บทที่ 1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

สายพันธุ์พืชในโลกนี้มีหลายชนิดจากการสำรวจที่ได้รับการตั้งชื่อมีประมาณ 250,000 – 270,000 สายพันธุ์ (Guo et al., 2004) ในการที่จะระบุชนิดของพันธุ์พืชให้ถูกต้องจากการที่ได้พบเห็นในที่ต่างๆ เป็นเรื่องที่ยากมากเนื่องจากสายพันธุ์พืชมีมากมายหลายสายพันธุ์ดังนั้นการที่จะระบุชนิดสายพันธุ์พืชได้ถูกต้องๆ ต้องอาศัยผู้ที่มีความเชี่ยวชาญเกี่ยวกับสายพันธุ์พืชโดยตรง ปัจจุบันคอมพิวเตอร์ถูกพัฒนาให้สามารถทำงานด้านการตัดสินใจที่ฉลาดมากขึ้นทำให้ได้มีการนำเทคโนโลยีคอมพิวเตอร์เข้ามาช่วยในการจำแนกพันธุ์พืช ทำให้ช่วยย่น ระยะเวลา และมีความแม่นยำในการระบุชนิดพันธุ์พืชได้

การคำนวณเชิงวิวัฒนาการ (evolutionary computation) ใช้สำหรับค้นหาผลลัพธ์ของปัญหาที่มีความซับซ้อน เช่น การคำนวณสัดส่วนของสารเคมีเพื่อให้ได้องค์ประกอบของสารที่มีประสิทธิภาพ การบริหารทรัพยากรในการผลิตให้มีประสิทธิภาพ วิธีการหาผลลัพธ์ที่ดีที่สุด ปัญหาต่างๆเหล่านี้ใช้การคำนวณเชิงวิวัฒนาการทั้งสิ้น การคำนวณเชิงวิวัฒนาการมีหลายวิธีแบบที่ใช้กันแพร่หลาย คือ ขั้นตอนวิธีเชิงพันธุกรรม (Genetic Algorithm : GA) เป็นวิธีค้นหาคำตอบที่ดีที่สุดโดยใช้หลักการคัดเลือกแบบธรรมชาติและหลักการทางสายพันธุ์ การทำงานของ GA ใช้การจำลองปัญหาให้อยู่ในรูปแบบเหมือนสิ่งมีชีวิตที่ประกอบด้วยเซลล์ ซึ่งแต่ละเซลล์จะมีกลุ่มโคโมโซม (Chromosomes) ที่เหมือนกัน โดยแต่ละโคโมโซมจะประกอบไปด้วยยีน (Gene) ซึ่งจะสะท้อนลักษณะเฉพาะของสิ่งมีชีวิตนั้น เช่น สีตา สีผม เป็นต้น การค้นหาผลลัพธ์ของปัญหาโดยการใช้ขั้นตอนวิธีทางพันธุกรรม จึงเปรียบเสมือนกับการนำพารามิเตอร์ของปัญหาแทนในรูปแบบสิ่งมีชีวิตและกำหนดให้สภาพแวดล้อมเป็นปัญหาหรือฟังก์ชั่นที่ต้องการค้นหาพารามิเตอร์ที่เหมาะสม ปัจจุบันเป็นที่ยอมรับในประสิทธิภาพและมีการนำไปประยุกต์ใช้อย่างกว้างขวางในการแก้ปัญหาการหาค่าเหมาะสมที่สุด

ด้วยคุณลักษณะของใบพืชประกอบไปด้วย รูปใบ (Leaf Shape) เส้นใบ (Leaf Vein) ปลายใบ (Leaf apex) โคนใบ (Leaf base) ขอบใบ (Leaf margin) ขนาดของใบ (Leaf Size) จากคุณลักษณะดังกล่าว ผู้วิจัยจึงมีแนวความคิดในการนำขั้นตอนวิธีเชิงพันธุกรรมมาประยุกต์ใช้ในการจำแนกใบพืช เนื่องจากงานวิจัยส่วนใหญ่ได้นำขั้นตอนวิธีเชิงพันธุกรรมไปประยุกต์ใช้ในด้านต่างๆมากมายเพื่อเพิ่มประสิทธิภาพร่วมกับของโมเดลอื่นๆ แต่ด้วยประสิทธิภาพของขั้นตอนวิธีเชิงพันธุกรรมซึ่งมีความสามารถในด้านการปรับคุณลักษณะของข้อมูล ทำให้สามารถใช้คุณลักษณะต่างๆของข้อมูลในการจำแนกข้อมูลได้อย่างมีประสิทธิภาพ ดังนั้นในงานวิจัยนี้จึงได้นำขั้นตอนวิธีเชิงพันธุกรรมเพียงวิธีการเดียวมาประยุกต์ใช้ในการแก้ปัญหาการหาค่าพารามิเตอร์ที่เหมาะสมในการจำแนกใบพืช

วัตถุประสงค์ของการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยมีวัตถุประสงค์หลัก 2 ข้อ คือ

1. เพื่อนำขั้นตอนวิธีเชิงพันธุกรรมแต่ละขั้นตอนมาหาค่าพารามิเตอร์ที่เหมาะสมสำหรับใช้ในการจำแนกใบพืชโดยใช้คุณลักษณะรูปทรงและพื้นผิวของใบพืช

2. เพื่อเปรียบเทียบผลการจำแนกใบพืชระหว่างขั้นตอนวิธีเชิงพันธุกรรมกับขั้นตอนวิธีเพื่อนบ้านใกล้สุด

คำถามการวิจัย

เมื่อได้ศึกษาถึงทฤษฎี แนวคิด และงานวิจัยที่เกี่ยวข้องกับหลักการขั้นตอนวิธีเชิงพันธุกรรมแล้ว ผู้วิจัยได้พบข้อสงสัยจนนำไปสู่ข้อสังเกตและข้อคำถามในการวิจัย ดังนี้

1. ขั้นตอนวิธีเชิงพันธุกรรมในแต่ละขั้นตอนจะมีวิธีการใดและจะมีค่าพารามิเตอร์ใดที่เหมาะสมเพื่อนำไปใช้ในการจำแนกใบพืช โดยใช้คุณลักษณะรูปทรงและพื้นผิวมีความถูกต้องอยู่ในระดับใด

2. ในการจำแนกใบพืชระหว่างขั้นตอนวิธีเชิงพันธุกรรมกับขั้นตอนวิธีเพื่อนบ้านใกล้สุด วิธีการใดสามารถจำแนกใบพืชได้ดีกว่า

สมมติฐานการวิจัย

ในการวิจัยครั้งนี้ ผู้วิจัยได้กำหนดสมมติฐานการวิจัยไว้ 2 ข้อ คือ

1. การเลือกใช้ขั้นตอนวิธีเชิงพันธุกรรมที่ถูกต้องและการปรับค่าพารามิเตอร์ที่เหมาะสมโดยใช้คุณลักษณะรูปทรงและพื้นผิวทำให้สามารถจำแนกใบพืชให้ความถูกต้องอยู่ในระดับมาก

2. ขั้นตอนวิธีเชิงพันธุกรรมมีประสิทธิในการจำแนกใบพืชมากกว่าวิธีขั้นตอนเพื่อนบ้านใกล้สุดอยู่ในระดับใด

ขอบเขตของการวิจัย

1. ใช้กระบวนขั้นตอนวิธีเชิงพันธุกรรมที่เหมาะสมสำหรับปรับค่าพารามิเตอร์ที่เหมาะสมในการจำแนกชนิดใบพืช

2. ประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมโดยใช้คุณลักษณะรูปทรงและพื้นผิวในการจำแนกใบพืชสำหรับคำนวณหาค่าความถูกต้อง ค่าความครบถ้วน และค่าความแม่นยำ

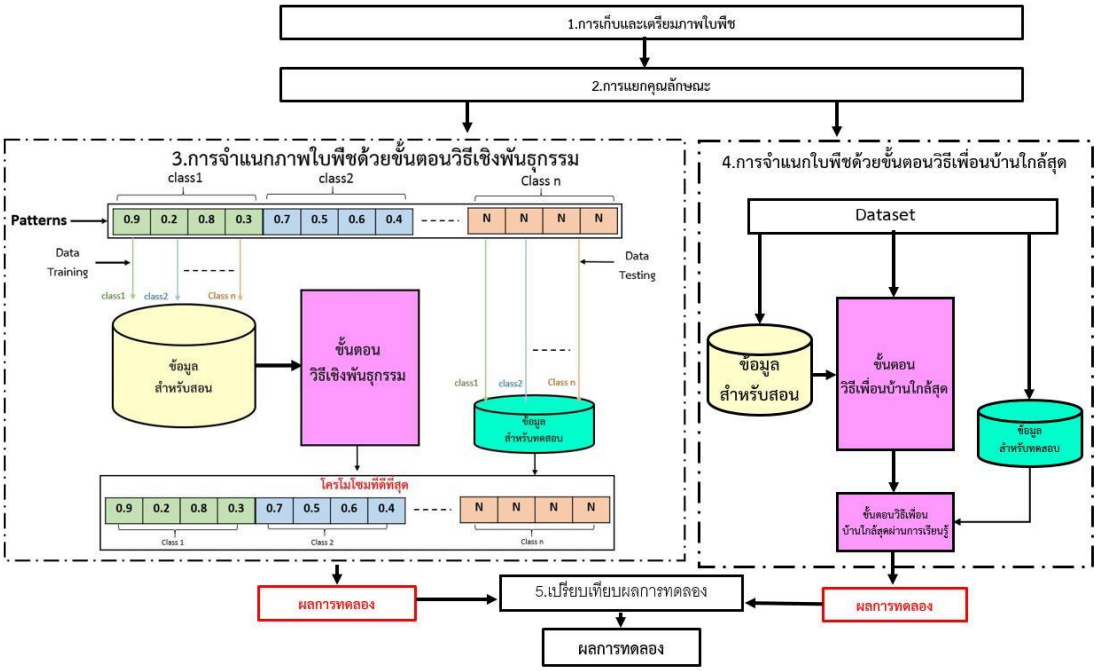
3. ประยุกต์ใช้ขั้นตอนวิธีเพื่อนบ้านใกล้สุดโดยใช้คุณลักษณะรูปทรงและพื้นผิวในการจำแนกใบพืชสำหรับคำนวณหาค่าความถูกต้อง ค่าความครบถ้วน และค่าความแม่นยำ

4. เปรียบเทียบผลการทดลองในการจำแนกใบพืชโดยใช้ Dataset ชุดเดียวกันระหว่างขั้นตอนวิธีเชิงพันธุกรรมกับขั้นตอนวิธีเพื่อนบ้านใกล้สุด

5. ข้อมูลสำหรับการทดลองในการจำแนกใบพืชเป็นกลุ่มข้อมูลตัวอย่างของใบพืช 30 สายพันธุ์ จำนวน 340 ใบ โดยใช้ข้อมูลใบพืช “Leaf” dataset (Pedro F.B.Silva et al.,2014)

กรอบแนวคิดในการวิจัย

งานวิจัยเรื่องการจำแนกใบพืชโดยการใช้คุณลักษณะรูปทรงและพื้นผิวของใบพืชด้วยขั้นตอนวิธีเชิงพันธุกรรมโดยการประยุกต์ใช้ขั้นตอนวิธีเชิงพันธุกรรมมาใช้ในการจำแนกชนิดใบพืช กรอบแนวคิดในการวิจัยประกอบด้วยขั้นตอนในการพัฒนาระบบดังภาพที่ 1.1



ภาพที่ 1.1 กรอบแนวคิดในการวิจัย (Conceptual Framework)

กรอบแนวคิดในการวิจัยเรื่องการจำแนกใบพืชโดยการใช้คุณลักษณะรูปทรงและพื้นผิวของใบพืชด้วยวิธีเชิงพันธุกรรม ที่นำเสนอใน ภาพที่ 1.1 เป็นการสรุปภาพรวมของการดำเนินงานสำหรับงานวิจัยนี้ ซึ่งสามารถอธิบายได้ดังนี้

ขั้นตอนที่ 1 การเก็บและเตรียมภาพภาพใบพืช

ขั้นตอนที่ 2 การแยกคุณลักษณะ

ขั้นตอนที่ 3 การจำแนกภาพใบพืชด้วยขั้นตอนวิธีเชิงพันธุกรรม

ขั้นตอนที่ 4 การจำแนกภาพใบพืชด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด

ขั้นตอนที่ 5 เปรียบเทียบผลการทดลองระหว่างขั้นตอนวิธีเชิงพันธุกรรมกับขั้นตอนวิธีเพื่อนบ้านใกล้สุด

ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำขั้นตอนวิธีเชิงพันธุกรรมเพียงวิธีการเดียวไปประยุกต์ใช้ในการหาค่าความเหมาะสมคุณลักษณะของข้อมูลได้

2. สามารถนำขั้นตอนวิธีเชิงพันธุกรรมไปประยุกต์ใช้ในการค้นหาคำตอบจากชุดข้อมูลจำนวนมากแทนการพิจารณาข้อมูลทีละคำตอบ

3. สามารถนำไปพัฒนาเป็น Application บนสมาร์ทโฟนสำหรับการเรียนรู้สายพันธุ์พืชนอกห้องเรียนในระบบนิเวศหรือสถานที่ต่างๆได้

นิยามศัพท์

1. Classification หมายถึง การจำแนกออกเป็นประเภท หรือการแบ่งกลุ่มข้อมูล
2. Feature Extraction หมายถึง การดึงลักษณะเด่น
3. Training หมายถึง การสอนให้ระบบเรียนรู้
4. Testing หมายถึง กระบวนการทดสอบการหาผลลัพธ์
5. Genetic Algorithm หมายถึง ขั้นตอนวิธีเชิงพันธุกรรม
6. k-Nearest Neighbor หมายถึง ขั้นตอนวิธีเพื่อนบ้านใกล้สุด