกเตอร์ของค่าคุณลักษณะ โดยที่ class label ถูกดึงมาจากชุดที่มีขอบเขตจำกัด ไม่มีอัลกอริธึมเดียวสำหรับการฝึกตัวแยกประเภท แต่มีกลุ่มของอัลกอริธึมตามหลักการทั่วไป : ตัวแยกประเภท Bayes ไร้เดียงสา ทั้งหมดถือว่าค่าของคุณลักษณะเฉพาะนั้นไม่ขึ้นกับค่าของคุณลักษณะอื่นๆ เมื่อพิจารณาจากตัวแปรคลาส

3.4 แอลเอ็มที (LMT) [4] การจำแนกรูปแบบที่มีการเชื่อมโยงและผสมผสานการถดถอยโลจิสติกส์ (LR) และการตัดสินใจการเรียนรู้ในการตัดสินของต้นไม้

3.5 แรนดอมฟอเรส (Random Forest) [5] เป็นหนึ่งในกลุ่มของโมเดลที่เรียกว่า Ensemble learning ที่มีหลักการคือการเทรนโมเดลที่เหมือนกันหลายๆ ครั้ง (หลาย Instance) บนข้อมูลชุดเดียวกัน โดยแต่ละครั้งของการเทรนจะเลือกส่วนของข้อมูลที่เทรนไม่เหมือนกัน แล้วเอาการตัดสินใจของโมเดลเหล่านั้นมาโหวตกันว่า Class ไหนถูกเลือกมากที่สุด

3.6 การวัดประสิทธิภาพแบบจำลอง [6] ในการวัดประสิทธิภาพของแบบจำลองในแต่ละเทคนิควิธีต่างๆ โดยใช้ค่าความแม่นยำของแบบจำลอง (Accuracy) จะต้องทำการเลือกข้อมูลสำหรับเรียนรู้ (Training Set) และข้อมูลสำหรับทดสอบ (Testing Set) งานวิจัยนี้เลือกใช้วิธีการสุ่มเลือกข้อมูลแบบความเที่ยงตรง 3 กลุ่ม (3-fold Cross Validation) การสุ่มข้อมูลแบบความเที่ยงตรงเป็นการสุ่มตัวอย่างข้อมูลโดยเริ่มจากการแบ่งชุดข้อมูลออกเป็นส่วนๆ เท่า ๆ กัน นำข้อมูลบางส่วนมาทำการ

เรียนรู้ และนำข้อมูลส่วนที่เหลือมาทาการทดสอบแบบจำลองที่ได้จากเรียนรู้ โดยในการทำงานจะทำการแบ่งกลุ่มข้อมูลออกเป็นสิบชุดเท่าๆ กันในการทดลองครั้งแรกจะเลือกข้อมูลชุดที่หนึ่ง เป็นข้อมูลชุดทดสอบและข้อมูลที่เหลือเป็นชุดเรียนรู้ในการทดลองครั้งถัดไปจะเลือกข้อมูลชุดที่สอง เป็นข้อมูลชุดทดสอบและข้อมูลชุดที่เหลือเป็นชุดเรียนรู้ ทำจนกระทั้งข้อมูลได้ถูกนำมาเป็นข้อมูลทดสอบและชุดเรียนรู้ ซึ่งจะมีการทดลองทั้งหมด 3 ครั้ง การวัดค่าประสิทธิภาพของแบบจำลองนั้นจะอาศัย Confusion matrix

3.7 ฮีสโตแกรมสี [6] ประกอบด้วยค่าความน่าจะเป็นของสีระดับต่าง ๆ ซึ่งจะมีจำนวนของสีในแต่ละระดับที่ปรากฏในภาพ ฮีสโตแกรมสามารถอธิบายความมืด ความสว่าง ความคมชัดสูงหรือต่ำโดยทั่วไปแล้วรูปภาพแต่ละรูปจะมีฮีสโตแกรมที่แน่นอน ซึ่งฮีสโตแกรมจะมีค่าไม่เปลี่ยนแปลงเมื่อมีการเคลื่อนย้าย (Translation) หมุน (Rotation) แกนของรูปภาพ และการเปลี่ยนมุมเพียงเล็กน้อย ดังนั้นจึงเป็นไปได้ที่จะนำฮีสโตแกรมสีมาใช้สำหรับเปรียบเทียบความคล้ายคลึงของรูปภาพในการคำนวณค่าฮีสโตแกรมสี ภาพแต่ละภาพจะถูกควอนไทซ์สี ภายในภาพเพื่อลดมิติของเวกเตอร์และลดความซับซ้อนในการคำนวณลง โดยการแบ่งกลุ่มสีออกเป็น m ถังสี (Bins) ซึ่งส่วนใหญ่นิยมใช้ 32, 64 หรือ 256 ถังสี เนื่องจากการแยกแยะความแตกต่างของระดับค่าสีของสายตามนุษย์มีความละเอียดไม่มากนัก ไพโรจน์

3.8 งานวิจัยที่เกี่ยวข้อง

ชื่นนภา และวงกต [7] ได้พัฒนาระบบวินิจฉัยโรคพริกโดยใช้เทคนิคจำแนกข้อมูล ซึ่งเทคนิคที่นำมาประยุกต์ใช้คือต้นไม้ตัดสินใจ ข้อมูลที่นำมาใช้ในการวินิจฉัยโรคพริกเป็นข้อมูลที่ใช้ได้มาจากการเก็บข้อมูลโดยแบบสอบถามจากบุคลากรในร้านทรัพย์มงคล ระบบนี้ถูกพัฒนาด้วย ภาษาพีเอชพี (PHP) ร่วมกับฐานข้อมูลมายเอสคิวแอล (MySQL) ผลการพัฒนาระบบ พบวาระบบสามารถวินิจฉัยโรคพริกและแนะนำสารเคมีที่ช่วยในการรักษาโรคได้ สวนการทดสอบความถูกต้องของโมเดลโดย ใช้โปรแกรม Weka 3.7 พบว่าโมเดลที่ใช้ในการวินิจฉัยโรคพริกมีค่าความถูกต้องเท่ากับ 98.90% ดังนั้นระบบที่พัฒนาขึ้นนี้มีผลประเมินคาความพึงพอใจในระดับดีและสามารถนำไปประยุกต์ใช้งานได้

ประภาส และคณะ [8] ได้พัฒนาระบบการจำแนกสาเหตุของการเกิดโรคในมันสำปะหลังตามลักษณะอาการ โดยวิธีการเหมืองข้อมูล โดยใช้กฎผังต้นไม้โดยใช้โปรแกรม weka ทำการตรวจสอบความถูกต้อง ของกฎ โดยมีผู้เชี่ยวชาญด้านโรคข้าวจำนวน 10คน เป็นผู้ตรวจสอบความถูกต้อง โดยจะทดสอบความถูกต้อง โดยใช้ตัวอย่างของโรคข้าว 20 ข้อ มาเปรียบเทียบเพื่อทำการวินิจฉัย พบว่าการวินิจฉัยโรคขาวมีความถูกต้อง 94.5% ดังนั้นระบบที่พัฒนาขึ้นนี้มีผลประเมินค่าความพึงพอใจในระดับดี และสามารถนำไปประยุกต์ใช้งานได้

ไพโรจน์ [6] ได้พัฒนาระบบการจำแนกสาเหตุของการเกิดโรคในแตงกวาตามลักษณะอาการ ซึ้งแบ่งได้แก่ โรคราน้ำค้าง โรคผลเน่า โรคใบด่าง โรคราแป้ง และใบปกติ ซึ่งเทคนิคในงานวิจัยฉบับนี้ใช้เทคนิคต้นไม้ตัดสินใจร่วมกับการสกัดคุณลักษณะภาพ ด้วยอัลกอริทึม Simple color histogram filter กับ Color layout filter และนำข้อมูลเข้าสู่กระบวนการจำแนกข้อมูลด้วยเทคนิคต้นไม้ตัดสินใจได้แก่ J48,RandomForest, RandomTree และ HorffdingTree การดำเนินงานได้ทำการเปรียบเทียบระหว่างการสกัดคุณลักษณะภาพด้วย Simple color histogram filter กับ Color layout filter เพื่อวัดประสิทธิภาพของตัวแบบ ข้อมูลที่สกัดคุณลักษณะด้วย Color layout filter กับเทคนิค RandomForest ให้ค่าความถูกต้องสูงสุดร้อยละ 92.5 ให้ค่าความแม่นยาสูงสุดร้อยละ 92.5 ให้ค่าความระลึกสูงสุดร้อยละ 92.5 ให้ค่าความระลึกสูงสุดร้อยละ92.5 ดังนั้นตัวแบบการจำแนกข้อมูลโรคแตงกวาที่ได้จากการสกัดคุณลักษณะภาพด้วยอัลกอริทึมColor layout filter ร่วมกับการจำแนกประเภทข้อมูลด้วย RandomForest มีเปอร์เซ็นต์ความถูกต้องมากที่สุด 92.5% จึงเป็นตัวแบบที่เหมาะสมต่อการนำไปประยุกต์ใช้สำหรับการวินิจฉัยโรคแตงกวาต่อไป

4.วิธีดำเนินงานวิจัย

วิธีการและขั้นตอนการพัฒนาตัวแบบวินิจฉัยโรคในใบสับปะรด ด้วยฮีสโตแกรมสีและจำแนกด้วยเทคนิคเหมืองข้อมูล มาใช้ในงานวิจัย โดยผู้วิจัยได้ทำการนำภาพโรคในใบสับปะรด มาแปลงเป็นฮีสโตแกรม จากนั้นนำข้อมูลที่ได้พัฒนาเป็นตัวแบบวินิจฉัยจำแนกประเภทด้วยเทคนิคต้นไม้ตัดสินใจโดยใช้โปรแกรม weka มีรายละเอียดและขั้นตอนดังนี้

4.1 เก็บรวบรวมข้อมูลโรคสับปะรด

โรคเหี่ยว [9] อาการแรกนั้นเกิดกับระบบรากก่อน โดยรากจะไม่มีการสร้างเซลล์ส่วนปลายราก ชะงักการเจริญเติบโต รากจะไม่ทำงานและเซลล์จะตาย ซึ่งต่อมาเนื้อเยื่อส่วนราก จะเน่า แล้วสับปะรดจะแสดงอาการให้เห็น ทางส่วนปลายใบและตัวใบในเวลาต่อมา คือใบจะอ่อนนิ่ม มีเขียวอ่อน หรือสีเหลืองอ่อน ปลายใบแห้งเป็นสี น้ำตาลจนถึงสีแดงลามสู่โคนใบ ใบลู่ลง แผ่แบน ไม่ตั้งขึ้นเหมือนใบปกติ ต่อมาต้นเหี่ยวและแห้ง รากสั้นกุด ถอนต้นง่าย การทำลายเริ่มตั้งแต่ ปลูกจนถึงเก็บเกี่ยว ทั้งในแปลงต้นปลูก และแปลงต้นตอ แสดงอาการเด่นชัดหลัง การบังคับดอก ผลสับปะรดจะไม่พัฒนา มีขนาดเล็ก การทำลาย ทำให้คุณภาพ และผลผลิตเสียหายมาก ไม่สามารถ เก็บเกี่ยวผลผลิตได้

ภาพที่ 1 โรคเหี่ยวสับปะรด

โรครากเน่า [9] ใบยอดแสดงอาการปลายซีดเหลือง ดึงออกได้ง่าย เพราะโคนใบเน่าเป็นสีน้ำตาล อาการเน่าจะลุกลามสู่ ส่วนล่างของลำต้น และเชื้อราในดินจะทำให้รากเน่า เชื้อ P. parasitica ทำให้ลำต้นชะงักการเจริญเติบโต ทำให้ รากเน่าน้อย แต่เชื้อรา ทำให้รากเน่ามาก และทำให้ผลเน่าเป็นจุดสีเขียวเข้ม เมื่อผ่าดูจะพบว่าเนื้อเยื่อภายในจะเน่าเป็นสีน้ำตาล

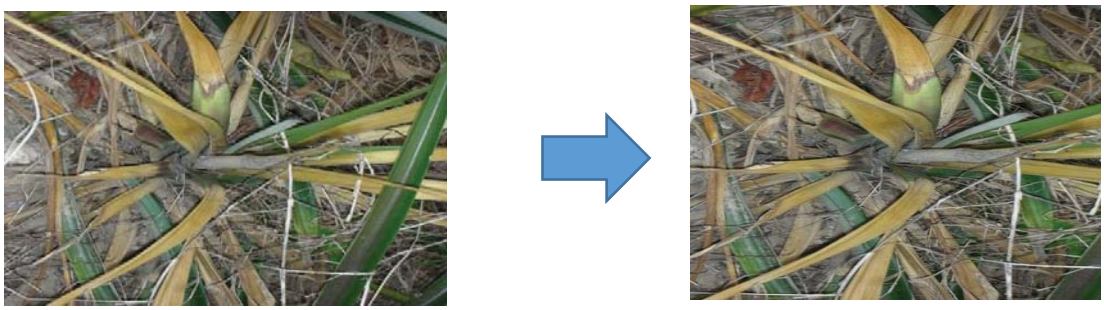
ภาพที่ 2 โรครากเน่าสับปะรด

ใบปกติ เป็นใบสับปะรดที่สมบูรณ์ไม่เป็นโรค

ภาพที่ 3 ต้นปกติสับปะรด

4.2 ขั้นตอนการกรองข้อมูล

ผู้วิจัยนำข้อมูลรูปภาพจากขั้นตอนการเก็บข้อมูลมาทำการกรองข้อมูลโดยการนำรูปภาพมาตัดขอบส่วนเกิน



ภาพที่ 4 ตัวอย่างการกรองข้อมูล

4.3 ขั้นตอนการแปลงข้อมูล

ต่อจากขั้นตอนการกรองข้อมูลก็จะเป็นขั้นตอนการแปลงข้อมูลให้อยู่ในรูปแบบของไฟล์ Arff โดยมีรูปแบบไฟล์ดังนี้

@relation คือ ชื่อข้อมูล

@attribute filename string คือ attribute ชื่อไฟล์รูปภาพ

@attribute class {normal, wilt, top rot, root rot } คือ ชื่อ

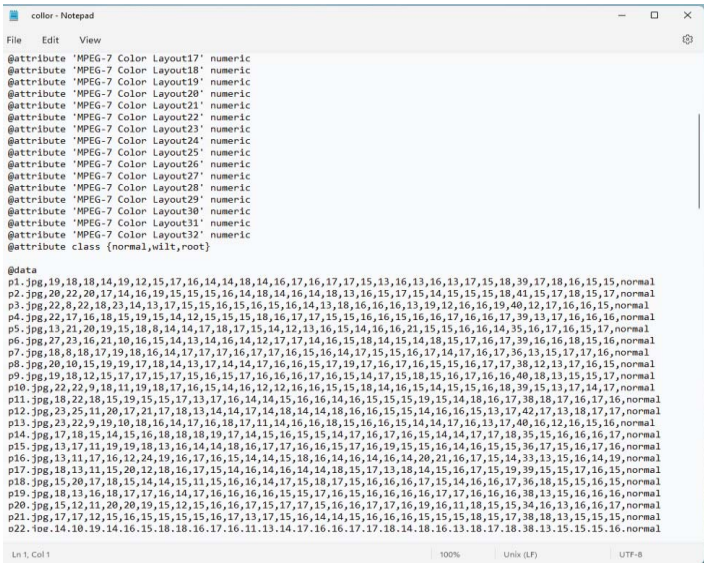
class

@data

Normal1.jpg,Normal คือ ชื่อไฟล์รูปภาพ,ชื่อclass

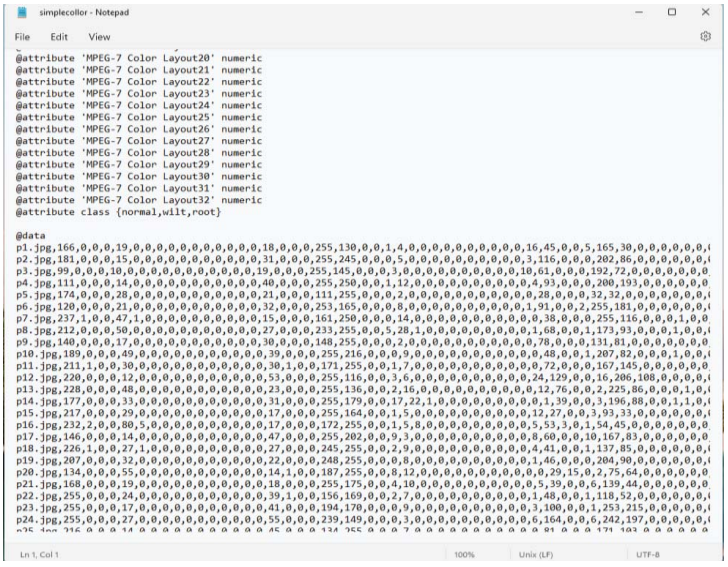
4.4 ขั้นตอนการสร้างตัวแบบ

ในการสร้างตัวแบบงานวิจัยนี้ได้ใช้โปรแกรม weka ซึ่งได้มีการรวบรวมอัลกอริทึมที่จำเป็นสำหรับงานวิจัยนี้ สำหรับขั้นตอนนี้จะทำการแปลงข้อมูลรูปภาพเป็นข้อมูลฮีสโตแกรมสี ได้เลือกอัลกอริทึม Color layout filter และ Simple color histogram filter



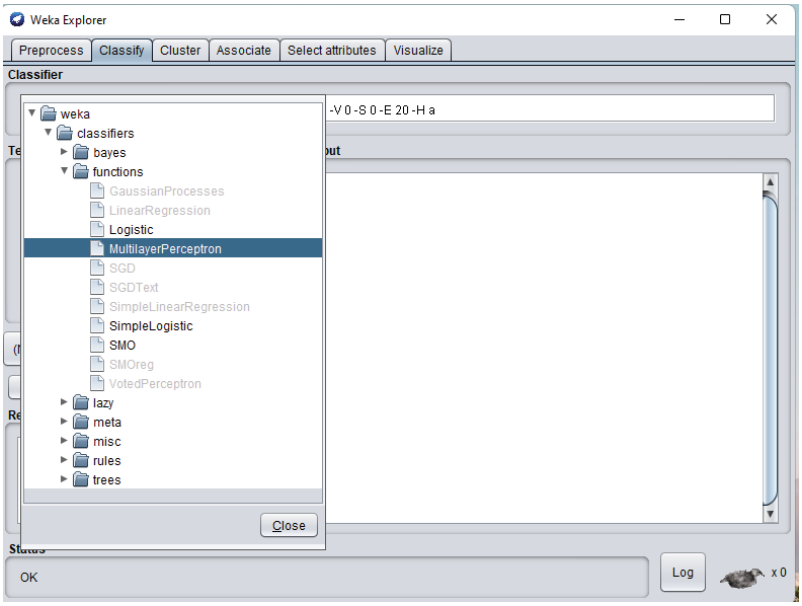
ภาพที่ 5 ข้อมูลหลังการประมวลผลข้อมูลด้วยอัลกอริทึม Color layout filter

ภาพที่ 5 คือ ข้อมูลหลังการสกัดคุณลักษณะภาพด้วย Color layout filter



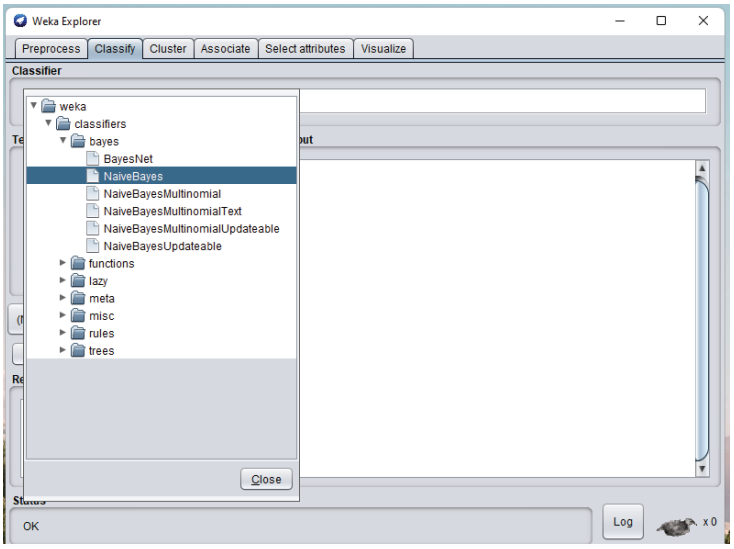
ภาพที่ 6 ข้อมูลหลังการประมวลผลข้อมูลด้วยอัลกอริทึม Simple color histogram filter

ภาพที่ 6 คือ ข้อมูลหลังการสกัดคุณลักษณะภาพด้วยอัลกอริทึม Simple color histogram filter



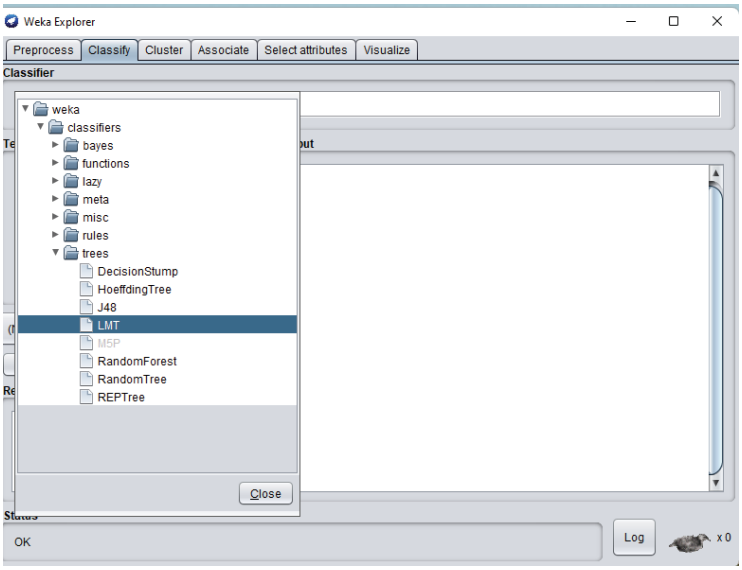
ภาพที่ 7 ขั้นตอนจำแนกข้อมูลอัลกอริทึม MultilayerPerceptron

ภาพที่ 7 คือ ขั้นตอนจำแนกข้อมูลอัลกอริทึม Function.MultilayerPerceptron เลือก Classify , Function. แล้วเลือก MultilayerPerceptron



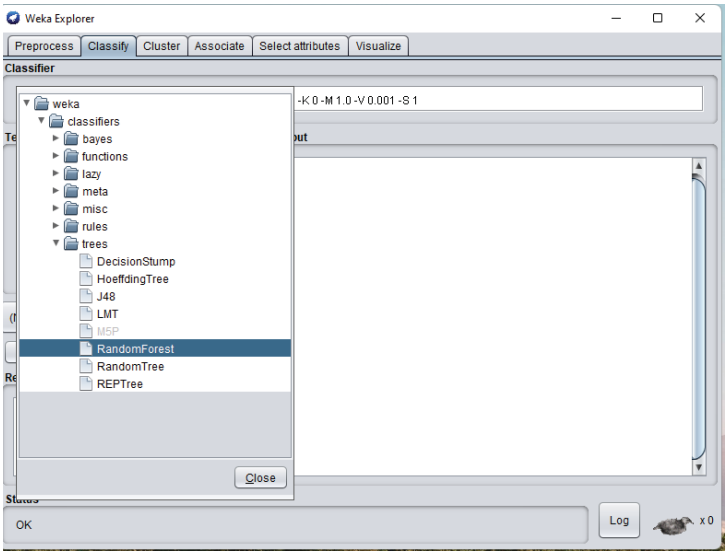
ภาพที่ 8 ขั้นตอนจำแนกข้อมูลอัลกอริทึม NaiveBayes

ภาพที่ 8 คือ ขั้นตอนจำแนกข้อมูลอัลกอริทึม bayes.NaiveBayes เลือก Classify , bayes แล้วเลือก NaiveBayes



ภาพที่ 9 ขั้นตอนจำแนกข้อมูลอัลกอริทึม LMT

ภาพที่ 9 คือ ขั้นตอนจำแนกข้อมูลอัลกอริทึม tree.LMT เลือก Classify , tree แล้วเลือก LMT



ภาพที่ 10 ขั้นตอนจำแนกข้อมูลอัลกอริทึม Random Forest

ภาพที่ 10 คือ ขั้นตอนจำแนกข้อมูลอัลกอริทึม Random Forest เลือก Classify , tree แล้วเลือก RandomForest

5.ผลการวิจัย

จากการจำแนกประเภทข้อมูลด้วยอัลกอริทึมต่าง ๆ สามารถสรุปผลได้ดังตารางที่ 1

ตารางที่ 1 ผลการจำแนกประเภทข้อมูลของอัลกอริทึมแต่ละชนิด

อัลกอริทึมก่อนการประมวลผล (Preprocess)	อัลกอริทึมจำแนกประเภทข้อมูล (Classify)	ค่าความถูกต้อง (Accuracy)	ค่าความแม่นยำ (Precision)	ค่าความลึก (Recall)	ค่าความเหวี่ยง (F - Measure)
Color layout filter	Multilayer Perceptron	74.00 %	73.32 %	74.03 %	73.52 %
	Naive Bayes	74.00 %	74.12 %	74.03 %	74.02 %
	LMT	80.00 %	80.40 %	80.01 %	80.08 %
	Random Forest	75.33 %	74.83 %	75.31 %	74.52 %
Simple color Histogram filter	Multilayer Perceptron	84.00 %	84.01 %	84.02 %	83.92 %
	Naive Bayes	84.66 %	84.60 %	84.70 %	84.69 %
	LMT	84.66 %	84.60 %	84.70 %	84.60 %
	Random Forest	84.66 %	84.50 %	84.70 %	84.33 %

6.สรุปผล

จากผลการเปรียบเทียบอัลกอริทึมจำแนกประเภทข้อมูล พบว่าอัลกอริทึม Simple color histogram filter ร่วมกับ Naive Bayes , LMT , Random Forest มีเปอร์เซ็นความถูกต้องเท่ากัน แต่ที่เลือก Naive Bayes เพราะว่าเมื่อเทียบค่าในตาราง ที่ 1 Naive Bayes มีเปอร์เซ็นค่าความเหวี่ยงที่ดีกว่า รองลงมาคือ อัลกอริทึม Simple color histogram filter ร่วมกับ LMT 84.66 % และ อัลกอริทึม Simple color histogram filter ร่วมกับ Random Forest 84.66 % , อัลกอริทึม Simple color histogram filter ร่วมกับ Multilayer Perceptron 84 % , อัลกอริทึม Color layout filter ร่วมกับ LMT 80 % , อัลกอริทึม Color layout filter ร่วมกับ Random Forest 75.33 % , อัลกอริทึม Color layout filter ร่วมกับ Naive Bayes 74 % , อัลกอริทึม Color layout filter ร่วมกับ Multilayer Perceptron 74 %

7.ข้อเสนอแนะ

7.1 การศึกษาในครั้งนี้ได้เก็บข้อมูลโรคสับปะรด ซึ่งยังไม่ครอบคลุมทุกโรค และใช้ภาพในการสร้างตัวแบบอย่างละ 50 ดังนั้นการศึกษาในครั้งต่อไปควรเก็บข้อมูลโรคสับปะรดแบบต่างๆ และเพิ่มภาพข้อมูลแต่ละโรคให้มากขึ้น

7.2 นำตัวแบบในการศึกษาในครั้งนี้ ไปพัฒนาเป็นแอปพลิเคชัน เพื่อวิเคราะห์โรคสับปะรด

8.เอกสารอ้างอิง

[1] ชิดชนก ส่งศิริ ธนาวินท์ รักธรรมานนท์ และกฤษณะ ไวยมัย. (2560). **การใช้เทคนิค Data Mining เพื่อค้นหาภาควิชา**

ที่เหมาะสมที่สุดให้กับนิสิต. กรุงเทพฯ : มหาวิทยาลัยเกษตรศาสตร์.

[2] Wikipedia. **multilayer perceptron** ค้นเมื่อ [20 ธันวาคม 2565] จาก https.//th.wikipedia.org/wiki/

[3] อนัตต์ชัย ชุติภาสเจริญ และ ดร.จรัญ แสนราช.(2561). **รายงานวิจัยเรื่องการเปรียบเทียบประสิทธิภาพของอลักอริทึมและการคัดเลือกคุณลักษณะที่ เหมาะสมเพื่อการพยากรณ์โอกาสความสำเร็จในการโอนเงินข้ามประเทศ ของบุคคลทั่วไป.** ขอนแก่น : มหาวิทยาลัยขอนแก่น.

[4] เพชรรัตน์ ม่วงน้อย จักรพันธ์ พลาผล และ ภรัณยา ปาลวิสุทธิ์. (2564). **รายงานวิจัยเรื่องตัวแบบประเมินภาวะความเสี่ยงการเป็นโรคซึมเศร้าของนักศึกษา ด้วยเทคนิคเหมืองข้อมูล.** นครปฐม : มหาวิทยาลัยราชภัฏนครปฐม.

[5] ชิตพงษ์ กิตตินราดา. (2563). **Random Forest** ค้นเมื่อ [12 มกราคม 2565] จาก https://guopai.github.io/ml-blog10.html

[6] ไพโรจน์ จงดี. (2562). **การพัฒนาตัวแบบวินิจฉัยโรคแตงกวาด้วยการสกัดคุณลักษณะภาพร่วมกับต้นไม้ตัดสินใจ.**ราชบุรี.

[7] ชื่นนภา บุญตาเพศ และ วงกต ศรีอุไร. (2560). **ระบบวินิจฉัยโรคพริกโดยใช้เทคนิคการจำแนกข้อมูล**.ค้นเมื่อ [12 มกราคม 2565] จาก URL:http://202.44.34.144/nccitedoc/admin/nccit_files/NCCIT-20142810140420.pdf

[8] ประภาส ตั้งธนาบด อุสมาน หวังหมัน และมูหามัดนูร มะเส็ง. (2560). **ได้พัฒนาระบบการจำแนกสาเหตุของการเกิดโรคในมันสำปะหลังตามลักษณะอาการ โดยวิธีการเหมืองข้อมูล.**ค้นเมื่อ [12 มกราคม 2565] จาก http://aucc.csit.rru.ac.th/Upload/205-124-cameraredy.pdf

[9] โรคสับปะรด (2565). **บทที่ 6 การจัดการโรค แมลง และวัชพืช.** ค้นเมื่อ [12 มกราคม 2565] จาก http://www.samrancom.com/บทที่%206%20การจัดการโรค%20แมลง%20และวัชพืช.pdf