

การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางค่านพฤติกรรมทางสังคมของมนุษย์

CLASSIFICATION ALGORITHM BASED ON HUMAN SOCIAL BEHAVIOR

จิตรภรณ์ มูลวงศ์  
JITTRAPORN MOOLWONG

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาคณะหลังสูตปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. 2550

KMITL-2007-IT-M-001-120

การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์

CLASSIFICATION ALGORITHM BASED ON HUMAN SOCIAL BEHAVIOR



จิตรภรณ์ มุลวงศ์

JITRAPORN MOOLWONG

อพ.  
๗๕๔๗  
๒๕๕๐

เลขหมู่.....  
เลขทะเบียน..... 77946  
วัน,เดือน,ปี..... ๒๒ ก.พ. ๒๕๕๑

b. 11880995  
i. ....

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

บัณฑิตวิทยาลัย

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ. ๒๕๕๐

KMITL-2007-IT-M-001-120

**CLASSIFICATION ALGORITHM BASED ON HUMAN SOCIAL BEHAVIOR**

**JITTRAPORN MOOLWONG**

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT  
OF THE REQUIREMENT FOR THE DEGREE OF  
MASTER OF SCIENCE IN INFORMATION TECHNOLOGY  
SCHOOL OF GRADUATE STUDIES  
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

**2007**

**KMITL-2007-IT-M-001-120**

**COPYRIGHT 2007**

**SCHOOL OF GRADUATE STUDIES**

**KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

**บัณฑิตวิทยาลัย**  
**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**  
**ใบรับรองวิทยานิพนธ์**

-----

หัวข้อวิทยานิพนธ์      การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์  
Classification Algorithm Based on Human Social Behavior

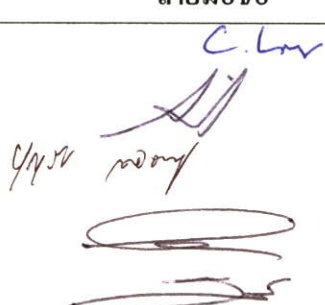
ชื่อนักศึกษา              นางสาวจิตรราภรณ์      มุลวงศ์

รหัสประจำตัว              47066245

ปริญญา                      วิทยาศาสตรมหาบัณฑิต

สาขาวิชา                    เทคโนโลยีสารสนเทศ

อาจารย์ที่ปรึกษาวิทยานิพนธ์      รศ.ดร.อาริต      ชรรมนโน

คณะกรรมการสอบวิทยานิพนธ์		ลายมือชื่อ
ศ.ดร.ชิตชนก	เหลือสินทรัพย์	
รศ.ดร.อาริต	ชรรมนโน	
รศ.ดร.บุญธีร์	เครือตราชู	
รศ.ดร.วรพจน์	กรีสุระเดช	
ผศ.ดร.ชนารัตน์	ชลิดาพงศ์	

วัน/เดือน/ปี ที่สอบ 2 ตุลาคม 2550 เวลา 10.00 น. เป็นต้นไป  
สถานที่สอบ ณ ห้อง 331 คณะเทคโนโลยีสารสนเทศ

  
บัณฑิตวิทยาลัยรับรองแล้ว  
(รศ.ดร.จารุวัตร เจริญสุข)  
คณบดีบัณฑิตวิทยาลัย

วันที่.....13.....เดือน.....ธันวาคม.....พ.ศ.....๒๕๕๐.....

หัวข้อวิทยานิพนธ์	การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์
นักศึกษา	นางสาว จิตรภรณ์ มูลวงศ์
รหัสประจำตัว	47066245
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
พ.ศ.	2550
อาจารย์ที่ปรึกษาวิทยานิพนธ์	รศ.ดร.อาริต ธรรมโน

### บทคัดย่อ

วิทยานิพนธ์ฉบับนี้นำเสนอวิธีการจัดกลุ่มข้อมูลแบบใหม่ที่ใช้แนวคิดทางด้านการลอกเลียนแบบวิวัฒนาการของสิ่งมีชีวิต โดยลอกเลียนพฤติกรรมการอยู่ร่วมกันทางสังคมของมนุษย์ที่มีการแบ่งกลุ่มทางสังคมเป็นกลุ่มวงในและกลุ่มวงนอก สมาชิกที่อยู่ในกลุ่มเดียวกันจะมีการรักษาระยะห่างทางสังคมระหว่างกันและมีทัศนคติต่อกันในแง่บวก ส่วนสมาชิกที่อยู่ต่างกลุ่มกันจะมีทัศนคติต่อกันในแง่ลบ ในงานวิจัยนี้ได้มีการนำขั้นตอนต่างๆ มาใช้ร่วมกันอย่างเป็นระบบไม่ว่าจะเป็นการสร้างคลัสเตอร์ การนำพฤติกรรมการอยู่ร่วมกันทางสังคมของมนุษย์มาประยุกต์ใช้ในการปรับค่าจุดศูนย์กลางของคลัสเตอร์ การลบคลัสเตอร์ และการสุ่มค่าระยะทางให้กับแต่ละคลัสเตอร์ เพื่อหาจุดศูนย์กลางใหม่ที่ดีกว่า โดยผลลัพธ์ที่ได้จากโมเดลคือจำนวนคลัสเตอร์และจุดศูนย์กลางของแต่ละคลัสเตอร์ ซึ่งจากการทดลองพบว่าโมเดลที่นำเสนอจะใช้จำนวนคลัสเตอร์น้อยกว่าโมเดลที่ใช้ในการจัดกลุ่มข้อมูลแบบอื่นๆ ที่อัตราความถูกต้องที่ใกล้เคียงกัน

<b>Thesise</b>	Classification Algorithm based on Human Social Behavior
<b>Student</b>	Miss Jittraporn Moolwong
<b>Student ID.</b>	47066245
<b>Degree</b>	Master of Science
<b>Program</b>	Information Science
<b>Year</b>	2007
<b>Thesis Advisor</b>	Assoc. Prof. Dr. Arit Thammano

## **ABSTRACT**

This thesis proposes a new classification algorithm that imitates an organism evolutionary concept. A model focuses on imitative social behavior of human, which has social categorize to in-group and out-group boundary. A same group member will be preserve social distance between groups , have a positive attitude for each other and negative attitude toward out-group member. In this research systematically combine a various method as follows create cluster, apply social behavior to update center, delete cluster and find better center by random value. Results of model are a cluster center and a cluster number that created less than another classification model while having equal or better accuracy.

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จได้อย่างดีด้วยคำแนะนำ คำปรึกษา กำลังใจ และการดูแลเอาใจใส่จาก รศ.ดร. อาริต ธรรมโน ซึ่งเป็นอาจารย์ผู้ควบคุมวิทยานิพนธ์ ข้าพเจ้ารู้สึกทราบบ้างในความอนุเคราะห์จากท่านอาจารย์ และขอขอบพระคุณเป็นอย่างสูง

ขอขอบคุณเพื่อนๆ พี่ๆ น้องๆ ในห้องวิจัย Computational Intelligence Laboratory และในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกคนที่ให้คำแนะนำ และคอยให้กำลังใจเสมอมา

ขอขอบคุณบัณฑิตศึกษาและบัณฑิตวิทยาลัยคณะเทคโนโลยีสารสนเทศที่ให้ความช่วยเหลือในเรื่องต่างๆ

สุดท้ายนี้ข้าพเจ้าขอกราบขอบพระคุณ บิดา มารดา และครอบครัวของข้าพเจ้าที่เป็นกำลังใจ และให้การสนับสนุนในทุกเรื่องๆ ทำให้ข้าพเจ้าสามารถทำวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดี

คุณค่าและประโยชน์อันพึงมาจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอบอบแด่ผู้มีพระคุณทุกท่าน

จิตรารกรณ์ มุลวงศ์

# สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VII
สารบัญรูป.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	1
1.3 ขอบเขตการวิจัย.....	1
1.4 ขั้นตอนของการศึกษา.....	2
1.5 โครงสร้างของวิทยานิพนธ์.....	2
บทที่ 2 ทฤษฎีพื้นฐานที่เกี่ยวข้อง.....	3
2.1 จีเนติกอัลกอริทึม.....	3
2.1.1 วิธีการเลือกโครโมโซมพ่อแม่.....	5
2.1.1.1 Pairing of top to bottom.....	5
2.1.1.2 Random pairing.....	5
2.1.1.3 Weighted random pairing.....	6
2.1.1.4 Tournament selection.....	9
2.1.2 ตัวดำเนินการทางพันธุกรรม.....	10
2.1.2.1 การไขว้เปลี่ยน.....	10
2.1.2.2 การกลายพันธุ์.....	12
2.1.3 ตัวอย่างการทำงานของ Genetic algorithm.....	13
2.2 Particle Swarm Optimization.....	16
2.2.1 วิธีการหาค่า pBest และ gBest.....	17
2.2.2 ตัวดำเนินการของ Particle Swarm Optimization.....	18
2.2.2.1 การปรับค่าอัตราเร็ว (Velocity update).....	18
2.2.2.2 การปรับค่าตำแหน่ง (Position update).....	19

## สารบัญ (ต่อ)

	หน้า
2.3.4 ตัวอย่างการทำงานของ Particle Swarm Optimization.....	19
2.3 บทความงานวิจัยที่เกี่ยวข้อง.....	22
2.3.1 A GAs based approach for mining breast cancer pattern.....	22
2.3.2 Intelligence Particle Swarm Classifier (IPS-classifier).....	24
2.3.3 Classification by evolutionary ensembles.....	25
2.3.4 New evolutionary neural network.....	28
2.3.5 Isan Dhamma Character Recognition Using Genetic Algorithm.....	31
บทที่ 3 การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์.....	34
3.1 บทนำ.....	34
3.2 โครงสร้างการทำงาน.....	35
3.3 กระบวนการเรียนรู้โมเดล.....	36
3.3.1 ขั้นตอนการสร้างคลัสเตอร์.....	36
3.3.2 ขั้นตอนการคำนวณค่าความถูกต้อง.....	37
3.3.3 ขั้นตอนการปรับค่าจุดศูนย์กลางของทุกคลัสเตอร์.....	38
3.3.4 ขั้นตอนการลบคลัสเตอร์.....	40
3.3.5 ขั้นตอนการพิจารณาค่าเปอร์เซ็นต์ความถูกต้อง.....	41
3.3.6 ขั้นตอนการสุ่มค่าระยะทางให้แก่แต่ละจุดศูนย์กลาง.....	42
3.3.7 ขั้นตอนการวนทำซ้ำ.....	43
3.4 กระบวนการทดสอบโมเดล.....	43
3.5 ตัวอย่างการทำงาน.....	44
บทที่ 4 การทดสอบการจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์.....	52
4.1 ข้อมูลที่ใช้ในการทดลอง.....	52
4.1.1 ข้อมูลมาตรฐาน.....	52
4.1.2 ข้อมูลที่ผู้วิจัยสร้างขึ้นเอง.....	54
4.2 วิธีทำการทดลอง.....	55

## สารบัญ (ต่อ)

	หน้า
4.3 การกำหนดพารามิเตอร์ของระบบ .....	57
4.4 ผลการทดลอง .....	57
4.4.1 การทดลองที่ 1 .....	57
4.4.1.1 ชุดข้อมูลมาตรฐาน .....	52
4.4.1.2 ชุดข้อมูลที่ผู้วิจัยสร้างขึ้นเอง.....	60
4.4.2 การทดลองที่ 2 .....	70
4.4.2.1 ชุดข้อมูลมาตรฐาน .....	70
4.4.2.2 ชุดข้อมูลที่ผู้วิจัยสร้างขึ้นเอง.....	72
4.5 สรุปผลการทดลอง.....	74
บทที่ 5 สรุปผลการวิจัย และข้อเสนอแนะ .....	77
5.1 สรุปผลงานวิจัย.....	77
5.2 ข้อดีของงานวิจัย .....	78
5.3 ปัญหาที่พบในงานวิจัย.....	78
5.4 แนวทางการพัฒนาในอนาคต.....	79
บรรณานุกรม.....	80
ภาคผนวก.....	82
ภาคผนวก ก. ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่.....	83
ประวัติผู้เขียน .....	90

# สารบัญตาราง

ตารางที่	หน้า
2.1 แสดงตัวอย่างการจัดเรียงโครโมโซมตามค่าความเหมาะสมจากมากไปน้อย.....	5
2.2 แสดงตัวเลขที่สุ่มได้และลำดับของโครโมโซมที่ได้จากการคำนวณ.....	6
2.3 แสดงค่าความน่าจะเป็นของแต่ละโครโมโซมโดยใช้วิธี Rank weighting .....	7
2.4 แสดงค่าความน่าจะเป็นของแต่ละโครโมโซมโดยใช้วิธี Roulette wheel.....	8
2.5 แสดงรายละเอียดการแข่งขันของโครโมโซมที่ได้จากการสุ่ม .....	9
3.1 แสดงคลัสเตอร์ที่ชนะเมื่อใช้คลัสเตอร์ 1 และ 2 ที่ผ่านการปรับค่าจุดศูนย์กลาง .....	48
3.2 แสดงจำนวนรูปแบบอินพุตที่เป็นสมาชิกของแต่ละคลัสเตอร์.....	49
3.3 แสดงการสุ่มระยะทางให้กับจุดศูนย์กลางของคลัสเตอร์ที่ 1.....	50
3.4 แสดงการสุ่มระยะทางให้กับจุดศูนย์กลางของคลัสเตอร์ที่ 2.....	50
4.1 แสดงการแบ่งข้อมูลของชุดข้อมูล Iris .....	55
4.2 แสดงการแบ่งข้อมูลของชุดข้อมูล Wine.....	55
4.3 แสดงการแบ่งข้อมูลของชุดข้อมูล Ionosphere.....	56
4.4 แสดงการแบ่งข้อมูลของชุดข้อมูล Haberman.....	56
4.5 แสดงการแบ่งข้อมูลของชุดข้อมูล Heart-Statlog.....	56
4.6 แสดงการแบ่งข้อมูลของชุดข้อมูล Flower1, Flower2, Sawtooth.....	56
4.7 แสดงการแบ่งข้อมูลของชุดข้อมูล Fan.....	57
4.8 แสดงผลการทดลองของข้อมูล Iris .....	58
4.9 แสดงผลการทดลองของข้อมูล Wine.....	58
4.10 แสดงผลการทดลองของข้อมูล Ionosphere.....	59
4.11 แสดงผลการทดลองของข้อมูล Haberman.....	59
4.12 แสดงผลการทดลองของข้อมูล Heart-Statlog.....	60
4.13 แสดงผลการทดลองของข้อมูล Flower1 .....	60
4.14 แสดงผลการทดลองของข้อมูล Flower2 .....	63
4.15 แสดงผลการทดลองของข้อมูล Sawtooth .....	65
4.16 แสดงผลการทดลองของข้อมูล Fan.....	68
4.17 แสดงผลการทดลองของข้อมูล Iris เมื่อเปรียบเทียบกับ RBF .....	70
4.18 แสดงผลการทดลองของข้อมูล Wine เมื่อเปรียบเทียบกับ RBF .....	71

## สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.19 แสดงผลการทดลองของข้อมูล Ionosphere เมื่อเปรียบเทียบกับ RBF.....	71
4.20 แสดงผลการทดลองของข้อมูล Haberman เมื่อเปรียบเทียบกับ RBF .....	72
4.21 แสดงผลการทดลองของข้อมูล Heart-Statlog เมื่อเปรียบเทียบกับ RBF .....	72
4.22 แสดงผลการทดลองของข้อมูล Flower1 เมื่อเปรียบเทียบกับ RBF .....	73
4.23 แสดงผลการทดลองของข้อมูล Flower2 เมื่อเปรียบเทียบกับ RBF .....	73
4.24 แสดงผลการทดลองของข้อมูล Sawtooth เมื่อเปรียบเทียบกับ RBF.....	73
4.25 แสดงผลการทดลองของข้อมูล Fan เมื่อเปรียบเทียบกับ RBF.....	74

# สารบัญรูป

รูปที่	หน้า
2.1 แสดงโครงสร้างการทำงานของจินตคณิตอรรถวิธีม .....	4
2.2 แสดงวงล้อรูเลตที่ที่ได้โดยใช้วิธี Rank weighting .....	7
2.3 แสดงวงล้อรูเลตที่ที่ได้โดยใช้วิธี Roulette wheel .....	9
2.4 แสดงการไขว้เปลี่ยนด้วยวิธี Single-point crossover .....	10
2.5 แสดงการไขว้เปลี่ยนด้วยวิธี 2-point crossover .....	10
2.6 แสดงการไขว้เปลี่ยนด้วยวิธี Multi-point Crossover .....	11
2.7 แสดงการไขว้เปลี่ยนด้วยวิธี Variable to variable Crossover .....	11
2.8 แสดงการไขว้เปลี่ยนด้วยวิธี Uniform Crossover .....	12
2.9 แสดงวงล้อรูเลตที่ของตัวอย่าง โดยใช้วิธี Rank weighting .....	14
2.10 แสดงโครงสร้างการทำงานของ Particle Swarm Optimization .....	17
2.11 แสดงรูปแบบของโคร โมโซมที่ใช้ในการค้นหาภูการทำนาย .....	23
2.12 แสดงสถาปัตยกรรมของ Classification by evolutionary ensembles .....	26
2.13 แสดงตัวอย่าง โคร โมโซมที่สร้างขึ้น .....	32
2.14 แสดงโคร โมโซมก่อนและหลังการครอสโอเวอร์ .....	32
2.15 แสดงยีนตัวอักษรก่อนและหลังการมิวเตต .....	33
3.1 แสดงแสดงรูปแบบอินพุตที่ใช้ในการเรียนรู้โมเดล โดย $\diamond$ แทนรูปแบบอินพุตในคลาสที่ 1 และ $\triangle$ แทนรูปแบบอินพุตในคลาสที่ 2 .....	44
3.2 แสดงจุดศูนย์กลางเริ่มต้นของคลัสเตอร์ โดย * แทนจุดศูนย์กลางของคลัสเตอร์ที่ 1 และ + แทนจุดศูนย์กลางของคลัสเตอร์ที่ 2 .....	45
3.3 แสดงจุดศูนย์กลางของคลัสเตอร์ 1 และ 2 หลังการปรับค่าจุดศูนย์กลาง .....	48
3.4 แสดงจุดศูนย์กลางของคลัสเตอร์ 2 หลังการสุ่มระยะทาง .....	51
4.1 แสดงลักษณะของดอก Iris ทั้ง 3 ประเภท .....	52
4.2 แสดงลักษณะของชุดข้อมูลที่ทำวิจัยสร้างขึ้นเอง .....	54
4.3 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Flower1 .....	61
4.4 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบโมเดลของชุดข้อมูล Flower1 .....	61
4.5 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Flower1 .....	62

## สารบัญรูป (ต่อ)

รูปที่	หน้า
4.6 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Flower1 .....	62
4.7 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Flower2.....	63
4.8 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบ โมเดลของชุดข้อมูล Flower2.....	64
4.9 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Flower2 .....	64
4.10 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Flower2 .....	65
4.11 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Sawtooth.....	66
4.12 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบ โมเดลของชุดข้อมูล Sawtooth .....	66
4.13 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Sawtooth.....	67
4.14 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Sawtooth.....	67
4.15 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Fan .....	68
4.16 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบ โมเดลของชุดข้อมูล Fan .....	69
4.17 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Fan.....	69
4.18 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน ของชุดข้อมูล Fan.....	70

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาและความสำคัญของปัญหา

การจัดกลุ่มข้อมูล (Classification) เป็นขั้นตอนในการค้นหาเซตของโมเดลที่ใช้อธิบายหรือจำแนกความแตกต่างของข้อมูลแต่ละประเภทโดยพิจารณาจากคุณสมบัติที่มีอยู่ ตั้งแต่อดีตจนถึงปัจจุบันมีผู้คิดค้น โมเดลที่ใช้ในการจัดกลุ่มข้อมูลมากมายหลายโมเดลด้วยแนวคิดที่ต่างกันไป ซึ่งแนวความคิดที่ค่อนข้างได้รับความนิยมคือ โมเดลที่เลียนแบบวิวัฒนาการของสิ่งมีชีวิต (Evolutionary Algorithms) เช่น โมเดลที่เลียนแบบการถ่ายทอดทางพันธุกรรมของมนุษย์หรือที่เรียกว่า จีเนติกอัลกอริทึม โมเดลที่เลียนแบบระบบทางสังคมของสิ่งมีชีวิตที่มีการรวมกันเป็นกลุ่ม และมีการติดต่อสื่อสารระหว่างกัน ในสภาพแวดล้อมเดียวกันหรือที่เรียกว่า Swarm Intelligence โดยโมเดลที่ได้รับความนิยมคือ Ant Colony (ACO) ซึ่งลอกเลียนแบบลักษณะการเดินของมด และ Particle Swarm Optimization (PSO) ที่ลอกเลียนแบบการหาอาหารของฝูงนก วิทยานิพนธ์ฉบับนี้พยายามสร้างโมเดลใหม่ที่มีการทำงานในลักษณะลอกเลียนแบบพฤติกรรมทางสังคมของมนุษย์ (human social behavior) ที่มีการสร้างวงเขตของกลุ่ม (group boundary) ทำให้เกิดกลุ่มวงใน (in-group) และ กลุ่มวงนอก (out-group) ซึ่งสมาชิกที่อยู่ในกลุ่มเดียวกันจะมีทัศนคติต่อกันในแง่บวก แต่จะมีทัศนคติในแง่ลบกับสมาชิกที่อยู่นอกกลุ่ม และสมาชิกของกลุ่มวงในจะมีการรักษาระยะห่างทางสังคม (Social distance) กับกลุ่มวงนอก โดยการสร้างความรู้สึกอคติกับกลุ่มอื่น

### 1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1. เพื่อพัฒนาโมเดลใหม่ที่ใช้ในการจัดกลุ่มข้อมูลที่มีแนวคิดในการลอกเลียนแบบวิวัฒนาการของสิ่งมีชีวิต โดยเป็นการลอกเลียนแบบพฤติกรรมการอยู่ร่วมกันทางสังคมของมนุษย์
2. เพื่อพัฒนาโมเดลที่มีการสร้างจำนวนคลัสเตอร์ที่เหมาะสมกับข้อมูลแต่ละแบบ

### 1.3 ขอบเขตการวิจัย

1. เป็นการประมวลผลแบบ off-line
2. โมเดลนี้ใช้ได้กับข้อมูลที่เป็นตัวเลขเท่านั้น
3. ใช้อัตราความถูกต้องในการจัดกลุ่มข้อมูลเป็นตัววัดประสิทธิภาพของโมเดล

## 1.4 ขั้นตอนของการศึกษา

1. ศึกษาโมเดลพื้นฐานที่ใช้ในการจัดกลุ่มข้อมูล
2. ศึกษาโมเดลที่มีแนวคิดในการลอกเลียนแบบวิวัฒนาการของสิ่งมีชีวิต
3. ศึกษาทฤษฎีที่เกี่ยวกับพฤติกรรมการอยู่ร่วมกันทางสังคมของมนุษย์
4. ทดลองนำแนวคิดการอยู่ร่วมกันทางสังคมของมนุษย์มาประยุกต์ใช้ร่วมกับการจัดกลุ่มข้อมูล
5. วิเคราะห์และสรุปผลเพื่อเลือกหาแนวทางที่เหมาะสมจะนำมาสร้างโมเดลใหม่
6. สร้างโมเดลที่ใช้ในการจัดกลุ่มข้อมูลแบบใหม่
7. นำโมเดลใหม่ไปทดลองกับข้อมูลลักษณะต่างๆ
8. แก้ไขปรับปรุงโมเดลให้มีความเหมาะสมมากขึ้น
9. สรุปผลการทดลอง

## 1.5 โครงสร้างของวิทยานิพนธ์

วิทยานิพนธ์ฉบับนี้ได้แบ่งเนื้อหาออกเป็น 5 บทด้วยกันคือ

บทที่ 1 กล่าวถึงความเป็นมาของงานวิจัย ความมุ่งหมาย วัตถุประสงค์ ขอบเขตของการวิจัย และขั้นตอนการศึกษา

บทที่ 2 กล่าวถึงทฤษฎีพื้นฐานที่ใช้ในการวิจัย

บทที่ 3 กล่าวถึงทฤษฎีการจัดกลุ่มข้อมูล โดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์

บทที่ 4 กล่าวถึงการทดลองและผลการทดลอง

บทที่ 5 เป็นบทสรุปผลการวิจัยและข้อเสนอแนะ

## บทที่ 2

# ทฤษฎีพื้นฐานที่เกี่ยวข้อง

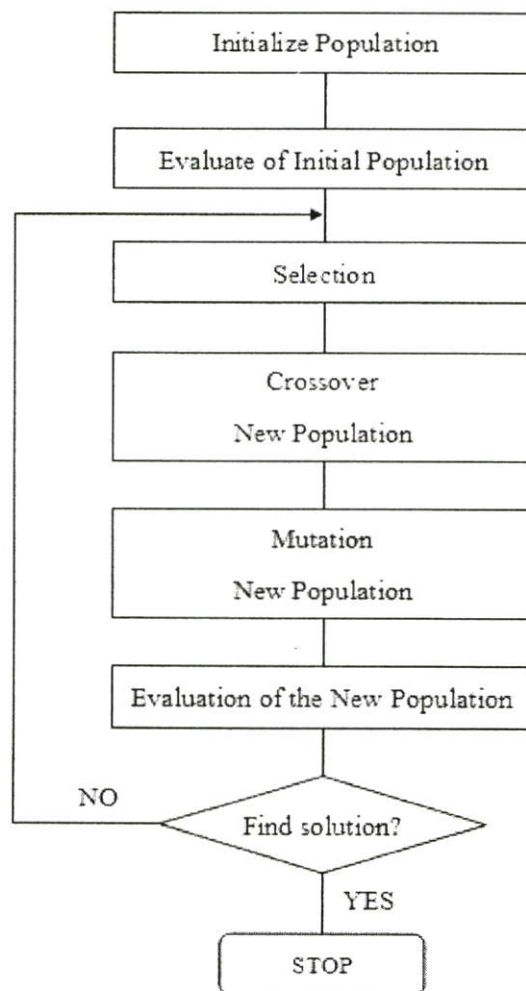
การจัดกลุ่มข้อมูล (Classification) เป็นขั้นตอนในการค้นหาเขตของโมเดลที่อธิบายหรือจำแนกความแตกต่างของข้อมูลแต่ละประเภทโดยพิจารณาจากคุณสมบัติที่มีอยู่ [1] การจัดกลุ่มข้อมูลเป็นการเรียนรู้แบบมีเป้าหมาย (Supervised Learning) คือ มีการนำข้อมูลในอดีตที่มีการกำหนดรูปแบบข้อมูลเข้า กลุ่มข้อมูล และจำนวนกลุ่มข้อมูลไว้ล่วงหน้าแล้วมาใช้ในการสร้างโมเดล ข้อมูลในอดีตที่ใช้สร้างโมเดลจะถูกแบ่งเป็น 2 ส่วน คือ ข้อมูลทดลอง (Training Data) คือ ข้อมูลที่ใช้ในการสร้างโมเดล และข้อมูลทดสอบ (Testing Data) คือ ข้อมูลที่ใช้ในการทดสอบความถูกต้องของโมเดลที่ได้ การจัดกลุ่มข้อมูลจะแบ่งขั้นตอนการทำงานออกเป็น 3 ขั้นตอนดังนี้ (1) Model Construction เป็นขั้นตอนการสร้างโมเดลโดยเรียนรู้จากข้อมูลทดลองว่าข้อมูลโดยอยู่ในกลุ่มเดียวกันบ้างโดยผลลัพธ์ที่ได้จากขั้นตอนนี้คือ โมเดลที่ใช้ในการแบ่งกลุ่มข้อมูล (Classifier Model) โดยโมเดลที่ได้จากการจัดกลุ่มอาจอยู่ในรูปแบบของ Classification (IF-THEN) Rule, Decision tree, Mathematical formulae หรือ Neural Network [1] (2) Model Evaluation เป็นขั้นตอนในการนำข้อมูลทดสอบมาทดสอบความถูกต้องของโมเดลที่ได้ โดยในขั้นตอนนี้เราสามารถปรับเปลี่ยนโมเดลได้จนกว่าความถูกต้องจะอยู่ในระดับที่น่าพอใจ (3) Model Usage เป็นขั้นตอนในการนำข้อมูลใหม่ที่ยังไม่ถูกแบ่งกลุ่ม (Unseen Data) มาผ่านโมเดลเพื่อจัดหากกลุ่มข้อมูลที่เหมาะสม

### 2.1 จีเนติกอัลกอริทึม (Genetic Algorithm)

ตามกฎทางธรรมชาติสิ่งมีชีวิตทุกชีวิตจะมีชุดของโครโมโซมซึ่งใช้อธิบายถึงที่มาของสิ่งมีชีวิต โดยโครโมโซมคือโครงสร้างที่บรรจุสารพันธุกรรมหรือยีนในเซลล์ของสิ่งมีชีวิต ยีนจะทำหน้าที่ควบคุม และถ่ายทอดลักษณะทางพันธุกรรมของสิ่งมีชีวิตจากพ่อแม่ไปสู่ลูก เช่น ลักษณะของเส้นผม ลักษณะดวงตา เพศ และผิว เป็นต้น

จีเนติกอัลกอริทึมได้นำแนวคิดเกี่ยวกับพันธุศาสตร์ทางชีววิทยาและวิวัฒนาการทางธรรมชาติของสิ่งมีชีวิตมาประยุกต์ใช้ในการหาคำตอบให้ปัญหาโดยนำคำตอบมาสร้างเป็นโครโมโซม [3] ซึ่งแต่ละโครโมโซมจะประกอบด้วยยีนหลายๆ ยีนที่แสดงถึงคุณลักษณะ (Variable) ของโครโมโซม โดยโครโมโซมที่ดีจะถูกเลือกมาเป็นต้นแบบในการสร้างคำตอบใหม่ๆ ซึ่งเหมือนกับระบบธรรมชาติที่ทำการคัดกรองสิ่งมีชีวิต มีการแพร่พันธุ์ และมีการถ่ายทอดทางพันธุศาสตร์ให้ประชากรในรุ่นถัดไป [3]

จินตริกัลกอริทึมมีลักษณะการทำงานเหมือนกับ Directed search algorithm หรือ Heuristic algorithm คือ ทำหน้าที่ค้นหาเขตของคำตอบที่เป็นไปได้ โดยจินตริกัลกอริทึมจะใช้ฟังก์ชันความเหมาะสม (Fitness Function) เพื่อเลือกโครโมโซมที่มีความเหมาะสมมาเป็นโครโมโซมพ่อแม่ (parents) โครโมโซมพ่อแม่จะถูกทำการไขว้เปลี่ยนคุณลักษณะ (crossover) เพื่อสร้างเป็นโครโมโซมลูก (offspring) และ โครโมโซมลูกจะถูกกลายพันธุ์ (mutation) เพื่อให้ได้โครโมโซมที่มีความหลากหลาย หลังจากนั้นจะหาค่าความเหมาะสมของโครโมโซมพ่อแม่และโครโมโซมลูกทั้งหมด และทำการเลือกโครโมโซมชุดที่ดีที่สุดมาเป็นประชากรในรุ่นถัดไป ส่วนโครโมโซมที่ไม่ดีจะถูกตัดทิ้ง



รูปที่ 2.1 แสดงโครงสร้างการทำงานของจินตริกัลกอริทึม

โครโมโซมลูกจะถูกสร้างโดยการจับคู่กันของโครโมโซมพ่อแม่ โดยโครโมโซมพ่อแม่ 1 คู่จะทำให้เกิดโครโมโซมลูก 2 โครโมโซม จากประชากรทั้งหมดเราจะนำโครโมโซมมาเรียงตามค่า

Cost function จากน้อยไปมาก โดยโครโมโซมที่มีค่า cost น้อยจะเป็นโครโมโซมที่ดี ส่วนโครโมโซมที่มีค่า cost มากจะเป็นโครโมโซมที่ไม่ดี หรือเรียงตามค่าความเหมาะสมจากมากไปน้อย โดยโครโมโซมที่มีค่าความเหมาะสมมากจะเป็นโครโมโซมที่ดี ส่วนโครโมโซมที่มีค่าความเหมาะสมน้อยจะเป็นโครโมโซมที่ไม่ดี

ตัวอย่าง ถ้ากำหนดให้จำนวนประชากร= 6, จำนวนบิต= 14, จำนวนยีน= 7 และ ฟังก์ชันความเหมาะสม คือ  $\text{Max: } f(x, y) = x \sin(4x) + 1.1y \sin(2y)$  ดังนั้น ถ้าโครโมโซม คือ 00001101010110 เราจะได้ค่า  $x = 0000110 = 6$ ,  $y = 1010110 = 86$  และ ค่าความเหมาะสม = 61.6057

เราจะทำการเรียงโครโมโซมตามค่าความเหมาะสมจากมากไปน้อยได้ดังตารางที่ 2.1

ตารางที่ 2.1 แสดงตัวอย่างการจัดเรียงโครโมโซมตามค่าความเหมาะสมจากมากไปน้อย

ลำดับโครโมโซม	โครโมโซม	x	y	ฟังก์ชันความเหมาะสม	ค่าความเหมาะสม
1	00001101010110	6	86	$6\sin(4*6) + 1.1(86)\sin(2*86)$	61.6057
2	00100110011101	19	29	$19\sin(4*19) + 1.1(29)\sin(2*29)$	42.4287
3	01010110000100	43	4	$43\sin(4*43) + 1.1(4)\sin(2*4)$	34.8255
4	01011000011010	44	26	$44\sin(4*44) + 1.1(26)\sin(2*26)$	31.3306
5	00001000001010	4	10	$4\sin(4*4) + 1.1(10)\sin(2*10)$	8.8908
6	00101000111111	20	63	$20\sin(4*20) + 1.1(63)\sin(2*63)$	2.9906

### 2.1.1 วิธีการเลือกโครโมโซมพ่อแม่

**2.1.1.1 Pairing of top to bottom** ทำการจัดเรียงโครโมโซมจากโครโมโซมที่ดีที่สุดไปหาโครโมโซมที่แย่ที่สุด และทำการจับคู่โครโมโซมพ่อแม่จากบนลงล่าง [4] เช่นถ้าดูจากตารางที่ 2.1 คู่ของโครโมโซมที่ได้โดยใช้วิธีนี้เป็นดังนี้

คู่ที่ 1 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 2

คู่ที่ 2 : โครโมโซมที่ 3 คู่กับ โครโมโซมที่ 4

คู่ที่ 3 : โครโมโซมที่ 5 คู่กับ โครโมโซมที่ 6

**2.1.1.2 Random pairing** ทำการจัดคู่โครโมโซมพ่อแม่โดยการสุ่มตัวเลขตามจำนวนประชากรที่กำหนดไว้เพื่อหาลำดับของโครโมโซม [4]

$$parent = roundup (N * random) \quad (2.1)$$

โดยที่  $N$  คือ จำนวนประชากร  
 $random$  คือ ตัวเลขจำนวนจริงในช่วง 0-1 ที่ได้จากการสุ่มค่า  
 $roundup$  คือ ฟังก์ชันที่ใช้ในการปัดเลขจำนวนจริงขึ้นเป็นเลขจำนวนเต็ม

ตารางที่ 2.2 แสดงตัวเลขที่สุ่มได้และลำดับของโครโมโซมที่ได้จากการคำนวณ

เลขที่สุ่มได้	6 * เลขที่สุ่มได้	ลำดับของโครโมโซมที่ได้
0.1535	0.921	1
0.6781	4.0686	5
0.0872	0.5232	1
0.1936	1.1616	2
0.7021	4.2126	5
0.3933	2.3598	3

จากตารางที่ 2.2 จะได้คู่ของโครโมโซมดังนี้

- คู่ที่ 1 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 5
- คู่ที่ 2 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 2
- คู่ที่ 3 : โครโมโซมที่ 5 คู่กับ โครโมโซมที่ 3

**2.1.1.3 Weighted random pairing** การหาค่าความน่าจะเป็นให้กับโครโมโซมโดยค่าความน่าจะเป็นจะขึ้นอยู่กับค่าความเหมาะสมของโครโมโซม โดยโครโมโซมที่มีค่าความเหมาะสมมากจะมีความน่าจะเป็นที่จะถูกเลือกสูง ส่วนโครโมโซมที่มีค่าความเหมาะสมน้อยจะมีความน่าจะเป็นที่จะถูกเลือกต่ำ บ่อยครั้งที่เราจะเรียกวิธีนี้ว่า Roulette wheel weighting ซึ่งจะนำเสนอ 2 เทคนิค ดังนี้

**1. Rank weighting** การหาค่าความน่าจะเป็นจากลำดับของโครโมโซม เมื่อชุดของโครโมโซมต้นแบบถูกเรียงจากโครโมโซมที่ดีที่สุดไปหาโครโมโซมที่แย่ที่สุด ดังนั้นลำดับของโครโมโซมจึงสามารถบอกได้ว่าโครโมโซมนั้นๆ มีความเหมาะสมที่จะเป็นคำตอบมากน้อยแค่ไหน [4]

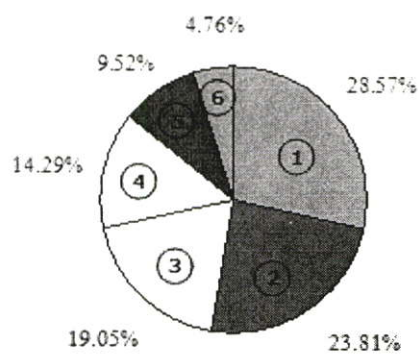
$$P_n = \frac{N - n + 1}{\sum_{n=1}^N n} \quad (2.2)$$

โดยที่  $P_n$  คือ ความน่าจะเป็นของโครโมโซมลำดับที่  $n$   
 $N$  คือ จำนวนประชากร  
 $n$  คือ ลำดับของโครโมโซม

จากตารางที่ 2.1 ถ้า  $N=6$  แล้วจะได้ว่า  $P_n = \frac{6-n+1}{1+2+3+4+5+6} = \frac{7-n}{21}$

ตารางที่ 2.3 แสดงค่าความน่าจะเป็นของแต่ละโครโมโซมโดยใช้วิธี Rank weighting

ลำดับโครโมโซม	โครโมโซม	$P_n$	$\sum_{i=1}^n P_i$
1	00001101010110	$\frac{7-1}{21} = 0.2857$	0.2857
2	00100110011101	$\frac{7-2}{21} = 0.2381$	0.5238
3	01010110000100	$\frac{7-3}{21} = 0.1905$	0.7143
4	001011000011010	$\frac{7-4}{21} = 0.1429$	0.8572
5	00001000001010	$\frac{7-5}{21} = 0.0952$	0.9524
6	00101000111111	$\frac{7-6}{21} = 0.0476$	1.0000



รูปที่ 2.2 แสดงวงล้อรูปที่ได้อาศัยวิธี Rank weighting

เมื่อได้ค่าความน่าจะเป็นของแต่ละโครโมโซมแล้ว ทำการสุ่มตัวเลขทศนิยมที่มีค่าอยู่ในช่วง 0-1 ตามจำนวนประชากรที่กำหนดไว้ แล้วดูว่าเลขที่สุ่มได้แต่ละครั้งตกอยู่ในช่วงของโครโมโซมไหน ก็เลือกโครโมโซมนั้นมาเป็นโครโมโซมพ่อแม่

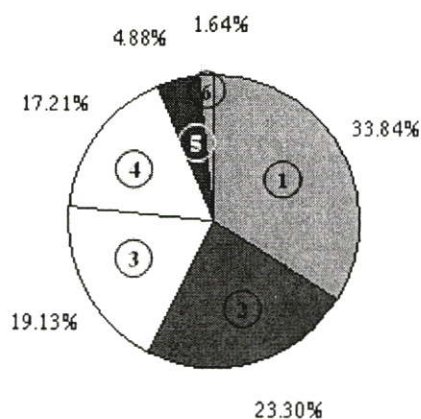
**2. Roulette wheel** การหาค่าความน่าจะเป็นโดยการนำค่าความเหมาะสมที่คำนวณได้มาหารออกด้วยผลรวมของค่าความเหมาะสมทั้งหมด [2]

$$P_n = \frac{f_n}{\sum_{p=1}^N f_p} \quad (2.3)$$

โดยที่  $P_n$  คือ ค่าความน่าจะเป็นของโครโมโซมลำดับที่  $n$   
 $f_n$  คือ ค่าความเหมาะสมของโครโมโซมลำดับที่  $n$   
 $N$  คือ จำนวนประชากร  
 $n$  คือ ลำดับของโครโมโซม

ตารางที่ 2.4 แสดงค่าความน่าจะเป็นของแต่ละโครโมโซมโดยใช้วิธี Roulette wheel

ลำดับโครโมโซม	โครโมโซม	ค่าความเหมาะสม	$P_n = \frac{f_n}{\sum_{p=1}^N f_p}$
1	00001101010110	61.6057	$\frac{61.6057}{182.0719} = 0.34$
2	00100110011101	42.4287	$\frac{42.4287}{182.0719} = 0.23$
3	01010110000100	34.8255	$\frac{34.8255}{182.0719} = 0.19$
4	001011000011010	31.3306	$\frac{31.3306}{182.0719} = 0.17$
5	00001000001010	8.8908	$\frac{8.8908}{182.0719} = 0.05$
6	00101000111111	2.9906	$\frac{2.9906}{182.0719} = 0.02$
ผลรวมของค่าความเหมาะสม		182.0719	



รูปที่ 2.3 แสดงวงล้อรูเลตต์ที่ได้โดยใช้วิธี Roulette wheel

เมื่อได้ค่าความน่าจะเป็นของแต่ละโครโมโซมแล้ว ทำการสุ่มตัวเลขทศนิยมที่มีค่าอยู่ในช่วง 0-1 ตามจำนวนประชากรที่กำหนดไว้ แล้วดูว่าเลขที่สุ่มได้แต่ละครั้งตกอยู่ในช่วงของโครโมโซมใด ก็เลือกโครโมโซมนั้นมาเป็นโครโมโซมพ่อแม่

**2.1.1.4 Tournament selection** นำคู่ของโครโมโซมมาจัดการแข่งขัน โดยในแต่ละการแข่งขันจะมีการสุ่มลำดับของโครโมโซมขึ้นมา 2 ตัว และทำการเปรียบเทียบค่าความเหมาะสมของโครโมโซม โดยโครโมโซมที่มีค่าความเหมาะสมมากที่สุดจะเป็นผู้ชนะ [4]

ตารางที่ 2.5 แสดงรายละเอียดการแข่งขันของโครโมโซมที่ได้จากการสุ่ม

ลำดับการแข่งขัน	ลำดับของโครโมโซมที่สุ่มได้	โครโมโซมที่ชนะ
1	2 VS 1	1
2	5 VS 5	5
3	6 VS 3	3
4	4 VS 5	4
5	1 VS 1	1
6	4 VS 5	4

จากโครโมโซมที่ชนะในแต่ละการแข่งขันให้ทำการจับคู่โครโมโซมจากบนลงล่าง ดังนั้น จะได้คู่ของโครโมโซมพ่อแม่ ดังนี้

คู่ที่ 1 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 5

คู่ที่ 2 : โครโมโซมที่ 3 คู่กับ โครโมโซมที่ 4

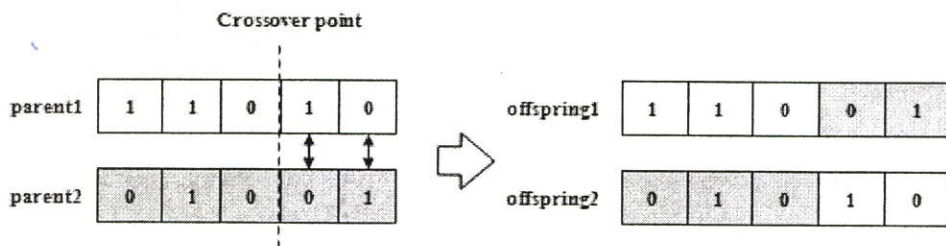
คู่ที่ 3 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 4

## 2.1.2 ตัวดำเนินการทางพันธุกรรม (Genetic Operation)

### 2.1.2.1 การไขว้เปลี่ยน (Crossover)

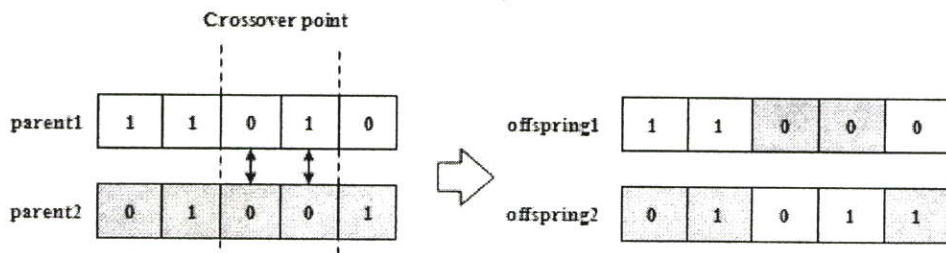
หลังจากที่เราได้คู่ของโครโมโซมพ่อแม่แล้วเราจะทำการไขว้เปลี่ยนคุณลักษณะของโครโมโซมพ่อแม่เพื่อสร้างเป็นโครโมโซมลูก โดยจะมีการกำหนดค่าคงที่เพื่อนำมาเป็นค่าความน่าจะเป็นที่จะไขว้เปลี่ยน ( $P_C$ ) ซึ่งโดยทั่วไปเราจะกำหนด  $P_C$  ให้มีค่าสูงๆ หรือประมาณ 0.9 [8] หลังจากนั้นทำการสุ่มตัวเลขในช่วง 0-1 ซึ่งถ้าค่าที่สุ่มได้มีค่าน้อยกว่าค่าความน่าจะเป็นที่ตั้งไว้โครโมโซมนั้นจะถูกไขว้เปลี่ยน ให้กำหนดตำแหน่งที่จะทำการไขว้เปลี่ยนโดยการสุ่มเลขจำนวนเต็มที่มีค่าอยู่ในช่วง 1 ถึง  $m-1$  ( $m$  คือ จำนวนยีนในโครโมโซม) จำนวน  $n$  ตัว ตามแต่เทคนิคที่เลือกใช้ [3] ในที่นี้จะนำเสนอเทคนิคการไขว้เปลี่ยน 5 เทคนิค ดังนี้

1. **Single-point Crossover** ทำการสุ่มตำแหน่งยีนที่จะทำการไขว้เปลี่ยน 1 ตำแหน่ง และสลับค่ายีนพ่อแม่ที่อยู่หลังตำแหน่งที่สุ่มได้ทุกยีน



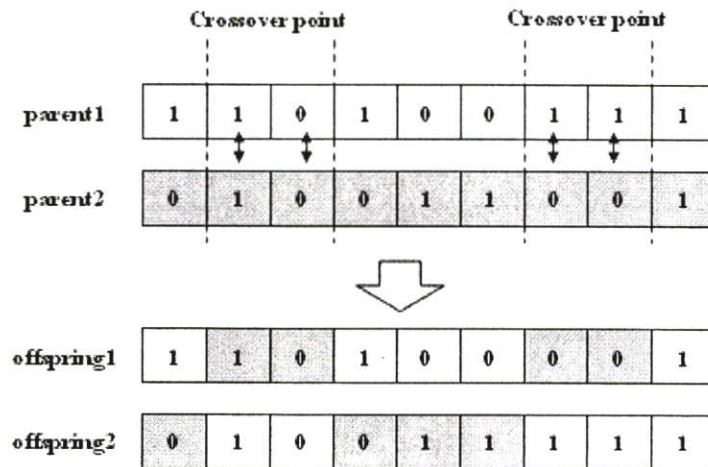
รูปที่ 2.4 แสดงการไขว้เปลี่ยนด้วยวิธี Single-point Crossover

2. **2-point Crossover** สุ่มตำแหน่งยีนที่จะทำการไขว้เปลี่ยน 2 ตำแหน่ง และทำการสลับค่ายีนพ่อแม่ที่อยู่ภายในช่วงตำแหน่งที่สุ่มได้ หรือสลับค่ายีนพ่อแม่ที่อยู่ภายนอกช่วงตำแหน่งที่สุ่มได้ [8]



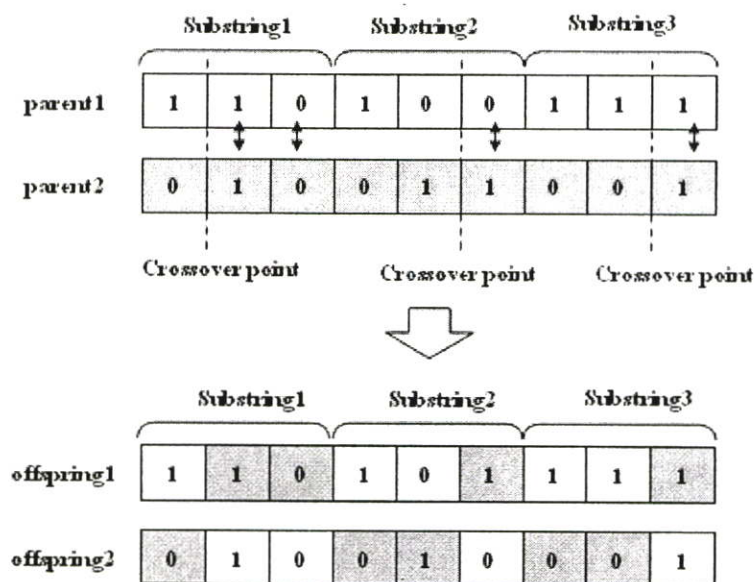
รูปที่ 2.5 แสดงการไขว้เปลี่ยนด้วยวิธี 2-point Crossover

3. **Multi-point Crossover** มีลักษณะการทำคล้ายกับ Single-point Crossover และ 2-point Crossover โดยทำการสุ่มตำแหน่งที่จะไขว้เปลี่ยน ( $n_c \geq 3$ ) โดยจะทำการไขว้เปลี่ยนขึ้นที่อยู่ระหว่างจุดไขว้เปลี่ยนแต่ละจุด



รูปที่ 2.6 แสดงการไขว้เปลี่ยนด้วยวิธี Multi-point Crossover

4. **Variable to variable Crossover** การทำจีเนติกอัลกอริทึมแบบมาตรฐาน กลุ่มของข้อมูลหรือกลุ่มของตัวเลข (string) ซึ่งแสดงถึงตัวแปรแต่ละตัวจะถูกแบ่งออกเป็นส่วนย่อย (substring) ในรูปของเลขฐานสอง โดยวิธีนี้จะเป็นการทำ Single-point Crossover ให้กับแต่ละ substring หรือให้กับแต่ละตัวแปรนั่นเอง [8]

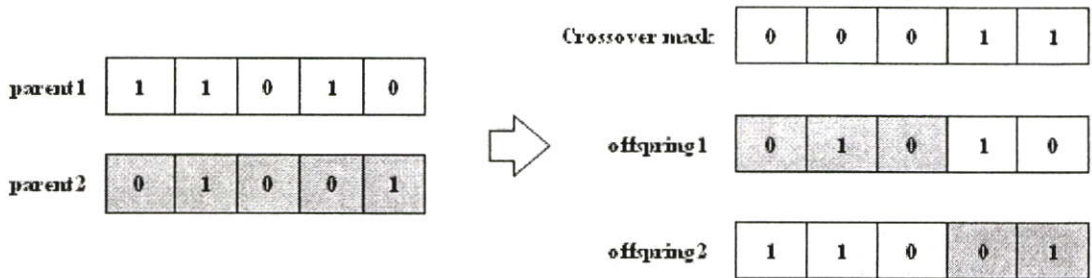


รูปที่ 2.7 แสดงการไขว้เปลี่ยนด้วยวิธี Variable to variable Crossover

5. **Uniform Crossover** เริ่มต้นด้วยการสร้าง Crossover mask โดยการสุ่มตัวเลข (0,1) ตามจำนวนยีนในโครโมโซม โดยตัวเลขที่อยู่ในแต่ละยีนจะเป็นตัวบอกว่าโครโมโซมลูกจะต้องทำการคัดลอกยีนจากโครโมโซมพ่อแม่ตัวไหน โดยโครโมโซมลูกตัวที่ 1 และ 2 จะมีขั้นตอนการคัดลอกยีนจากโครโมโซมพ่อแม่ดังนี้

– โครโมโซมลูกตัวที่ 1 : จะดูว่าถ้ายีนใน Crossover mask เป็นเลข 1 โครโมโซมลูกจะทำการคัดลอกยีนจากโครโมโซมพ่อแม่ตัวที่ 1 แต่ถ้ายีนใน Crossover mask เป็นเลข 0 โครโมโซมลูกจะทำการคัดลอกยีนจากโครโมโซมพ่อแม่ตัวที่ 2

– โครโมโซมลูกตัวที่ 2 : ทำการสร้าง Crossover mask ขึ้นมาใหม่อีก 1 ชุด แล้วทำการบวกรวมกับโครโมโซมลูกตัวที่ 1 [8] หรือใช้ Crossover mask ร่วมกับโครโมโซมลูกตัวที่ 1 แต่ใช้กฎในลักษณะที่ตรงข้ามกับโครโมโซมลูกตัวที่ 1 ก็ได้ [7]



รูปที่ 2.8 แสดงการไขว้เปลี่ยนด้วยวิธี Uniform Crossover

### 2.1.2.2 การกลายพันธุ์ (Mutation)

เป็นการเพิ่มลักษณะพิเศษให้กับโครโมโซมลูก ซึ่งอาจจะเป็นลักษณะที่ไม่มีในลักษณะพื้นฐานของประชากรเริ่มต้นที่เรากำหนด โดยทั่วไปในแต่ละรอบจะมียีนที่ถูกกลายพันธุ์เพียง 1-5% การกลายพันธุ์ส่วนใหญ่มักจะทำให้ค่าความเหมาะสมของโครโมโซมลดลงแต่จะเป็นการเพิ่มความหลากหลาย และทำให้ประชากรมีความแข็งแรงมากขึ้น เพราะการกลายพันธุ์เป็นการหันเหเส้นทางมาบรรจบกันของคำตอบช่วยให้คำตอบที่เหมาะสมมาบรรจบกันช้าลง [4]

ขั้นตอนการกลายพันธุ์เริ่มต้นด้วยการกำหนดค่าคงที่ซึ่งเป็นค่าความน่าจะเป็นที่แต่ละยีนจะถูกกลายพันธุ์ ( $P_M$ ) ซึ่งโดยทั่วไปเราจะกำหนด  $P_M$  ให้มีค่าต่ำๆ หรือประมาณ 0.01 (มีการกลายพันธุ์ 1%) [8] หลังจากนั้นแต่ละยีนจะทำการสุ่มค่าตัวเลขระหว่าง 0-1 โดยถ้าค่าที่สุ่มได้มีค่าน้อยกว่าค่าความน่าจะเป็นที่ตั้งไว้ยีนนั้นๆ จะถูกกลายพันธุ์ [3] ให้ทำการสุ่มตัวเลขที่อยู่ในช่วงที่กำหนดไว้มาแทนที่ค่าที่อยู่ในตำแหน่งนั้นๆ หรือทำการสุ่มตัวเลขเพื่อนำมาใช้คำนวณร่วมกับค่าเดิมในยีนนั้นๆ เพื่อให้ได้เป็นค่าใหม่

### 2.1.3 ตัวอย่างการทำงานของ Genetic algorithm

ทำการกำหนดค่าเริ่มต้นให้ตัวแปรต่างๆ ดังนี้

- จำนวนประชากร = 4
- จำนวนคุณลักษณะ = 2 ( x และ y)
- ฟังก์ชันความเหมาะสม

$$\max : f(x, y) = x \sin(4x) + 1.1y \sin(2y) \quad \text{โดยที่ } 1 \leq x \leq 10 \text{ และ } 1 \leq y \leq 10$$

- ค่าความน่าจะเป็นในการไขว้เปลี่ยน = 0.9
- ค่าความน่าจะเป็นในการกลายพันธุ์ = 0.05

1. ทำการสุ่มเลขจำนวนจริงให้กับ x และ y เพื่อสร้างเป็นโครโมโซมเริ่มต้นตามจำนวนประชากรที่กำหนดไว้ โดยจำนวนยีนเท่ากับจำนวนคุณลักษณะ

$$\text{โครโมโซมที่ 1} : x = 2.7708, y = 8.4617$$

$$\text{โครโมโซมที่ 2} : x = 7.6151, y = 9.1032$$

$$\text{โครโมโซมที่ 3} : x = 9.1382, y = 5.2693$$

$$\text{โครโมโซมที่ 4} : x = 9.0465, y = 8.3097$$

2. หาค่าความเหมาะสมของแต่ละโครโมโซมจากฟังก์ชันความเหมาะสมที่กำหนดไว้ และทำการจัดเรียงโครโมโซมตามค่าความเหมาะสมที่ได้จากมากไปน้อย

$$\text{โครโมโซมที่ 1} : x = 2.7708, y = 8.4617 \quad \text{ค่าความเหมาะสม} = -11.4865$$

$$\text{โครโมโซมที่ 2} : x = 7.6151, y = 9.1032 \quad \text{ค่าความเหมาะสม} = -12.224$$

$$\text{โครโมโซมที่ 3} : x = 9.1382, y = 5.2693 \quad \text{ค่าความเหมาะสม} = -13.5287$$

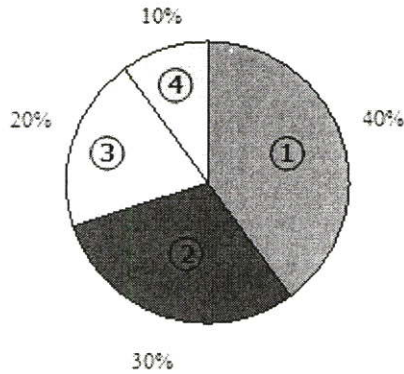
$$\text{โครโมโซมที่ 4} : x = 9.0465, y = 8.3097 \quad \text{ค่าความเหมาะสม} = -16.2561$$

3. เลือกโครโมโซมพ่อแม่จำนวน 2 โครโมโซม ซึ่งในที่นี้จะใช้วิธี Rank weighting เพื่อหาค่าความน่าจะเป็นของแต่ละโครโมโซม ดังนั้นจากสมการที่ 2.2 จะได้สมการดังนี้

$$P_n = \frac{4 - n + 1}{1 + 2 + 3 + 4} = \frac{5 - n}{10}$$

$$\text{โครโมโซมที่ 1} : x = 2.7708, y = 8.4617 \quad \text{มีค่าความน่าจะเป็น} = \frac{5 - 1}{10} = 0.4$$

- โครโมโซมที่ 2 :  $x = 7.6151, y = 9.1032$     มีค่าความน่าจะเป็น  $= \frac{5-2}{10} = 0.3$
- โครโมโซมที่ 3 :  $x = 9.1382, y = 5.2693$     มีค่าความน่าจะเป็น  $= \frac{5-3}{10} = 0.2$
- โครโมโซมที่ 4 :  $x = 9.0465, y = 8.3097$     มีค่าความน่าจะเป็น  $= \frac{5-4}{10} = 0.1$



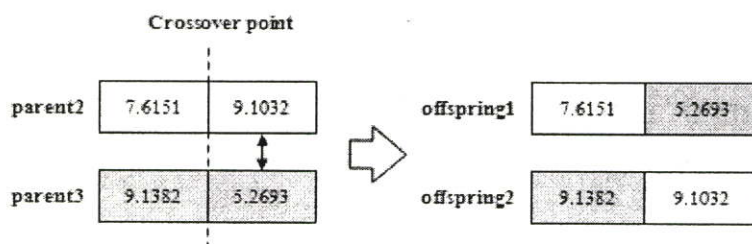
รูปที่ 2.9 แสดงวงล้อรูปเลขที่ของตัวอย่างโดยใช้วิธี Rank weighting

เมื่อได้ช่วงความน่าจะเป็นของแต่ละโครโมโซมแล้ว เราจะทำการสร้างโครโมโซมลูก 4 โครโมโซม ดังนั้น จะทำการจับคู่โครโมโซมพ่อแม่จำนวน 2 คู่ โดยทำการสุ่มตัวเลขจำนวนจริงที่มีค่าอยู่ในช่วง 0-1 และดูว่าตัวเลขที่สุ่มได้ตกอยู่ในช่วงของโครโมโซมใด

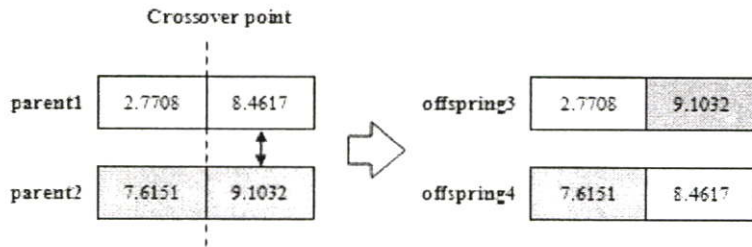
- คู่ที่ 1 : (0.4679, 0.7344) → โครโมโซม 2 คู่กับ โครโมโซม 3
- คู่ที่ 2 : (0.2872, 0.6985) → โครโมโซม 1 คู่กับ โครโมโซม 2

4. นำโครโมโซมพ่อแม่แต่ละคู่มาไขว้เปลี่ยนคุณลักษณะซึ่งกันและกัน เพื่อสร้างเป็นโครโมโซมลูก

- โครโมโซมชุดที่ 1 → สุ่มค่าความน่าจะเป็นในการไขว้เปลี่ยน = 0.61  
→ สุ่มตำแหน่งการไขว้เปลี่ยน = 2



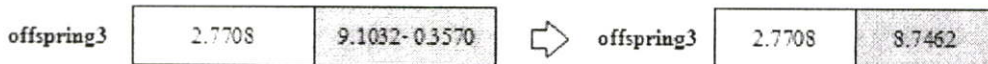
- โครโมโซมชุดที่ 2 → สุ่มค่าความน่าจะเป็นในการไขว้เปลี่ยน = 0.67  
 → สุ่มตำแหน่งการไขว้เปลี่ยน = 2



5. นำโครโมโซมลูกที่ได้มาทำการกลายพันธุ์เพื่อปรับเปลี่ยนคุณลักษณะของยีน โดยทำการสุ่มค่าความน่าจะเป็นในการกลายพันธุ์ให้แต่ละยีน ดังนี้

- โครโมโซมลูกที่ 1 : 0.84 , 0.09
- โครโมโซมลูกที่ 2 : 0.82 , 0.50
- โครโมโซมลูกที่ 3 : 0.11 , 0.04
- โครโมโซมลูกที่ 4 : 0.99 , 0.10

ดังนั้น โครโมโซมลูกที่มีค่าความน่าจะเป็นน้อยกว่าค่าความน่าจะเป็นที่ตั้งไว้ ( $P_M = 0.05$ ) มี 1 จุด คือ โครโมโซมที่ 3 ยีนที่ 2 ดังนั้น ทำการสุ่มตัวเลข 1 ตัว เพื่อนำมาคำนวณร่วมกับค่าเดิมของโครโมโซมที่ 3 ยีนที่ 2 ซึ่งถ้าเลขที่สุ่มได้ คือ -0.3570 จะได้โครโมโซมลูกดังนี้



6. นำโครโมโซมเดิมและโครโมโซมลูกที่ได้มาหาค่าความเหมาะสม เพื่อเลือกชุดของโครโมโซมที่เหมาะสมที่สุดตามจำนวนประชากรที่ตั้งไว้

โครโมโซมเดิม

- โครโมโซมที่ 1 :  $x = 2.7708, y = 8.4617$  มีค่าความเหมาะสม = -11.4865
- โครโมโซมที่ 2 :  $x = 7.6151, y = 9.1032$  มีค่าความเหมาะสม = -12.224
- โครโมโซมที่ 3 :  $x = 9.1382, y = 5.2693$  มีค่าความเหมาะสม = -13.5287
- โครโมโซมที่ 4 :  $x = 9.0465, y = 8.3097$  มีค่าความเหมาะสม = -16.2561

โครโมโซมลูก

- โครโมโซมที่ 5 :  $x = 7.6151, y = 5.2693$  มีค่าความเหมาะสม = -11.4201
- โครโมโซมที่ 6 :  $x = 9.1382, y = 9.1032$  มีค่าความเหมาะสม = -14.3325

โครโมโซมที่ 7 :  $x = 2.7708, y = 8.7462$  มีค่าความเหมาะสม = -12.1623

โครโมโซมที่ 8 :  $x = 7.6151, y = 8.4617$  มีค่าความเหมาะสม = -14.945

จัดเรียงค่าความเหมาะสมจากมากไปน้อยเพื่อเลือกโครโมโซมชุดที่ดีที่สุดมาเป็นประชากรในรุ่นถัดไป ซึ่งได้ประชากรในรุ่นถัดไป ดังนี้

โครโมโซมที่ 5 :  $x = 7.6151, y = 5.2693$  มีค่าความเหมาะสม = -11.4201

โครโมโซมที่ 1 :  $x = 2.7708, y = 8.4617$  มีค่าความเหมาะสม = -11.4865

โครโมโซมที่ 7 :  $x = 2.7708, y = 8.7462$  มีค่าความเหมาะสม = -12.1623

โครโมโซมที่ 2 :  $x = 7.6151, y = 9.1032$  มีค่าความเหมาะสม = -12.2240

7. ทำซ้ำกระบวนการที่ 3 ถึง 7 จนกระทั่งครบตามจำนวนรุ่นที่กำหนดไว้หรือค่าของโครโมโซมที่ดีที่สุดไม่เปลี่ยนแปลง

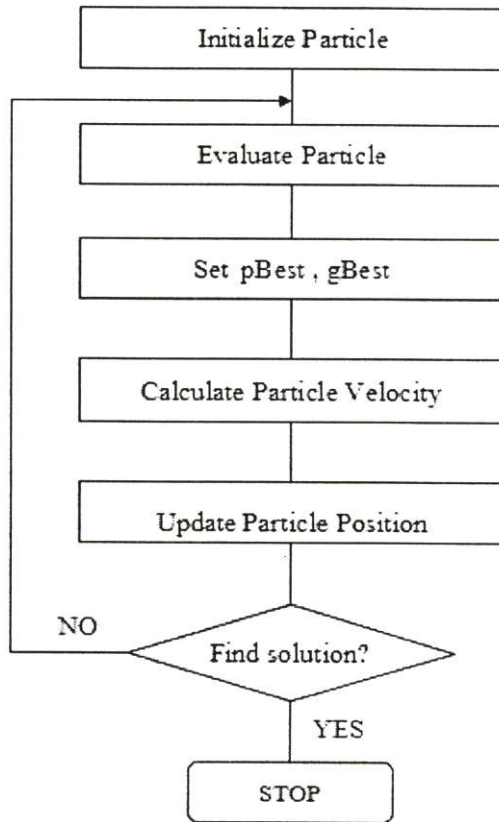
## 2.2 Particle Swarm Optimization (PSO)

เทคนิคทางด้าน การคำนวณที่มีแรงคลอจมาจากระบบทางชีววิทยามีหลายเทคนิค เช่น Artificial Neural Network ซึ่งเป็น โมเดลที่เลียนแบบการทำงานของสมองมนุษย์ หรือจินตริก อัลกอริทึมซึ่งเป็น โมเดลที่เลียนแบบวิวัฒนาการของมนุษย์ และเทคนิคทางชีววิทยาแบบอื่นเช่น เทคนิคที่เลียนแบบระบบทางสังคมของสิ่งมีชีวิตที่มีการรวมกันเป็นกลุ่ม และมีการติดต่อสื่อสารระหว่างกัน ในสภาพแวดล้อมเดียวกัน อาจเรียกว่า swarm intelligence โดยเทคนิคที่ได้รับความนิยม คือ Ant Colony (ACO) และ Particle Swarm Optimization (PSO) [6]

PSO ถูกพัฒนาโดย Dr. Russell C. Eberhart และ Dr. James Kennedy ในปี 1995 โดยแนวคิดของโมเดลมีการลอกเลียนแบบระบบพื้นฐานในการรวมกลุ่มกัน โดยมุ่งจำลองการรวมกลุ่มกันของฝูงนกหรือฝูงปลา ซึ่งมีการปรับตัวโดยใช้หลักจิตวิทยาในการรวมกลุ่มกันทางสังคม [5] PSO เลียนแบบพฤติกรรมของฝูงนก ซึ่งกลุ่มของนกจะมีการล่าหาอาหารในพื้นที่ ซึ่งแต่ละพื้นที่จะมีแหล่งอาหารอยู่เพียงจุดเดียว ดังนั้น กลุ่มนกจึงใช้กลยุทธ์ในการหาอาหารที่ได้ผลดีวิธีหนึ่ง คือ การบินตามนกตัวที่อยู่ใกล้กับอาหารมากที่สุด [6]

PSO ทำการเรียนรู้จากพฤติกรรมข้างต้นและนำมาใช้เป็นวิธีในการหาคำตอบที่เหมาะสมให้กับปัญหา ใน PSO นกแต่ละตัวหมายถึงแต่ละคำตอบที่เป็นไปได้ซึ่งเราจะเรียกว่าพาร์ทิเคิล (particle) ทุกๆ พาร์ทิเคิลจะมีค่าความเหมาะสมซึ่งถูกคำนวณโดยใช้ฟังก์ชันความเหมาะสม (Fitness Function) และมีอัตราเร็วซึ่งใช้นำทางการบินของพาร์ทิเคิล แต่ละพาร์ทิเคิลจะบินไปในทิศทางเดียวกับพาร์ทิเคิลที่ดีที่สุด ซึ่ง PSO ไม่ได้รับประกันว่าจะได้คำตอบที่ดีที่สุด (global optimum) [6]

PSO จะมีการกำหนดกลุ่มของพาร์ทิเคิลเริ่มต้น แต่ละพาร์ทิเคิลจะมีการหาค่าตำแหน่งที่ดีที่สุดของตัวเอง (pBest) และ ตำแหน่งที่ดีที่สุดของกลุ่ม (gBest) ซึ่งจะมีผลต่อการปรับเปลี่ยนตำแหน่งของพาร์ทิเคิล หลังจากนั้นจะทำการปรับค่าอัตราเร็วและตำแหน่งของแต่ละพาร์ทิเคิล ซึ่งโครงสร้างการทำงานของ PSO แสดงได้ดังรูปที่ 2.10 [6]



รูปที่ 2.10 แสดงโครงสร้างการทำงานของ Particle Swarm Optimization

### 2.1.4 วิธีการหาค่า pBest และ gBest

2.1.4.1 ตำแหน่งที่ดีที่สุดของตัวเอง (personal best: pBest) ทำการหาค่า pBest ของแต่ละพาร์ทิเคิลโดยการเปรียบเทียบค่าความเหมาะสมระหว่างตำแหน่งที่ดีที่สุดที่เคยไปถึง (pBest เดิม) กับตำแหน่งที่คำนวณได้จากรอบปัจจุบันโดย

- ถ้าค่าความเหมาะสมของตำแหน่งใหม่ > ค่าความเหมาะสมของ pBest เดิมแล้ว ให้ pBest มีค่าเท่ากับตำแหน่งใหม่
- ถ้าค่าความเหมาะสมของตำแหน่งใหม่  $\leq$  ค่าความเหมาะสมของ pBest เดิมแล้ว ให้ pBest มีค่าเท่ากับ pBest เดิม

2.1.4.2 ตำแหน่งที่ดีที่สุดของกลุ่ม (global best: gBest) เมื่อได้ค่า pBest ของแต่ละพาร์ทิเคิลแล้ว ให้ทำการเลือก pBest ที่ให้ค่าความเหมาะสมมากที่สุดมาเป็น gBest

## 2.1.5 ตัวดำเนินการของ Particle Swarm Optimization [5]

### 2.1.5.1 การปรับค่าอัตราเร็ว (Velocity update)

1. การปรับค่าอัตราความเร็วแบบมาตรฐาน (Standard PSO) แต่ละพาร์ทิเคิลจะมีการปรับอัตราเร็ว (Velocity Vector) ที่ใช้ในการบอกทิศทางให้กับแต่ละพาร์ทิเคิล การปรับค่าอัตราเร็วจะดูจากอัตราความเร็วปัจจุบัน ระยะทางจากตำแหน่งปัจจุบันไปหาตำแหน่งที่ดีที่สุดของตัวเอง และระยะทางจากตำแหน่งปัจจุบันไปหาตำแหน่งที่ดีที่สุดของกลุ่ม โดยค่าอัตราความเร็วใหม่ที่ได้อาจจะถูกนำไปใช้ในการคำนวณหาตำแหน่งถัดไปของพาร์ทิเคิล การปรับอัตราเร็วแสดงได้ดังสมการที่ 2.4

$$v_{id}(t+1) = v_{id}(t) + c_1 rand_1(pbest_{id}(t) - x_{id}(t)) + c_2 rand_2(gbest_d(t) - x_{id}(t)) \quad (2.4)$$

โดย	$i$	คือ	ลำดับของพาร์ทิเคิล
	$t$	คือ	รอบการทำงานปัจจุบัน
	$d$	คือ	ขนาดของอินพุต โดย $d = 1, \dots, N_d$
	$N_d$	คือ	จำนวนขนาดของอินพุต
	$v_{id}$	คือ	อัตราความเร็ว
	$x_{id}$	คือ	ตำแหน่งปัจจุบัน
	$pbest_{id}$	คือ	ตำแหน่งที่ดีที่สุดของตัวเอง
	$gbest_d$	คือ	ตำแหน่งที่ดีที่สุดของกลุ่ม
	$c_1, c_2$	คือ	ค่าคงที่ที่แสดงถึงปัจจัยทางสังคมหรือปัจจัยในการเรียนรู้
	$rand_1, rand_2$	คือ	ค่าที่ได้จากการสุ่มเลขจำนวนจริงที่อยู่ในช่วง 0 ถึง 1

2. การปรับค่าอัตราความเร็วโดยใช้แรงผลักดัน (PSO with Inertia) ในปี 1998 Shi และ Eberhart ได้เสนอ PSO ที่มีการใช้ค่าน้ำหนักของแรงผลักดัน (inertia weight) โดยค่าน้ำหนักของแรงผลักดันมีไว้เพื่อควบคุมผลกระทบที่ค่าอัตราเร็วเก่าจะมีต่อค่าอัตราเร็วปัจจุบัน โดยถ้าค่าน้ำหนักของแรงผลักดันมากจะสามารถค้นหา global ได้ง่ายขึ้น แต่ถ้าค่าน้ำหนักของแรงผลักดันน้อยจะทำให้ค้นหา local ได้ง่ายขึ้น

ค่าน้ำหนักของแรงผลักดันที่เหมาะสมควรจะสมดุลกันระหว่างความสามารถในการค้นหา global และ local ซึ่งโดยทั่วไปแล้วเราจะทำการกำหนดค่าเริ่มต้นของค่าน้ำหนักของแรงผลักดันให้มากกว่า และค่อยๆ ลดค่าลงทีละน้อยตามระยะเวลาการทำงานที่ผ่านมา [5] เพื่อให้การหาค่าตอบมีความละเอียดมากขึ้น การปรับอัตราเร็วโดยใช้ค่าน้ำหนักของแรงผลักดันแสดงได้ดังสมการที่ 2.5

$$v_{id}(t+1) = w(t+1)v_{id}(t) + c_1rand_1(pbest_{id}(t) - x_{id}(t)) + c_2rand_2(gbest_d(t) - x_{id}(t)) \quad (2.5)$$

ในแต่ละรอบการทำงานจะทำการลดค่า  $w(t+1)$  จากค่าน้ำหนักเริ่มต้น ( $w_{start}$ ) ไปหาค่าน้ำหนักสุดท้าย ( $w_{end}$ ) ซึ่งค่าของ  $w_{start}$  และ  $w_{end}$  ที่นิยมใช้คือ 0.9 และ 0.4 ตามลำดับ

$$w(t+1) = w_{start} - (w_{start} - w_{end}) \left( \frac{t}{T_{max}} \right) \quad (2.6)$$

โดยที่	$w(t+1)$	คือ	ค่าน้ำหนักของแรงผลักดัน
	$T_{max}$	คือ	รอบการทำงานที่สูงที่สุด
	$t$	คือ	รอบการทำงานปัจจุบัน
	$w_{start}$	คือ	ค่าเริ่มต้นของค่าน้ำหนักของแรงผลักดัน
	$w_{end}$	คือ	ค่าสิ้นสุดของค่าน้ำหนักของแรงผลักดัน

**2.1.5.2 การปรับค่าตำแหน่ง (Position update)** แต่ละพาร์ติเคิลจะทำการปรับค่าตำแหน่งของตัวเอง โดยนำค่าอัตราความเร็วใหม่ที่ได้อมาเป็นตัวบอกทิศทางการเคลื่อนที่ของพาร์ติเคิลว่าควรจะเคลื่อนที่ไปในทิศทางใด

$$x_{id}(t+1) = x_i(t) + v_{id}(t+1) \quad (2.7)$$

**2.1.6 ตัวอย่างการทำงานของ Particle Swarm Optimization โดยใช้การปรับค่าอัตราความเร็วแบบมาตรฐาน**

ทำการกำหนดค่าเริ่มต้นให้ตัวแปรต่างๆ ดังนี้

- จำนวนพาร์ติเคิล = 4
- จำนวนคุณลักษณะ = 2 (x และ y)
- ฟังก์ชันความเหมาะสม

$$\max : f(x, y) = x \sin(4x) + 1.1y \sin(2y) \quad \text{โดยที่ } 1 \leq x \leq 10 \text{ และ } 1 \leq y \leq 10$$

- $v_{id} = 0$
- $c_1, c_2 = 2$
- $rand_1, rand_2 = 0.4642$  และ  $0.2749$  ตามลำดับ

1. กำหนดชุดของพาร์ทิเคิลเริ่มต้นตามจำนวนพาร์ทิเคิลที่กำหนดไว้ โดยตัวเลขที่สุ่มจะต้องอยู่ในช่วงของค่าตัวแปรที่กำหนดไว้

พาร์ทิเคิลที่ 1 :  $x = 2.7708, y = 8.4617$

พาร์ทิเคิลที่ 2 :  $x = 7.6151, y = 9.1032$

พาร์ทิเคิลที่ 3 :  $x = 9.1382, y = 5.2693$

พาร์ทิเคิลที่ 4 :  $x = 9.0465, y = 8.3097$

2. คำนวณหาค่าความเหมาะสมของแต่ละพาร์ทิเคิล

พาร์ทิเคิลที่ 1 :  $x = 2.7708, y = 8.4617$  ค่าความเหมาะสม = -11.4865

พาร์ทิเคิลที่ 2 :  $x = 7.6151, y = 9.1032$  ค่าความเหมาะสม = -12.224

พาร์ทิเคิลที่ 3 :  $x = 9.1382, y = 5.2693$  ค่าความเหมาะสม = -13.5287

พาร์ทิเคิลที่ 4 :  $x = 9.0465, y = 8.3097$  ค่าความเหมาะสม = -16.2561

3. ทำการตรวจสอบค่าความเหมาะสมของตำแหน่งใหม่ที่ได้กับค่าความเหมาะสมของ pBest เดิม ซึ่งถ้าค่าความเหมาะสมที่ได้จากตำแหน่งใหม่มีค่ามากกว่าค่าความเหมาะสมที่ได้จาก pBest เดิม ก็ให้ pBest เปลี่ยนเป็นเก็บค่าตำแหน่งใหม่

เนื่องจากการทำงานรอบแรก ดังนั้นจะถือว่าตำแหน่งปัจจุบันของแต่ละพาร์ทิเคิลเป็น pBest ของตัวเอง

พาร์ทิเคิลที่ 1 :  $x = 2.7708, y = 8.4617$  ค่าความเหมาะสม = -11.4865  
pBest คือ  $x = 2.7708, y = 8.4617$

พาร์ทิเคิลที่ 2 :  $x = 7.6151, y = 9.1032$  ค่าความเหมาะสม = -12.224  
pBest คือ  $x = 7.6151, y = 9.1032$

พาร์ทิเคิลที่ 3 :  $x = 9.1382, y = 5.2693$  ค่าความเหมาะสม = -13.5287  
pBest คือ  $x = 9.1382, y = 5.2693$

พาร์ทิเคิลที่ 4 :  $x = 9.0465, y = 8.3097$  ค่าความเหมาะสม = -16.2561  
pBest คือ  $x = 9.0465, y = 8.3097$

เลือก pBest ที่มีค่าความเหมาะสมสูงที่สุดมาเป็น gBest ดังนั้นจะเห็นว่าพาร์ทิเคิลที่มีค่าความเหมาะสมสูงที่สุดคือพาร์ทิเคิลที่ 1 ดังนั้น ค่า gBest ที่ได้คือ  $x=2.7708$  และ  $y=8.4617$

4. เมื่อได้ค่า pBest และ gBest แล้วให้ทำการคำนวณค่าอัตราเร็วใหม่ของแต่ละพาร์ติเคิล โดยใช้สมการที่ 2.4 ดังนี้

ข. พาร์ติเคิลที่ 1

$$\begin{aligned}v_{11} &= 0 + [2 * 0.4642(2.7708 - 2.7708)] + [2 * 0.2749(2.7708 - 2.7708)] \\ &= 0\end{aligned}$$

$$\begin{aligned}v_{12} &= 0 + [2 * 0.4642(8.4617 - 8.4617)] + [2 * 0.2749(8.4617 - 8.4617)] \\ &= 0\end{aligned}$$

ค. พาร์ติเคิลที่ 2

$$\begin{aligned}v_{21} &= 0 + [2 * 0.4642(7.6151 - 7.6151)] + [2 * 0.2749(2.7708 - 7.6151)] \\ &= -2.6634\end{aligned}$$

$$\begin{aligned}v_{22} &= 0 + [2 * 0.4642(9.1032 - 9.1032)] + [2 * 0.2749(8.4617 - 9.1032)] \\ &= -0.3527\end{aligned}$$

ง. พาร์ติเคิลที่ 3

$$\begin{aligned}v_{31} &= 0 + [2 * 0.4642(9.1382 - 9.1382)] + [2 * 0.2749(2.7708 - 9.1382)] \\ &= -3.5008\end{aligned}$$

$$\begin{aligned}v_{32} &= 0 + [2 * 0.4642(5.2693 - 5.2693)] + [2 * 0.2749(8.4617 - 5.2693)] \\ &= 1.7552\end{aligned}$$

จ. พาร์ติเคิลที่ 4

$$\begin{aligned}v_{41} &= 0 + [2 * 0.4642(9.0465 - 9.0465)] + [2 * 0.2749(2.7708 - 9.0465)] \\ &= -3.4504\end{aligned}$$

$$\begin{aligned}v_{42} &= 0 + [2 * 0.4642(8.3097 - 8.3097)] + [2 * 0.2749(8.4617 - 8.3097)] \\ &= 0.0836\end{aligned}$$

5. นำค่าอัตราเร็วใหม่ที่ได้มาใช้ในการปรับค่าตำแหน่งของแต่ละพาร์ติเคิล โดยใช้สมการที่ 2.7

$$\text{พาร์ติเคิลที่ 1 : } x_{11} = 2.7708 + 0 = 2.7708 \quad ; \quad x_{12} = 8.4617 + 0 = 8.4617$$

$$\text{พาร์ติเคิลที่ 2 : } x_{21} = 7.6151 - 2.6634 = 4.9517 \quad ; \quad x_{22} = 9.1032 - 0.3527 = 8.7505$$

$$\text{พาร์ติเคิลที่ 3 : } x_{31} = 9.1382 - 3.5008 = 5.6374 \quad ; \quad x_{32} = 5.2693 + 1.7552 = 7.0244$$

$$\text{พาร์ติเคิลที่ 4 : } x_{41} = 9.0465 - 3.4504 = 5.5961 \quad ; \quad x_{42} = 8.3097 + 0.0836 = 8.3932$$

6. ทำซ้ำข้อ 2 ถึง 6 จนกระทั่งครบรอบการทำงานที่กำหนดไว้ หรือค่าอัตราเร็วของทุกพาร์ติเคิลมีค่าเข้าใกล้ศูนย์คือแต่ละพาร์ติเคิลมีการเปลี่ยนแปลงค่าที่น้อยมาก

## ตัวอย่างการหาค่า pBest และ gBest ใหม่ ในการทำงานรอบที่ 2

พาร์ทิเคิลที่ 1	ตำแหน่ง pBest เดิม คือ $x = 2.7708, y = 8.4617$	ค่าความเหมาะสม คือ -11.4865
พาร์ทิเคิลที่ 2	ตำแหน่ง pBest เดิม คือ $x = 7.6151, y = 9.1032$	ค่าความเหมาะสม คือ -12.2240
พาร์ทิเคิลที่ 3	ตำแหน่ง pBest เดิม คือ $x = 9.1382, y = 5.2693$	ค่าความเหมาะสม คือ -13.5287
พาร์ทิเคิลที่ 4	ตำแหน่ง pBest เดิม คือ $x = 9.0465, y = 8.3097$	ค่าความเหมาะสม คือ -16.2561

หลังจากผ่านขั้นตอนการทำงานในรอบที่ 2 ได้ตำแหน่งใหม่และค่าความเหมาะสมใหม่ของพาร์ทิเคิลดังนี้

พาร์ทิเคิลที่ 1	ตำแหน่งใหม่ คือ $x = 2.7708, y = 8.4617$	ค่าความเหมาะสม คือ -11.4865
พาร์ทิเคิลที่ 2	ตำแหน่งใหม่ คือ $x = 4.9517, y = 8.7505$	ค่าความเหมาะสม คือ -5.3402
พาร์ทิเคิลที่ 3	ตำแหน่งใหม่ คือ $x = 5.6374, y = 7.0245$	ค่าความเหมาะสม คือ 4.7097
พาร์ทิเคิลที่ 4	ตำแหน่งใหม่ คือ $x = 5.5961, y = 8.3933$	ค่าความเหมาะสม คือ -10.2814

ทำการหาค่า pBest ใหม่ โดยเปรียบเทียบค่า pBest ที่ได้จากรอบก่อนหน้ากับ pBest ที่ได้จากรอบปัจจุบัน ดังนั้นได้ pBest ใหม่ดังนี้

พาร์ทิเคิลที่ 1	ตำแหน่ง pBest ใหม่ คือ $x = 2.7708, y = 8.4617$	ค่าความเหมาะสม คือ -11.4865
พาร์ทิเคิลที่ 2	ตำแหน่ง pBest ใหม่ คือ $x = 4.9517, y = 8.7505$	ค่าความเหมาะสม คือ -5.3402
พาร์ทิเคิลที่ 3	ตำแหน่ง pBest ใหม่ คือ $x = 5.6374, y = 7.0245$	ค่าความเหมาะสม คือ 4.7097
พาร์ทิเคิลที่ 4	ตำแหน่ง pBest ใหม่ คือ $x = 5.5961, y = 8.3933$	ค่าความเหมาะสม คือ -10.2814

จาก pBest ใหม่ที่ได้ ทำการเลือก pBest ที่มีค่าความเหมาะสมสูงที่สุดมาเป็น gBest ดังนั้นตำแหน่งใหม่ของ gBest คือ  $x=5.6374$  และ  $y=7.0245$

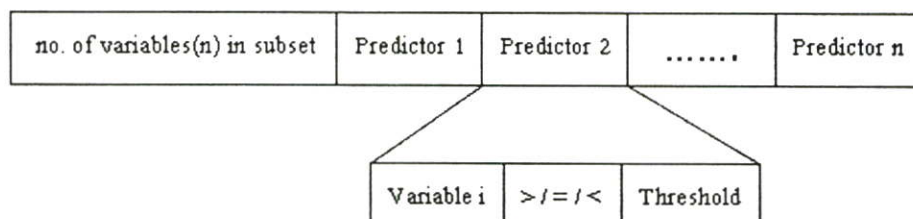
## 2.3 บทความงานวิจัยที่เกี่ยวข้อง

### 2.3.1 A GAs based approach for mining breast cancer pattern [12]

งานวิจัยนี้ถูกคิดค้นโดย Ta-Cheng Chen และ Tung-Chou Hsu ในปี 2006 โดยงานวิจัยนี้เป็นการนำ Genetic Algorithms (GAs) และกฎการทำนาย (Prediction rules) มาประยุกต์ใช้ร่วมกันเพื่อใช้วินิจฉัยโรคมะเร็งเต้านม โดยนำ GA มาใช้ในการค้นหากฎในการทำนาย

กฎในการทำนายจะถูกแสดงโดยชุดของเลขฐานสอง (String of binary digits) โดยแต่ละ string จะประกอบด้วย substring ที่มี 3 ส่วนประกอบ คือ Attribute ที่ใช้ในการทำนาย, เครื่องหมาย

ที่ใช้เปรียบเทียบ ( $>$ ,  $=$ ,  $<$ ) และ ค่าเกณฑ์ของแต่ละกฎ (Threshold) โดยโมเดลนี้จะมีการสร้างกฎเพิ่ม (additional rule) ในกรณีที่ถูกที่มีอยู่ก่อนไม่สามารถที่จะทำนายข้อมูลให้ถูกต้องได้ โดยกฎใหม่ที่เพิ่มขึ้นจะถูกสร้างจากชุดข้อมูลที่ทำนายผิด ซึ่งการสร้างกฎเพิ่มจะเป็นการช่วยปรับปรุงค่าความถูกต้องในการทำนายให้ดียิ่งขึ้น



รูปที่ 2.11 แสดงรูปแบบของโครโมโซมที่ใช้ในการค้นหากฎการทำนาย

โครงสร้างการทำงานของโมเดลประกอบด้วย 3 ส่วน ดังนี้

### 1) Data Preprocess

ทำการเตรียมข้อมูลโดยนำชุดของข้อมูลมาแบ่งออกเป็น 2 ส่วน คือ ชุดข้อมูลที่ใช้ในการเรียนรู้โมเดล (training dataset) และ ชุดข้อมูลที่ใช้ในการทดสอบโมเดล (validating dataset) หลังจากนั้นทำการเลือกคลาสที่ต้องการจะหาการทำนาย (เป็น/ไม่เป็นมะเร็ง) และทำการเลือกรูปแบบอินพุตจากชุดข้อมูลที่ใช้ในการเรียนรู้โมเดลตามคลาสที่เลือกไว้ ผลลัพธ์ที่ได้จากขั้นตอนนี้คือ ชุดของข้อมูลที่ใช้ในการเรียนรู้โมเดลที่ประกอบด้วยข้อมูลเฉพาะคลาสที่ถูกเลือก

### 2) Rule Mining Process

1. ทำการกำหนดประชากรเริ่มต้นโดยการสุ่ม
2. ทำการแปลงแต่ละโครโมโซมให้อยู่ในรูปแบบของ *If-Then* rule
3. นำข้อมูลที่ใช้ในการเรียนรู้โมเดลมาประเมินค่าความถูกต้องของแต่ละกฎ
4. ตรวจสอบการมาบรรจบกัน (Converge) ของโครโมโซม ถ้าโครโมโซมยังไม่บรรจบกันให้นำชุดของโครโมโซมไปผ่านตัวดำเนินการของ GA (Selection, Crossover และ Mutation) เพื่อให้ได้มาซึ่งโครโมโซมรุ่นใหม่ วนกลับไปทำขั้นตอนที่ 2. ซ้ำจนกระทั่งทุกโครโมโซมมาบรรจบกัน

5. ตรวจสอบการสร้างกฎเพิ่ม โดยถ้ากฎที่มีอยู่ไม่สามารถทำนายข้อมูลได้ถูกต้องทั้งหมดแล้วเราจะทำการสร้างกฎเพิ่มในขั้นตอนของ Data Modification Process แต่ถ้ากฎที่มีอยู่สามารถทำนายข้อมูลได้ถูกต้องทั้งหมดแล้วก็จะจบกระบวนการทำงาน โดยผลลัพธ์ที่ได้จากขั้นตอนนี้คือ กฎที่ใช้ในการตัดสินใจ (Decision rules)

### 3) Data Modification Process

ทำการปรับปรุงฐานข้อมูลโดยการตัดรูปแบบอินพุตที่ทำนายถูกออกจากชุดข้อมูลที่ใช้ในการเรียนรู้โมเดล ดังนั้น ฐานข้อมูลจะเหลือเพียงรูปแบบอินพุตที่ทำนายผิด ผลลัพธ์ที่ได้จากขั้นตอนนี้ คือ ฐานข้อมูลที่ถูกรับปรุงเพื่อนำไปใช้สร้างกฎใหม่ให้นำฐานข้อมูลใหม่ที่ได้อวนกลับไปทำ Rule Mining Process ขั้นตอนที่ 1

งานวิจัยนี้ได้ทำการทดลองโดยเปรียบเทียบกับซอฟต์แวร์ PolyAnalysis<sup>®</sup> โดยพบว่ากฎที่ถูกร่างจาก PolyAnalysis<sup>®</sup> มีการใช้ตัวแปรและค่าน้ำหนักจำนวนมากเมื่อเปรียบเทียบกับกฎที่ได้จากงานวิจัยนี้ และงานวิจัยนี้มีความถูกต้องในการทำนาย 96.995% ซึ่งมากกว่าค่าความถูกต้องที่ได้จาก PolyAnalysis<sup>®</sup> ถึง 0.425% โดยสามารถทำนายคนไข้ที่อยู่ในกลุ่มเป็นมะเร็งได้ถูกต้อง 100%

#### 2.3.2 Intelligence Particle Swarm Classifier (IPS-classifier) [13]

งานวิจัยนี้ถูกคิดค้นโดย Seyed-Hamid Zahiri และ Seyed-Alreza Seyedin ในปี 2005 โดยงานวิจัยนี้ทำการปรับปรุงประสิทธิภาพของ Particle Swarm Optimization (PSO) เพื่อให้มีประสิทธิภาพในการนำทางกลุ่มที่ดียิ่งขึ้น และพยายามนำทางกลุ่มเพื่อให้หลีกเลี่ยงจาก Local minimum โดยงานวิจัยมุ่งที่จะหาค่าที่เหมาะสมของ 3 ตัวแปร คือ ค่าน้ำหนักของแรงผลักดัน (Inertia weight:  $\omega_i$ ), ปัจจัยในการเรียนรู้ของตัวเอง (Cognitive parameter:  $c_{1i}$ ) และ ปัจจัยการเรียนรู้ทางสังคม (Social parameter:  $c_{2i}$ ) โดยพิจารณาจากค่าของ 3 ตัวแปร ดังนี้

1.  $f_{best}$  คือ ค่าความเหมาะสมที่สูงที่สุด
2. P-dist คือ ค่าความแตกต่างของตำแหน่งที่ดีที่สุดของแต่ละพาร์ทิเคิล ( $P_i$ ) และตำแหน่งที่ดีที่สุดของกลุ่ม ( $P_g$ ) ซึ่งคำนวณได้จากสมการที่ 2.8

$$P-dist = \|P_g - P_i\| = \sum_{k=1}^H |W_{kg} - W_{ki}| \quad (2.8)$$

3. UN คือ จำนวนรอบการทำงานที่ค่าความเหมาะสมของกลุ่มไม่มีการเปลี่ยนแปลง

งานวิจัยนี้จะใช้ค่า  $f_{best}$ , P-dist และ UN ในการหาค่าตัวแปร  $\omega_i$ ,  $c_{1i}$  และ  $c_{2i}$  ที่เหมาะสม โดยนำค่า  $f_{best}$ , P-dist และ UN มาสร้างเป็นกฎในการเรียนรู้ทั้งหมด 6 กฎ ดังนี้

- 1) **IF**  $f_{best}$  is low and P-dist is low **THEN**  $\omega_i$  is high,  $c_{1i}$  is high, and  $c_{2i}$  is low.
- 2) **IF**  $f_{best}$  is medium and UN is high and P-dist is high **THEN**  $\omega_i$  is medium,  $c_{1i}$  is high, and  $c_{2i}$  is medium.

- 3) IF  $f_{best}$  is medium and  $UN$  is low and  $P-dist$  is high THEN  $\omega_i$  is medium,  $c_{1i}$  is medium, and  $c_{2i}$  is medium.
- 4) IF  $f_{best}$  is high and  $UN$  is high THEN  $\omega_i$  is low,  $c_{1i}$  is low, and  $c_{2i}$  is high.
- 5) IF  $f_{best}$  is low and  $UN$  is high THEN  $\omega_i$  is high,  $c_{1i}$  is low, and  $c_{2i}$  is high.
- 6) IF  $f_{best}$  is high and  $P-dist$  is high THEN  $\omega_i$  is low,  $c_{1i}$  is low, and  $c_{2i}$  is high.

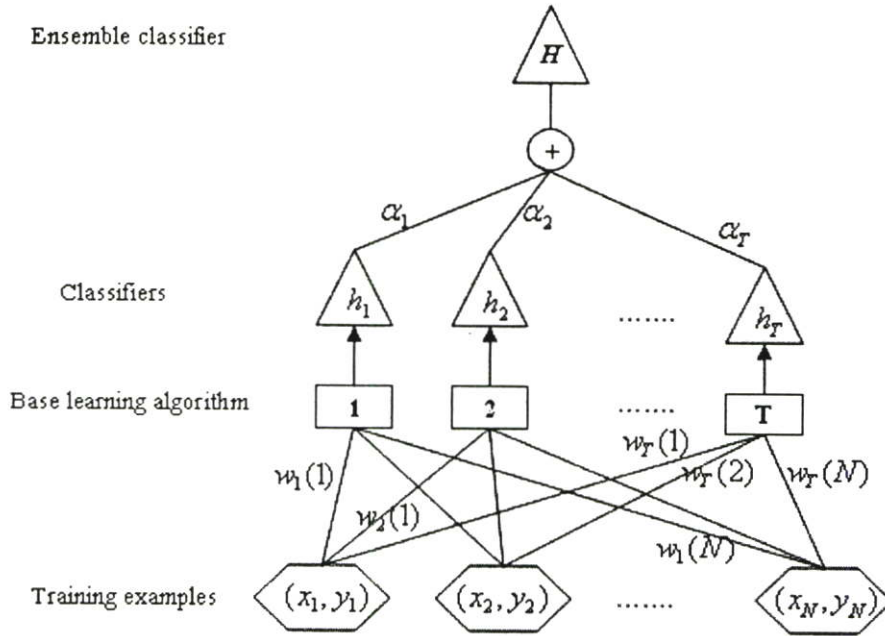
งานวิจัยนี้ได้ทำการทดลองกับชุดข้อมูล Iris, Wine และ Radar โดยเปรียบเทียบผลการทดลองที่ได้กับ MLP และ  $k$ -NN โดยในชุดข้อมูล Iris และ Wine พบว่าโดยเฉลี่ยแล้วงานวิจัยนี้ได้ค่าความถูกต้องมากกว่า MLP 2.942% แต่น้อยกว่า  $k$ -NN 1.095% และในชุดข้อมูล Radar พบว่าเมื่อ signal-to-noise ratios (SNR) มีค่าสูงขึ้นงานวิจัยนี้ก็จะมีค่าความถูกต้องในการทำนายสูงขึ้นด้วย โดยพบว่าที่ SNR=5 งานวิจัยนี้มีความถูกต้องมากกว่า MLP 3.2% และน้อยกว่า  $k$ -NN 1.4% แต่เมื่อ SNR=10 งานวิจัยนี้มีความถูกต้องมากกว่า MLP และ  $k$ -NN 22.6% และ 10.2% ตามลำดับ

### 2.3.3 Classification by evolutionary ensembles [14]

งานวิจัยนี้ถูกคิดค้นโดย Xiao Wang และ Han Wang ในปี 2006 งานวิจัยนี้ทำการสร้างโมเดลที่ใช้ทำนายกลุ่มข้อมูลโดยนำ Genetic Algorithm (GA) มาประยุกต์ใช้ในการหาชุดของค่าน้ำหนักที่เหมาะสมให้กับตัวทำนายกลุ่มข้อมูล (Classifier)

งานวิจัยนี้กำหนดให้ 1 โครโมโซมแทน 1 Classifier ในแต่ละ Generation เมื่อชุดของโครโมโซมผ่านการ mutate แล้วชุดของโครโมโซมจะถูกเก็บไว้เพื่อนำมาใช้เป็นตัวแทนของแต่ละ Classifier เมื่อได้โครโมโซมทั้งระบบแล้วจะทำการหาค่าความเหมาะสมของแต่ละโครโมโซมเพื่อนำมาคิดเป็นค่าน้ำหนักของแต่ละโครโมโซม ทำการคำนวณหาค่าเอาต์พุตโดยนำอินพุตมาผ่านโครโมโซมแต่ละตัว และนำค่าเอาต์พุตที่ได้จากโครโมโซมแต่ละตัวมารวมกับค่าน้ำหนักของโครโมโซมนั้น และทำการหาเอาต์พุตรวมของระบบโดยการนำเอาต์พุตสุดท้ายที่ได้จากแต่ละโครโมโซมมารวมกัน งานวิจัยนี้ใช้กับชุดข้อมูลที่มี 2 กลุ่มข้อมูลซึ่งสามารถหากลุ่มของข้อมูลได้ โดยดูจากเครื่องหมายของเอาต์พุตสุดท้ายที่ได้จากระบบ (+,-) สถาปัตยกรรมของงานวิจัยนี้แสดงได้ดังรูปที่ 2.12

โดยงานวิจัยนี้คาดว่าจะสามารถปรับปรุงค่าความถูกต้องและค่าความน่าเชื่อถือของระบบในกรณีที่มีบาง Classifier เกิดความล้มเหลวก็คาดว่าระบบโดยรวมจะสามารถกู้คืนค่าความผิดพลาดนั้นได้



รูปที่ 2.12 แสดงสถาปัตยกรรมของ Classification by evolutionary ensembles

2.3.3.1 ตัวดำเนินการทาง GA

ถ้ากำหนดให้โครโมโซมและค่าความเหมาะสมของ Generation ที่  $j-1$  คือ  $\{w_{j-1,k}\}_{k=1}^K$  และ  $\{f_{j-1,k}\}_{k=1}^K$

1) **Reproduction** ทำการเลือกโครโมโซมพ่อแม่โดยใช้ roulette-wheel โดยเราสามารถแบ่งพื้นที่ของแต่ละโครโมโซมตามค่าความเหมาะสมของโครโมโซมได้ดังสมการที่ 2.9 ซึ่งโครโมโซมที่มีค่าความเหมาะสมมากก็จะมีโอกาสที่จะถูกคัดลอกขึ้นไปสู่รุ่นถัดไปได้มาก

$$\left[ \frac{f_{j-1,1}}{\sum_{k=1}^K f_{j-1,k}}, \dots, \frac{f_{j-1,K}}{\sum_{k=1}^K f_{j-1,k}} \right] \tag{2.9}$$

2) **Crossover** โครโมโซมพ่อแม่ที่ได้จากการ Reproduction 1 คู่ จะถูกสร้างเป็นโครโมโซมลูกได้เพียง 1 โครโมโซมเท่านั้น โดยการนำโครโมโซมพ่อแม่มารวมกันแบบ linear combination ดังสมการที่ 2.10 ค่าน้ำหนักของ  $w_{offspring}$  เมื่อรวมกันแล้วจะมีค่าเท่ากับ 1

$$w_{offspring} = \frac{f_{j-1,k1} * w_{j-1,k1} + f_{j-1,k2} * w_{j-1,k2}}{f_{j-1,k1} + f_{j-1,k2}} \tag{2.10}$$

3) **Mutation** ทำการสุ่มค่าน้ำหนักจาก  $w_{offspring}$  มา  $N/2$  คู่ (ถ้า  $N$  เป็นเลขคี่แล้วจะสุ่มค่าน้ำหนักมา  $(N-1)/2$  คู่) แต่ละคู่จะประกอบด้วยค่าน้ำหนัก 2 ตัว เราจะทำการแลกเปลี่ยนค่าน้ำหนัก 2 ตัว ด้วยค่าความน่าจะเป็นในการ mutation ( $p_m$ )

ตัวดำเนินการทั้ง 3 ขั้นตอนของ GA จะถูกทำซ้ำทั้งหมด  $K$  ครั้ง ในแต่ละครั้งชุดของโครโมโซมที่ผ่านขั้นตอนการ mutate จะถูกเก็บเป็นสมาชิกใน generation ที่  $j$

### 2.3.3.2 ขั้นตอนการทำงาน

1) ทำการกำหนดค่าน้ำหนักของประชากรโดยแยกเป็น 2 กรณี ดังนี้

a. ใน generation แรก ( $j=1$ ) ทำการกำหนดค่าน้ำหนัก  $\{w_{1,k}\}_{k=1}^K$  โดยการสุ่ม ตาม uniform distribution ในช่วง  $(0,1)$

b. ใน generation ต่อๆ มา ( $j=2, \dots, J$ ) ค่าน้ำหนักของแต่ละโครโมโซม  $\{w_{j,k}\}_{k=1}^K$  จะถูกสร้างจากโครโมโซมในรุ่นก่อนหน้า  $\{w_{j-1,k}\}_{k=1}^K$  โดยใช้ตัวดำเนินการทาง GA ดังหัวข้อ 2.3.3.1 เพื่อสร้างเป็นประชากรใน generation ถัดไป

2) ป้อนค่า Training set  $\{S; w_{j,k}\}$  ให้กับอัลกอริทึมในการเรียนรู้

3) คำนวณค่าความผิดพลาดของโครโมโซมลำดับที่  $k$  ใน Generation ที่  $j$  ( $\varepsilon_{j,k}$ ) โดยการเปรียบเทียบกลุ่มของอินพุตที่ได้จากการทำนายโดยใช้โครโมโซมลำดับที่  $k$  ใน Generation ที่  $j$  ( $h_{j,k}(x_i)$ ) กับกลุ่มจริงที่ต้องการ ( $y_i$ ) โดยถ้า  $y_i = h_{j,k}(x_i)$  แล้ว  $I = 0$  แต่ถ้า  $y_i \neq h_{j,k}(x_i)$  แล้ว  $I = 1$  หลังจากนั้นนำค่าที่ได้มารวมกับค่าน้ำหนักของโครโมโซมนั้นดังสมการที่ 2.11

$$\varepsilon_{j,k} = \sum_{i=1}^N w_{j,k}(i) * I(y_i \neq h_{j,k}(x_i)) \quad (2.11)$$

4) หาค่าความเหมาะสมของโครโมโซม  $w_{j,k}$  ได้ดังสมการที่ 2.12

$$f_{j,k} = \frac{1}{2} \log \frac{1 - \varepsilon_{j,k}}{\varepsilon_{j,k}} \quad (2.12)$$

5) ทำซ้ำข้อ 1) จนครบตาม generation ที่กำหนดไว้ เมื่อครบตาม generation ที่กำหนดไว้ เราจะได้ classifier ตามจำนวนโครโมโซมที่ถูกสร้างขึ้น

6) ทำการหาค่าน้ำหนัก  $\alpha_{j,k}$  ของแต่ละ classifier ที่ถูกสร้างขึ้น ซึ่ง Classifier ที่มีค่าความเหมาะสมมากก็จะถูกโหวตให้มีค่าน้ำหนักมาก ดังสมการที่ 2.13

$$\alpha_{j,k} = \frac{f_{j,k}}{\sum_{j=1}^J \sum_{k=1}^K f_{j,k}} \quad (2.13)$$

7) ทำการหาผลลัพธ์โดยนำผลลัพธ์ที่ได้จากทุก Classifier มารวมกันตามค่าน้ำหนักของแต่ละ Classifier ดังสมการที่ 2.14 แล้วทำการดูเครื่องหมายของคำตอบที่ได้ว่าเป็นบวกหรือลบ เพื่อนำไปเปรียบเทียบกับคลาสต่อไป

$$H(.) = \text{sign} \left( \sum_{j=1}^J \sum_{k=1}^K \alpha_{j,k} h_{j,k}(\cdot) \right) \quad (2.14)$$

งานวิจัยนี้ทำการทดสอบกับชุดข้อมูล UCI จำนวน 20 ชุดข้อมูล โดยทำการเปรียบเทียบค่าความผิดพลาดเฉลี่ยที่ได้จากการทำนายระหว่าง 3 โมเดล คือ Bagging, Adaboost และ งานวิจัยนี้ (Classification by evolutionary ensembles) จากการทดลองพบว่า Adaboost ชนะ 11 ชุดข้อมูล Bagging ชนะ 5 ชุดข้อมูล และงานวิจัยนี้ชนะ 3 ชุดข้อมูล ซึ่งเมื่อพิจารณาจากค่าความผิดพลาดเฉลี่ยที่ได้จากทั้ง 3 โมเดลพบว่าค่าที่ได้จากงานวิจัยนี้จะเข้าใกล้ค่าของโมเดลที่ชนะและจะได้ค่าที่ดีกว่าโมเดลที่แพ้ในหลายๆ กรณี

#### 2.3.4 New evolutionary neural network [15]

งานวิจัยนี้ถูกคิดค้นโดย Wei Gao ในปี 2005 โดยงานวิจัยมุ่งที่จะปรับปรุงประสิทธิภาพของ Evolutionary Neural Network (ENN) ซึ่งเป็นโมเดลที่รวมการทำงานของ Evolutionary computation และ Neural Network เข้าด้วยกัน โดยงานวิจัยนี้จะพัฒนาสถาปัตยกรรมและค่าน้ำหนักที่เชื่อมต่อไปพร้อมๆ กัน เพื่อที่จะปรับปรุง neural network ดังนั้น โครงสร้างของ ENN ในงานวิจัยนี้จะใช้ 3 layers feedforward neural network

แต่ละโครง โมโซมหรือแต่ละเน็ตเวิร์คจะประกอบด้วยจำนวนโหนดในชั้น hidden ค่าน้ำหนักระหว่าง input กับ hidden ค่าน้ำหนักระหว่าง hidden กับ output และค่า threshold ของทั้ง network

### ขั้นตอนการทำงาน

1. กำหนดค่าเริ่มต้นของตัวแปร เช่น จำนวนโหนดอินพุต จำนวนโหนดเอาต์พุต ขอบเขตของโหนดในชั้น hidden และขอบเขตของค่าน้ำหนัก
2. สร้างโครโมโซมหรือเน็ตเวิร์คเริ่มต้นตามจำนวนประชากรที่กำหนดไว้ โดยทำการสุ่มจำนวนโหนดในชั้น hidden และ ค่าน้ำหนักตามขอบเขตที่กำหนดไว้ให้กับแต่ละโครโมโซม
3. คำนวณค่าความผิดพลาดและค่าความเหมาะสมของแต่ละโครโมโซมดังสมการที่ 2.15 และ 2.16 ตามลำดับ และเพื่อที่จะปรับปรุง ENN Model ให้เป็นลักษณะทั่วไป (generalization) โครโมโซมในแต่ละรุ่นจะถูกฝึกสอนด้วยวิธี “Sample counterchanging method” คือ แต่ละโครโมโซมจะถูกฝึกสอนด้วยข้อมูลที่ได้จากการสุ่มถึงจำนวน 80% ของข้อมูลที่ใช้ในการเรียนรู้

$$E = \frac{1}{2} \sum_{a=1}^N \sum_{k=1}^M [y_k(w_k : x^a) - t_k^a]^2 \quad (2.15)$$

$$f = \frac{1}{1 + E} \quad (2.16)$$

#### 4. ตัวดำเนินการ mutation

ค่าความน่าจะเป็นในการ Mutation ของแต่ละโครโมโซมจะถูกคำนวณจากค่า Network temperature ของแต่ละโครโมโซมดังสมการที่ 2.17 โดย  $f_{\max}$  คือ ค่าความเหมาะสมของโครโมโซมที่ดีที่สุดจากประชากรทั้งหมด

$$T_1 = 1 - \frac{f}{f_{\max}} \quad (2.17)$$

ถ้าค่า  $T_1$  มากประสิทธิภาพของเน็ตเวิร์คจะต่ำ และค่าความน่าจะเป็นของการเปลี่ยนแปลงสถาปัตยกรรมจะมาก ในที่นี้จะใช้ค่าความเหมาะสมในการตัดสินใจ Mutate ว่า จะทำการ Mutate สถาปัตยกรรมหรือค่าน้ำหนัก ซึ่งการ Mutate แต่ละแบบมีลักษณะดังนี้

**4.1 การ Mutate สถาปัตยกรรมของ Network** โดยการ Mutate จำนวนโหนดในชั้น hidden ด้วยการสุ่มจำนวนโหนดในชั้น hidden ตามขอบเขตที่กำหนดไว้เพื่อใช้แทนที่จำนวนโหนดในชั้น hidden เดิม ซึ่งถ้าจำนวนโหนดในชั้น hidden ถูกเพิ่ม ให้ทำการสุ่มค่าน้ำหนักตามค่าน้ำหนักที่เพิ่มขึ้น แต่ถ้าจำนวนโหนดในชั้น hidden ถูกลดให้ทำการตัดค่าน้ำหนักที่เชื่อมต่อถึง

**4.2 การ Mutate ค่าน้ำหนักที่เชื่อมต่อ** โดยค่าน้ำหนักแต่ละเส้นจะถูก Mutate โดยใช้ Adaptive Cauchy mutation method ตามสมการที่ 2.18, 2.19 และ 2.20

$$W_i = W_i + \sigma_i \cdot T \cdot C_i(0,1) \quad (2.18)$$

$$\sigma_i = \frac{1.0}{\sqrt{\beta_i F(X) + \lambda_i}} \quad (2.19)$$

$$T = \frac{T_0}{F_{\max} - F_{avr}} \quad (2.20)$$

โดย	$\sigma_i$	คือ	Normal deviation of parameter
	$C_i(0,1)$	คือ	ค่าจากการสุ่มตาม Cauchy distribution ที่อยู่ในช่วง 0 ถึง 1
	$F(X)$	คือ	ค่าความเหมาะสมของแต่ละโครโมโซม
	$\beta_i, \lambda_i$	คือ	ตัวแปรจำเพาะ ( $\beta_i = 1, \lambda_i = 0$ )
	T	คือ	adaptive parameter

### 5. ตัวดำเนินการ Selection

ตามวิธี Random competition method ค่าคะแนนใหม่ของแต่ละโครโมโซมจะถูกคำนวณจากค่าความเหมาะสมและค่า thickness ของแต่ละโครโมโซม โครโมโซมที่มีค่าความเหมาะสมมาก จะมีความน่าจะเป็นที่จะถูกเลือกมาก

ใน ENN model จะนำ thickness adjustment ใน immunized evolutionary programming ของแต่ละโครโมโซมมาใช้ คะแนนของแต่ละโครโมโซมจะถูกปรับโดย thickness term ซึ่งถ้าค่า thickness ของโครโมโซมมาก กระบวนการของ evolutionary จะมาบรรจบกันเร็ว โดยค่าคะแนน และค่า thickness ของแต่ละโครโมโซม (C) คำนวณได้จากสมการที่ 2.21 และ 2.22 ตามลำดับ

$$p'(i) \text{ scores} = p(i) \text{ scores} + \alpha \cdot C \cdot \left(1 - \frac{F(i)}{\max F}\right) \cdot p(i) \text{ scores} + \beta \cdot \frac{F(i)}{\max F} \cdot p(i) \text{ scores} \quad (2.21)$$

$$C = \frac{t \cdot (0.8 \cdot \max F \rightarrow \max F)}{N} \quad (2.22)$$

โดย	$\alpha, \beta$	คือ	adjustable parameter ( $\alpha = \beta = 0.5$ )
	MaxF	คือ	ค่าความเหมาะสมที่มากที่สุดจากประชากรทั้งหมด

## 6. เกณฑ์ในการหยุด

เกณฑ์ในการหยุดการทำงานจะขึ้นอยู่กับค่า Evolutionary generation และค่าความเหมาะสมที่สุดของแต่ละโครโมโซม

งานวิจัยนี้ได้ทำการทดลองกับปัญหา XOR โดยเปรียบเทียบผลการทดลองที่ได้กับ BP Network และ ENN ดังเดิม ซึ่งจะเปรียบเทียบในด้านจำนวนรอบในการทำงานและค่าความผิดพลาดที่เกิดขึ้น โดยแต่ละโมเดลมีค่าของตัวแปรต่างๆ ดังนี้

- BP Network : learning rate = 0.6, momentum = 0.2 และ สถาปัตยกรรมของเน็ตเวิร์ค คือ 2-2-1
- ENN ดังเดิม :  $P_m=0.05$ ,  $P_c=0.8$  และ  $N=30$
- New ENN : ขนาดประชากร=100, threshold=200 ขอบเขตของค่าน้ำหนักคือ -0.1 ถึง 0.1 และ สถาปัตยกรรมของเน็ตเวิร์ค คือ 2-2-1

จากการทดลองพบว่าจำนวนรอบการทำงานของ BP Network, ENN ดังเดิม และของงานวิจัยนี้ (New ENN) คือ 700000, 41 และ 14 ตามลำดับ และมีค่าความผิดพลาด คือ 0.00014, 0.00007 และ 0 ตามลำดับ

### 2.3.5 การรู้จำตัวอักษรธรรมอีสานด้วยจินตคณิตอัลกอริทึม (Isan Dhamma Character Recognition Using Genetic Algorithm) [16]

งานวิจัยนี้ถูกคิดค้นโดย A.Thammano และ T.Supin ในปี 2007 โดยงานวิจัยนี้ได้นำจินตคณิตมาประยุกต์ใช้เพื่อหาโครโมโซมที่ดีที่สุดและเหมาะสมที่สุด โดยใช้ฮาลอัสเตอร์ฟิสแทนซ์ คำนวณหาค่าความเหมาะสม ในงานวิจัยนี้ไม่มีการดึงคุณลักษณะเด่นของภาพออกมาแต่จะใช้ค่าข้อมูลของรูปภาพทั้งรูปภาพที่เป็นภาพแบบไบนารี (binary image) มาคำนวณเพื่อหาค่าความเหมาะสม และมีการเก็บค่าความผิดพลาดของคำตอบในประชากรรุ่นปัจจุบัน เพื่อใช้เป็นข้อมูลในการสร้างโครโมโซมใหม่และได้เสนอการทำมิวเตชัน โดยการเลือกยีนตัวอักษรที่มีความผิดพลาดสูง มาทำการมิวเตชันทำให้เกิดยีนตัวอักษรหรือรูปแบบตัวอักษรใหม่จึงเป็นการเพิ่มรูปแบบตัวอักษรให้มามากขึ้น

ในการเรียนรู้เริ่มต้นด้วยการนำรูปภาพแบบไบนารีที่มีขนาด 15 X 15 พิกเซล มาผ่านกระบวนการเตรียมข้อมูล (Preprocessing) เพื่อให้ได้รูปภาพที่มีเพียงลักษณะเส้นโครงร่างของรูปภาพ และจึงนำรูปภาพนั้นเข้าสู่ระบบการเรียนรู้

#### ขั้นตอนการเรียนรู้

1. สร้างประชากรเริ่มต้นหรือโครโมโซมเริ่มต้น โดยการสุ่มรูปแบบอินพุตแต่ละตัวมาเป็นยีนในโครโมโซม โดยแต่ละโครโมโซมที่สร้างขึ้นต้องแตกต่างกันอย่างน้อย 1 ยีน

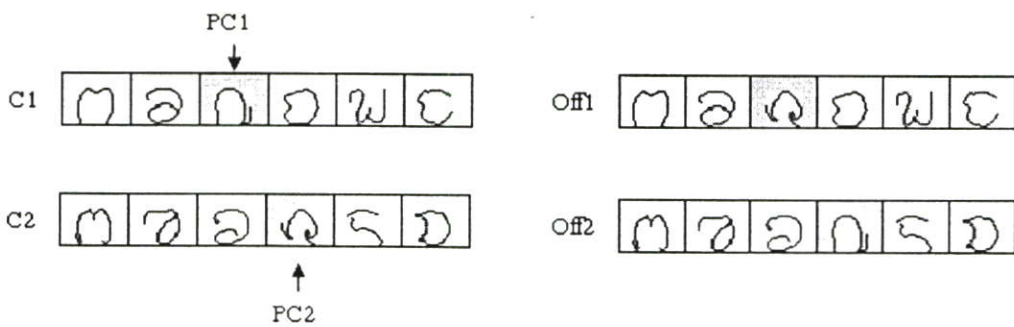
1	2	3	4	5	6	...	35
ก	อ	น	พ	ค	ย	...	ธ
ก	ข	น	พ	ย	ย	...	ล

รูปที่ 2.13 แสดงตัวอย่างโครโมโซมที่สร้างขึ้น

2. นำรูปแบบอินพุทที่ต้องการเรียนรู้ไปคำนวณหาค่าความเหมาะสมด้วยการคำนวณค่าฮาอูสคอรฟิดิสแทนส์ระหว่างรูปแบบอินพุทกับยีนตัวอักษรทุกยีนในโครโมโซม หายินที่ชนะโดยการพิจารณาค่าฮาอูสคอรฟิดิสแทนส์ที่น้อยที่สุดในแต่ละโครโมโซมว่าอยู่ที่ยีนใด ทำการเปรียบเทียบกลุ่มของรูปแบบอินพุทและกลุ่มของยีนที่ชนะ ค่าความเหมาะสมของแต่ละโครโมโซมคือผลรวมของค่าคำตอบของยีนที่ตรงกับค่าคำตอบของอินพุท

3. คัดเลือกโครโมโซมพ่อแม่ 2 โครโมโซมโดยพิจารณาจากค่าความเหมาะสมของโครโมโซม โดยใช้วงล้อถ่วงน้ำหนัก (Roulette - wheel selection)

4. สุ่มเลือกตำแหน่งของยีนที่จะทำการครอสโอเวอร์ในโครโมโซมพ่อแม่ตัวที่ 1 (PC1) ทำการหายีนในโครโมโซมพ่อแม่ตัวที่ 2 ที่มีกลุ่มตรงกับกลุ่มของยีนในตำแหน่ง PC1 (PC2) ทำการครอสโอเวอร์โดยการสลับยีนระหว่างยีนในตำแหน่ง PC1 กับ PC2 หลังผ่านการครอสโอเวอร์แล้วจะเกิดโครโมโซมลูกขึ้นมา 2 โครโมโซม โดยในการทดลองจะทำการครอสโอเวอร์ทั้งหมด  $N/2$  ครั้ง และนำเฉพาะประชากรที่เกิดจากการครอสโอเวอร์ไปเรียนรู้ในขั้นตอนที่ 5-7



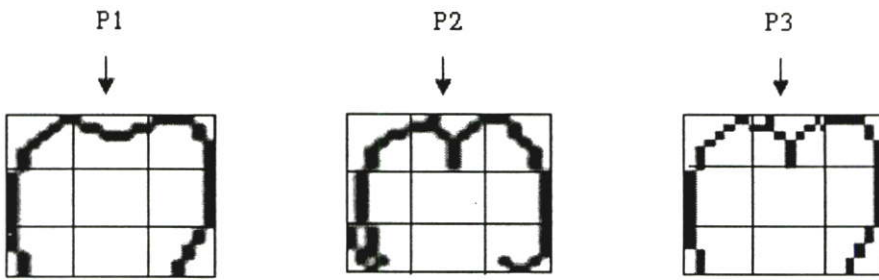
ก) โครโมโซมก่อนการครอสโอเวอร์

ข) โครโมโซมหลังการครอสโอเวอร์

รูปที่ 2.14 แสดงโครโมโซมก่อนและหลังการครอสโอเวอร์

5. ทำการมิวเตชันโดยการเลือกยีนตัวอักษรที่ให้คำตอบผิดมากที่สุดจากแต่ละโครโมโซม และทำการสุ่มเลือกยีนตัวอักษรที่จะมามิวเตทกับยีนที่อ้างอิงในโครโมโซมจากรูปแบบอินพุท

ทั้งหมด (P2) ทำการแบ่งยีนตัวอักษรทั้ง 2 ยีนออกเป็น 9 บล็อก ทำการสุ่มเลือกบล็อกที่จะทำการมิวเตท สลับค่าบล็อกที่สุ่มเลือกได้ระหว่าง 2 ยีน



รูปที่ 2.15 แสดงยีนตัวอักษรก่อนและหลังการมิวเตท

6. ตรวจสอบการเพิ่มยีนและการลบยีน โดยการสุ่มค่าความน่าจะเป็นที่จะเพิ่มหรือลบยีนในโครโมโซม (P)

ก. ถ้า  $P \geq 0.25$  จะทำการเพิ่มยีน

ทำการสุ่มค่ากลุ่มตัวอักษรที่ต้องการเพิ่ม สุ่มโครโมโซมที่จะทำการเพิ่ม และสุ่มเลือกตัวอักษรจากอินพุตแพทเทินที่มีกลุ่มตัวอักษรเดียวกับค่ากลุ่มที่สุ่มเลือกได้ ทำการเพิ่มเข้าไปในโครโมโซมที่ได้จากการสุ่มเลือก

ข. ถ้า  $P < 0.25$  จะทำการลบยีน

ทำการสุ่มเลือกโครโมโซม และตรวจสอบโครโมโซมนั้นว่ามียีนที่มีค่าความผิดพลาดสูงที่สุดอยู่หรือไม่ ถ้ามีให้ทำการลบยีนนั้นโดยมีเงื่อนไขว่าจะต้องมียีนตัวอักษรที่เป็นกลุ่มเดียวกับยีนที่ผิดพลาดสูงที่สุดมากกว่า 2 ยีนขึ้นไปจึงจะสามารถลบยีนนั้นได้

7. กำหนดค่าความเหมาะสมของชุดของโครโมโซมที่ได้

8. นำชุดของโครโมโซมก่อนการครอสโอเวอร์ (ขั้นตอนที่ 4) มารวมกับชุดของโครโมโซมที่ได้จากการครอสโอเวอร์ (ผ่านขั้นตอนการเรียนรู้ 5-7) และทำการเลือกโครโมโซมที่มีค่าความเหมาะสมสูงที่สุดตามจำนวนประชากรที่ตั้งไว้มาเป็นประชากรรุ่นต่อไป

9. วนกลับไปทำขั้นตอนที่ 2 ซ้ำจนกระทั่งครบตามจำนวนรอบการทำงานที่กำหนดไว้หรือค่าความเหมาะสมของโครโมโซมที่ดีที่สุดไม่มีการเปลี่ยนแปลง

งานวิจัยนี้ได้ทำการทดลองโดยทำการทำนายกลุ่มข้อมูลของตัวอักษรธรรมีสานจำนวน 35 กลุ่มตัวอักษร และนำไปเปรียบเทียบกับผลการทำนายที่ได้จากกลุ่มคนที่ใช้ตัวอักษรธรรมีสาน โดยจากผลการทดลองพบว่าคนสามารถทำนายกลุ่มของตัวอักษรธรรมีสานได้ถูกต้อง 41.10 % และงานวิจัยนี้สามารถทำนายได้ถูกต้อง 64.85% ซึ่งมากกว่าคน 23.75%

### บทที่ 3

## การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้าน พฤติกรรมทางสังคมของมนุษย์

### 3.1 บทนำ

สิ่งที่นักสังคมวิทยามุ่งศึกษาและพยายามอธิบายคือความสัมพันธ์ระหว่างพฤติกรรมของมนุษย์กับสิ่งที่เรียกว่าสังคม นักสังคมวิทยาของพฤติกรรมต่างๆ ของมนุษย์ว่าเป็น “พฤติกรรมทางสังคม (social behavior)” มนุษย์คิดและกระทำภายใต้การบีบคั้นโดยพลังของกฎเกณฑ์ทางสังคมอันได้แก่ความเชื่อทางศาสนา จารีตประเพณี กฎหมาย ค่านิยมของกลุ่ม และความกดดันทั้งหลายอันมาจากกลุ่มไม่ว่ากลุ่มเล็กๆ ที่เราเป็นสมาชิกโดยตรงหรือกลุ่มขนาดใหญ่ สังคมเป็นสิ่งที่มีความจริงดำรงอยู่ต่างหากจากตัวมนุษย์แต่ละคน และสามารถมีพลังบังคับควบคุมมนุษย์ได้อย่างกว้างขวางและตลอดเวลา Emile Durkheim ปรมาจารย์ทางสังคมวิทยาเรียกสิ่งนี้ว่า “social fact” [11]

นักสังคมวิทยาชี้ให้เห็นว่าพฤติกรรมทางสังคมของมนุษย์มีแบบแผนที่ซ้ำๆ กันจนสามารถวางเป็นกฎทั่วๆ ไปได้ และสามารถคาดคะเนได้ สังคมยังมีความซับซ้อนเพียงใดส่วนต่างๆ ก็ยังแบ่งแยกหน้าที่ออกจากกันมากเท่านั้น และยังคงพึ่งพาซึ่งกันและกันมากขึ้น ดังนั้นการดำรงอยู่ทางสังคมจึงขึ้นอยู่กับการทำหน้าที่ของส่วนต่างๆ ที่ต้องประสานสอดคล้องกัน โดยปกติเราใช้คำว่ากลุ่มเรียกคนจำนวนหนึ่งตั้งแต่สองคนขึ้นไปที่รวมกันเป็นหมู่คณะ บางครั้งกลุ่มหมายถึงคนจำนวนหนึ่งหรือคนประเภทหนึ่งที่มีลักษณะบางอย่างร่วมกัน (Categories) กลุ่มต่างๆ ที่เราเป็นสมาชิกมีมากมายหลายกลุ่ม บางกลุ่มมีขนาดเล็กประกอบด้วยสมาชิกเพียงไม่กี่คน บางกลุ่มมีสมาชิกรับร้อยนับพัน บางกลุ่มรวมตัวกันชั่วคราว บางกลุ่มดำรงอยู่อย่างถาวร [11]

นักสังคมวิทยาสร้างแนวความคิดที่เรียกว่ากลุ่มวงใน (In-group) และ กลุ่มวงนอก (Out-group) เพื่อชี้ให้เห็นถึงวงเขตของกลุ่ม (Group boundary) คนที่อยู่ในวงเขตของกลุ่มหนึ่งก็คือผู้ที่ได้รับการยอมรับว่าเป็นพวกเดียวกัน ส่วนพวกที่ถูกจัดว่าไม่ใช่พวกของเราคือคนที่อยู่นอกวงเขตของกลุ่ม สมาชิกของกลุ่มวงในมักพยายามรักษา “ระยะห่างทางสังคม (Social distance)” กับกลุ่มวงนอก โดยการสร้างความรู้สึกอคติหรือเป็นปฏิปักษ์กับกลุ่มอื่น [11]

มีนักคิดหลายๆ คนได้แสดงความคิดเห็นเกี่ยวกับกลุ่มวงในและกลุ่มวงนอกไว้มากมาย เช่น ในปี 1906 Sumner เห็นว่าธรรมชาติของกลุ่มวงในกับกลุ่มวงนอกมักจะมีทัศนคติที่แตกต่างกัน วัฒนธรรมของเผ่าพันธุ์ต่างๆ จะมีทัศนคติในแง่บวกต่อคนในกลุ่ม และมีความสัมพันธ์โดยตรงทางด้านความคิดที่เป็นศัตรู และมีความเกลียดชังต่อคนนอกกลุ่ม ในปี 1966 Sherif ได้นำเสนอทฤษฎีทางด้านพฤติกรรมระหว่างกลุ่มซึ่งเขาเชื่อว่าคนในกลุ่มเดียวกันมีรูปแบบการพึ่งพาอาศัยซึ่ง

กันและกันในแง่บวกเพื่อให้บรรลุเป้าหมายที่วางไว้ แต่ในทางตรงกันข้ามความสัมพันธ์ระหว่างกลุ่มจะมีลักษณะของการแข่งขันและการพึ่งพาอาศัยกัน ในแง่ลบ ในปี 1981 Tajfel ได้ตั้งสมมติฐานทางด้านความคล้ายคลึงและความแตกต่างว่าการรับรู้ที่เหมือนกันจะเพิ่มมากขึ้นเมื่อเป็นคนในกลุ่มเดียวกันซึ่งสัมพันธ์กับการรับรู้ความแตกต่างที่จะเพิ่มมากขึ้นเมื่อเป็นคนละกลุ่มกัน [9] ในปี 1969, 1978 Tajfel ได้เสนอว่าความขัดแย้งและความมีอคติในกลุ่มเป็นผลมาจากการที่แต่ละบุคคลมีการกระทำในลักษณะที่แยกตัวกันเป็นกลุ่ม ดังนั้นทั้งการขัดแย้งระหว่างกลุ่ม (Group conflict) และการจัดหมวดหมู่ทางสังคม (social categorization) จะนำมาสู่ความมีอคติ ความลำเอียง และการแบ่งแยก [10]

สมาชิกกลุ่มต่างๆ จะมีความสัมพันธ์กันมากขึ้นขึ้นอยู่กับองค์ประกอบที่สำคัญคือกลุ่มและการเป็นผู้นำของกลุ่ม กลุ่มที่มีขนาดใหญ่ย่อมมีความคิดเห็นแตกต่างกันออกไปมากกว่ากลุ่มที่มีสองหรือสามคนและอาจมีการจับกลุ่มกันเป็นกลุ่มย่อยมากกว่าหนึ่งหรือสองกลุ่ม

โมเดลที่นำเสนอจะมีลักษณะการทำงานที่เลียนแบบพฤติกรรมทางสังคมของมนุษย์คือสมาชิกที่อยู่ในกลุ่มเดียวกันหรือคลัสเตอร์ที่อยู่ในคลาสเดียวกันจะถูกดึงเข้าหากันแต่จะมีการเว้นระยะห่างระหว่างกันเพื่อไม่ให้แต่ละคลัสเตอร์อยู่ใกล้ชิดกันมากเกินไป และจะมีการผลักตัวเองให้ออกห่างจากคลัสเตอร์ที่อยู่คนละกลุ่มหรือคลัสเตอร์ที่อยู่คนละคลาสกัน

### 3.2 โครงสร้างการทำงาน

กำหนดให้เปอร์เซ็นต์ความถูกต้องเริ่มต้น = 0

3.2.1 ทำการการสุ่มรูปแบบอินพุท เพื่อสร้างเป็นคลัสเตอร์โดยดูจากค่าเปอร์เซ็นต์ความถูกต้อง ดังรายละเอียดที่แสดงในหัวข้อ 3.3.1

3.2.2 คำนวณค่าความถูกต้องหลังจากสร้างคลัสเตอร์ ดังรายละเอียดที่แสดงในหัวข้อ 3.3.2

3.2.3 ปรับค่าจุดศูนย์กลางของแต่ละคลัสเตอร์และคำนวณหาค่าความถูกต้องและนำไปเปรียบเทียบกับค่าความถูกต้องในขั้นตอนที่ 3.2.2 เพื่อปรับปรุงค่าตัวแปรของแต่ละคลัสเตอร์ ดังรายละเอียดที่แสดงในหัวข้อ 3.3.3

3.2.4 ทำการลบคลัสเตอร์โดยดูจากจำนวนรูปแบบอินพุทและคลาสของรูปแบบอินพุทที่อยู่ในคลัสเตอร์นั้นๆ ดังรายละเอียดที่แสดงในหัวข้อ 3.3.4

3.2.5 ตรวจสอบค่าเปอร์เซ็นต์ความถูกต้องหลังลบคลัสเตอร์เพื่อใช้เป็นเงื่อนไขในการสร้างคลัสเตอร์ในรอบถัดไป ดังรายละเอียดที่แสดงในหัวข้อ 3.3.5

3.2.6 ทำการสุ่มค่าระยะทางให้จุดศูนย์กลางของแต่ละคลัสเตอร์เพื่อหาจุดศูนย์กลางใหม่ที่ดีกว่า ดังรายละเอียดที่แสดงในหัวข้อ 3.3.6

3.2.7 วนกลับไปทำงานข้อ 3.2.1 จนกระทั่งครบรอบการเรียนรู้ที่กำหนดไว้ หรือความผิดพลาดที่เกิดจากการเรียนรู้มีค่าเป็น 0

### 3.3 กระบวนการเรียนรู้โมเดล

ทำการกำหนดค่าเริ่มต้นของตัวแปรต่างๆ ดังนี้

- เปอร์เซนต์ความถูกต้องเริ่มต้น ( $percent_{correct}$ ) = 0
- อัตราความเร็วระหว่างคลัสเตอร์  $i$  และ คลัสเตอร์  $j$  ที่จะใช้เดินเข้าหากัน ( $\delta_{ij}$ )
- อัตราเร็วของแต่ละคลัสเตอร์ ( $\beta_i$ )
- ค่าเกณฑ์ในการลบคลัสเตอร์ (Vigilance :  $\rho$ )
- ความกว้างของ Gaussian ( $\sigma$ ) = 0.1
- ระยะห่างระหว่างจุดศูนย์กลางของ 2 คลัสเตอร์ที่อยู่ในคลาสเดียวกัน (DisConst) = 0.3
- จำนวนรอบที่ใช้ดูเปอร์เซนต์ความถูกต้องว่ายังเพิ่มขึ้นอยู่หรือไม่ (MaxRound) = 30
- ขอบเขตของค่าระยะทางที่เราจะสุ่มให้กับแต่ละคลัสเตอร์ = -0.1 ถึง 0.1
- จำนวนครั้งที่เราจะสุ่มค่าระยะทางให้แต่ละคลัสเตอร์ = 5

กระบวนการเรียนรู้ของโมเดลมีขั้นตอนการทำงานดังนี้

#### 3.3.1 ขั้นตอนการสร้างคลัสเตอร์

ทำการสร้างคลัสเตอร์โดยดูจากเปอร์เซนต์ความถูกต้อง โดยถ้าการสร้างคลัสเตอร์เพิ่มทำให้เปอร์เซนต์ความถูกต้องเพิ่มขึ้นเราก็จะยังคงสร้างคลัสเตอร์เพิ่มต่อไป แต่ถ้าเมื่อใดก็ตามที่การสร้างคลัสเตอร์เพิ่ม ไม่สามารถทำให้เปอร์เซนต์ความถูกต้องเพิ่มขึ้นได้แล้วเราจะหยุดการสร้างคลัสเตอร์เพิ่มแต่จะทำการสร้างคลัสเตอร์แทนที่คลัสเตอร์ที่ถูกลบไปในรอบก่อนหน้าแทนซึ่งจะไม่ทำให้จำนวนคลัสเตอร์เพิ่มขึ้น ซึ่งเงื่อนไขในการสร้างคลัสเตอร์มีดังนี้

- ถ้าค่าเปอร์เซนต์ความถูกต้องยังคงเพิ่มขึ้น เราจะทำการสุ่มรูปแบบอินพุตจากคลาสที่มีค่าความผิดพลาดเกิดขึ้นคลาสนั้น 1 รูปแบบอินพุต
- ถ้าค่าเปอร์เซนต์ความถูกต้องไม่เพิ่มขึ้นแล้ว เราจะทำการสุ่มรูปแบบอินพุตจากคลาสนี้ที่มีคลัสเตอร์ถูกลบไปในรอบก่อนหน้าตามจำนวนที่ถูกลบไป

ขั้นตอนการพิจารณาเปอร์เซนต์ความถูกต้องจะถูกกล่าวถึงในขั้นตอนที่ 3.3.5

### 3.3.2 ขั้นตอนการคำนวณค่าความถูกต้อง

คำนวณค่าความถูกต้องหลังจากขั้นตอนสร้างคลัสเตอร์ โดยดูจากจำนวนรูปแบบอินพุทที่ทำนายถูก เพื่อนำไปใช้เปรียบเทียบในขั้นตอนที่ 3.3.3

การคำนวณค่าความถูกต้องมีขั้นตอนการทำงานดังนี้

3.3.2.1 กำหนดค่าเริ่มต้นของตัวนับจำนวนครั้งที่ทำนายถูก ( $Cnt$ ) = 0

3.3.2.2 รับรูปแบบอินพุทมาครั้งละ 1 รูปแบบอินพุท

3.3.2.3 ใช้ Euclidean Distance คำนวณหาค่าระยะห่างระหว่างรูปแบบอินพุทกับแต่ละจุดศูนย์กลางของคลัสเตอร์

$$dist_{ij} = \sqrt{\sum_{k=1}^K (x_{ik} - center_{jk})^2} \quad (3.1)$$

โดย	$i$	คือ ลำดับของรูปแบบอินพุท
	$j$	คือ ลำดับของคลัสเตอร์
	$k$	คือ ลำดับของคุณลักษณะ
	$K$	คือ จำนวนคุณลักษณะ
	$dist_{ij}$	คือ ระยะห่างระหว่างรูปแบบอินพุทลำดับที่ $i$ กับคลัสเตอร์ลำดับที่ $j$
	$x_{ik}$	คือ รูปแบบอินพุทลำดับที่ $i$ คุณลักษณะที่ $k$
	$center_{jk}$	คือ จุดศูนย์กลางของคลัสเตอร์ลำดับที่ $j$ คุณลักษณะที่ $k$

3.3.2.4 เลือกคลัสเตอร์ที่มีระยะห่างจากรูปแบบอินพุทน้อยที่สุดมาเป็นโหนดที่ชนะ

$$winner = \min (dist_j) \quad (3.2)$$

3.3.2.5 เปรียบเทียบคลาสของรูปแบบอินพุทกับคลาสของโหนดที่ชนะ

– ถ้าเป็นคลาสเดียวกันถือว่าทำนายถูก ให้ทำการนับจำนวนครั้งที่ถูก

$$Cnt = Cnt + 1 \quad (3.3)$$

– ถ้าเป็นคนละคลาสถือว่าทำนายผิด

3.3.2.6 วนกลับไปทำข้อ 3.3.2.2 จนกระทั่งครบทุกรูปแบบอินพุท โดยค่าความถูกต้องที่ได้คือจำนวนของรูปแบบอินพุทที่ทำนายถูก

### 3.3.3 ปรับค่าจุดศูนย์กลางของทุกคลัสเตอร์

ทำการปรับค่าจุดศูนย์กลางของทุกคลัสเตอร์ โดยจุดศูนย์กลางของคลัสเตอร์ที่อยู่ในคลาสเดียวกันจะถูกดึงเข้าหากัน แต่จะมีการรักษาระยะห่างไว้เพื่อไม่ให้คลัสเตอร์อยู่ใกล้กันมากเกินไป ส่วนจุดศูนย์กลางของคลัสเตอร์ที่อยู่คนละคลาสกันจะถูกผลักออกจากกัน โดยแรงดึงหรือแรงผลักจะขึ้นอยู่กับระยะทางที่คลัสเตอร์นั้นๆ มีต่อกัน โดยถ้าอยู่ใกล้แรงดึงหรือแรงผลักจะมาก แต่ถ้าอยู่ไกลแรงดึงหรือแรงผลักจะน้อย ขั้นตอนการปรับค่าจุดศูนย์กลางของคลัสเตอร์มีดังนี้

3.3.3.1 รับจุดศูนย์กลางของคลัสเตอร์ที่เราต้องการจะปรับค่าจุดศูนย์กลางมาครั้งละ 1 คลัสเตอร์ในที่นี้เราเรียกว่าคลัสเตอร์  $i$

3.3.3.2 รับจุดศูนย์กลางของคลัสเตอร์อื่นๆ ที่ไม่ใช่คลัสเตอร์  $i$  มาครั้งละ 1 คลัสเตอร์ในที่นี้เราเรียกว่าคลัสเตอร์  $j$  เพื่อนำมาใช้ในการปรับค่าจุดศูนย์กลางของคลัสเตอร์  $i$

3.3.3.3 ใช้ Euclidean Distance คำนวณระยะห่างระหว่างจุดศูนย์กลางของคลัสเตอร์  $i$  และจุดศูนย์กลางของคลัสเตอร์  $j$

$$dist_{ij}^k = \sqrt{(center_{ik} - center_{jk})^2} \quad (3.4)$$

โดย  $dist_{ij}^k$  คือ ระยะห่างระหว่าง  $center_{ik}$  กับ  $center_{jk}$   
 $center_{ik}$  คือ จุดศูนย์กลางของคลัสเตอร์ลำดับที่  $i$  คุณลักษณะที่  $k$   
 $center_{jk}$  คือ จุดศูนย์กลางของคลัสเตอร์ลำดับที่  $j$  คุณลักษณะที่  $k$

ถ้าคลัสเตอร์  $i$  และ คลัสเตอร์  $j$  เป็นคลัสเตอร์ที่อยู่ในคลาสเดียวกันเราจะทำการตรวจสอบว่าค่าระยะห่างที่ได้จากสมการ 3.4 ( $dist_{ij}$ ) น้อยกว่าค่าระยะห่างที่กำหนดไว้หรือไม่ โดยถ้า  $dist_{ij} < DistConst$  แสดงว่าจุดศูนย์กลางของคลัสเตอร์  $i$  และ จุดศูนย์กลางของคลัสเตอร์  $j$  อยู่ใกล้กันมากเกินไปอาจทำให้เกิดการทับซ้อนซึ่งกันและกัน ซึ่งจะเป็นผลทำให้เกิดความผิดพลาดในการทำนายมากยิ่งขึ้น ดังนั้นเราจะทำการปรับค่าอัตราเร็วระหว่างคลัสเตอร์  $i$  และ คลัสเตอร์  $j$  ให้เดินเข้าหากันด้วยอัตราเร็วที่ช้าลงดังสมการที่ 3.5

$$\delta_{ij}^k = \delta_{ij}^k - 0.1 \quad (3.5)$$

โดย  $\delta_{ij}^k$  คือ อัตราความเร็วระหว่างคลัสเตอร์  $i$  และ คลัสเตอร์  $j$  ที่คุณลักษณะลำดับที่  $k$

3.3.3.4 ทำการปรับค่าจุดศูนย์กลางของคลัสเตอร์  $i$  โดยพิจารณาจากจุดศูนย์กลางของคลัสเตอร์อื่นๆ ที่มีอยู่ โดยคลัสเตอร์  $i$  จะเดินเข้าหาคลัสเตอร์ที่อยู่ในคลาสเดียวกัน และผลัก

ออกจากคลัสเตอร์ที่อยู่คนละคลาสตามระยะห่างระหว่างคลัสเตอร์และตามอัตราเร็วที่มีอยู่ โดยถ้าคลัสเตอร์  $i$  อยู่ในคลาส  $P$  แล้ว จะทำการปรับค่าจุดศูนย์กลางของคลัสเตอร์  $i$  ได้ดังสมการที่ 3.6

$$center_{ik} = center_{ik} + \sum_{j \in P} \beta_i \delta_{ij}^k \eta_{ij}^k (center_{jk} - center_{ik}) - \sum_{j \notin P} \beta_i \delta_{ij}^k \eta_{ij}^k (center_{jk} - center_{ik}) \quad (3.6)$$

โดยที่  $k$  คือ ลำดับของคุณลักษณะ  
 $center_i$  คือ คลัสเตอร์ที่เราจะปรับค่าจุดศูนย์กลาง  
 $center_j$  คือ คลัสเตอร์ที่มีผลต่อการปรับค่าจุดศูนย์กลางของคลัสเตอร์  $i$   
 $\eta_{ij}^k$  คือ ค่าความเป็นสมาชิกที่คลัสเตอร์  $i$  มีต่อคลัสเตอร์  $j$  ที่คุณลักษณะลำดับที่  $k$

โดยในที่นี้จะใช้ Gaussian Function ดังสมการที่ 3.7 ในการคำนวณหาอัตราการเรียนรู้ของคลัสเตอร์  $j$  ที่มีต่อคลัสเตอร์  $i$  ที่คุณลักษณะลำดับที่  $k$  ( $\eta_{ij}^k$ ) โดยค่าอัตราการเรียนรู้ (แรงดึงหรือแรงผลัก) ที่  $center_j$  มีต่อ  $center_i$  จะขึ้นอยู่กับว่าคลัสเตอร์นั้นๆ มีความเป็นสมาชิกต่อกันมากหรือน้อยแค่ไหน ถ้ามีความเป็นสมาชิกมากแรงดึงหรือแรงผลักที่มีต่อกันก็จะมาก แต่ถ้ามีความเป็นสมาชิกลดลงแรงดึงหรือแรงผลักที่มีต่อกันก็จะน้อย

$$\eta_{ij}^k = e^{-\left(\frac{center_{ik} - center_{jk}}{\sigma}\right)^2} \quad (3.7)$$

โดย  $\sigma$  คือ ความกว้างของ Gaussian  
 $center_i$  คือ จุดศูนย์กลางของคลัสเตอร์ที่เราต้องการจะปรับค่า  
 $center_j$  คือ จุดศูนย์กลางของคลัสเตอร์ที่มีผลต่อการปรับค่า  $center_i$

3.3.3.5 วนไปทำข้อ 3.3.3.2 จนกระทั่งครบทุกคลัสเตอร์  $j$

3.3.3.6 ตรวจสอบค่าความถูกต้องที่เกิดจากการปรับค่าจุดศูนย์กลางของคลัสเตอร์  $i$  โดยมีขั้นตอนการทำงานเหมือนกับข้อ 3.3.2

3.3.3.7 เปรียบเทียบค่าความถูกต้องที่ได้จากครั้งก่อนหน้า ( $correct_{old}$ ) กับค่าความถูกต้องที่ได้จากข้อ 3.3.3.6 ( $correct_{new}$ ) โดยมีเงื่อนไขดังนี้

– ถ้า  $correct_{new} < correct_{old}$  หมายความว่า การปรับจุดศูนย์กลางของคลัสเตอร์  $i$  ทำให้เกิดความถูกต้องลดลง ดังนั้น เราจะกลับไปใช้จุดศูนย์กลางจุดเดิม (ก่อนปรับค่าจุดศูนย์กลาง)

– ถ้า  $correct_{new} > correct_{old}$  หมายความว่า การปรับจุดศูนย์กลางของคลัสเตอร์  $i$  ทำให้ค่าความถูกต้องเพิ่มขึ้น ดังนั้น จะใช้จุดศูนย์กลางใหม่ที่ได้ และทำการเก็บค่าความถูกต้องใหม่ เพื่อใช้เปรียบเทียบในการปรับจุดศูนย์กลางของคลัสเตอร์ถัดไป

– ถ้า  $correct_{new} = correct_{old}$  หมายความว่า การปรับจุดศูนย์กลางของคลัสเตอร์  $i$  ไม่ทำให้ค่าความผิดพลาดเปลี่ยนแปลง ดังนั้น ใช้จุดศูนย์กลางใหม่ที่ได้ แต่เราจะทำการลดอัตราเร็วในการเดินของคลัสเตอร์  $i$  ( $\beta_i$ ) ให้เดินด้วยอัตราเร็วที่ช้าลงเพื่อให้เกิดเป็นการเดินที่ระมัดระวังมากขึ้น โดยปรับอัตราเร็วของคลัสเตอร์  $i$  ดังสมการที่ 3.8

$$\beta_i = \beta_i * 0.6 \quad (3.8)$$

3.3.3.8 ทำซ้ำตั้งแต่ข้อ 3.3.3.1 จนกระทั่งปรับค่าจุดศูนย์กลางครบทุกคลัสเตอร์

### 3.3.4 ขั้นตอนการลบคลัสเตอร์

3.3.4.1 ทำการเก็บค่าว่าแต่ละคลัสเตอร์ประกอบด้วยรูปแบบอินพุทที่ทำนายถูกและทำนายผิดอย่างละกี่รูปแบบอินพุท

3.3.4.2 รับคลัสเตอร์เข้ามาทีละ 1 คลัสเตอร์เพื่อคำนวณหาประสิทธิภาพของคลัสเตอร์  $j$  ( $rate_j$ )

$$rate_j = \left[ -\frac{1}{2 \log_2 \frac{n_j}{N}} \right] * \left[ \frac{n_j^{correct} * \frac{\sum_{i=1, x, ecenter_j}^{n_j} \|x_i - center_j\|}{n_j}}{n_j} \right] \quad (3.9)$$

โดยที่	$N$	คือ	จำนวนข้อมูลที่ใช้ในการเรียนรู้โมเดล (training data)
	$n_j$	คือ	จำนวนรูปแบบอินพุทในคลัสเตอร์ $j$
	$n_j^{correct}$	คือ	จำนวนรูปแบบอินพุทในคลัสเตอร์ $j$ ที่ทำนายถูกคลาส
	$x_i$	คือ	รูปแบบอินพุทลำดับที่ $i$
	$center_j$	คือ	จุดศูนย์กลางของคลัสเตอร์ $j$

จากสมการที่ 3.9 สามารถอธิบายความหมายของแต่ละเทอมได้ดังนี้

$$\text{ก. เทอม} \frac{n_j^{\text{correct}}}{n_j} * \frac{\sum_{i=1, x_i \in \text{center}_j}^{n_j} \|x_i - \text{center}_j\|}{n_j}$$

โดยเทอมหน้า คือ อัตราส่วนของจำนวน

รูปแบบอินพุทในคลัสเตอร์  $j$  ที่ทำนายถูกเทียบกับจำนวนรูปแบบอินพุทในคลัสเตอร์  $j$  ทั้งหมด และเทอมหลัง คือ การวัดขนาดความกว้างโดยเฉลี่ยของคลัสเตอร์  $j$  โดยถ้าคลัสเตอร์  $j$  มีรูปแบบอินพุทที่ทำนายผิดก็จะส่งผลต่อเทอมหลัง คือ ขนาดความกว้างที่จะถูกลดค่าลงตามจำนวนรูปแบบอินพุทที่ทำนายผิด

$$\text{ข. เทอม} - \frac{1}{2 \log_2 \frac{n_j}{N}}$$

คือ การหาค่าอัตราส่วนของจำนวนรูปแบบอินพุทในคลัสเตอร์  $j$  เมื่อเปรียบเทียบกับจำนวนรูปแบบอินพุททั้งหมด (training data)

3.3.4.3 เปรียบเทียบค่า  $rate_j$  กับ ค่า  $\rho$  ที่กำหนดไว้ โดยถ้า  $rate_j < \rho$  แล้วแสดงว่าคลัสเตอร์  $j$  มีค่าต่ำกว่าเกณฑ์ที่ตั้งไว้ ดังนั้น คลัสเตอร์  $j$  จะถูกลบทิ้ง

3.3.4.4 วนทำซ้ำข้อ 3.3.4.2 จนกระทั่งครบทุกคลัสเตอร์

3.3.4.5 ตรวจสอบค่าความถูกต้องที่เกิดจากการลบคลัสเตอร์โดยมีขั้นตอนการทำงานเหมือนกับข้อ 3.3.2

### 3.3.5 ขั้นตอนการพิจารณาค่าเปอร์เซ็นต์ความถูกต้อง

เราจะทำการพิจารณาค่าเปอร์เซ็นต์ความถูกต้อง ( $percent_{\text{correct}}$ ) โดยดูว่าจำนวนคลัสเตอร์ที่มีอยู่เพียงพอต่อการเรียนรู้ของโมเดลหรือไม่ โดยดูจากความถูกต้องที่คำนวณได้จากขั้นตอนที่ 3.3.4.5 ซึ่งถ้าความถูกต้องยังคงเพิ่มขึ้นเรื่อยๆ ก็น่าจะบอกได้ว่าถ้าเราทำการเพิ่มจำนวนคลัสเตอร์ ความถูกต้องก็อาจจะยังเพิ่มขึ้นได้อีก แต่ถ้าความถูกต้องไม่เพิ่มขึ้นเกินกว่าค่าความถูกต้องสูงสุดที่เคยได้ ก็น่าจะบอกได้ว่าจำนวนคลัสเตอร์ที่มีอยู่เพียงพอต่อการเรียนรู้แล้ว ไม่จำเป็นต้องสร้างคลัสเตอร์เพิ่ม ซึ่งการตรวจสอบในขั้นตอนนี้จะมีผลต่อการสร้างคลัสเตอร์ในขั้นตอนที่ 3.3.1 ในรอบการเรียนรู้ถัดไป

จากการคำนวณค่าเปอร์เซ็นต์ความถูกต้องเราจะทำการเก็บค่าเปอร์เซ็นต์ความถูกต้องที่สูงที่สุด ( $percent_{\text{max}}$ ) ที่เคยได้ไว้เพื่อใช้ในการเปรียบเทียบค่าเปอร์เซ็นต์ความถูกต้องในรอบต่อไป โดย

ก. ถ้าเปอร์เซ็นต์ความถูกต้องในรอบปัจจุบันมีค่าน้อยกว่าหรือเท่ากับค่า  $percent_{\text{max}}$  ให้ทำการนับจำนวนรอบที่เปอร์เซ็นต์ความถูกต้องมีค่าน้อยกว่าหรือเท่ากับค่า  $percent_{\text{max}}$  โดย

- ถ้าจำนวนรอบยังคงน้อยกว่าจำนวนรอบ (MaxRound) ที่กำหนดไว้ ให้ทำการนับจำนวนรอบไปเรื่อยๆ จนกระทั่งเท่ากับ MaxRound หรือจนกระทั่งพบความถูกต้องที่มากกว่า  $percent_{\text{max}}$  ที่มีอยู่

– ถ้าจำนวนรอบเท่ากับจำนวนรอบที่กำหนดไว้ ให้หยุดการนับจำนวนรอบ และจบการทำงานของขั้นตอนการพิจารณาค่าเปอร์เซ็นต์ความถูกต้อง

ข. ถ้าเปอร์เซ็นต์ความถูกต้องในรอบปัจจุบันมีค่ามากกว่าค่า  $percent_{max}$  ให้ทำการเก็บค่าใหม่ที่ได้และเริ่มต้นนับจำนวนรอบใหม่

### 3.3.6 ทำการสุ่มค่าระยะทางให้จุดศูนย์กลาง

ในแต่ละรอบการทำงานเราจะทำการสุ่มค่าระยะทางเพื่อนำไปรวมกับจุดศูนย์กลางของแต่ละคลัสเตอร์เพื่อให้การพิจารณาเส้นทางการเดินของคลัสเตอร์เกิดความหลากหลายมากขึ้น โดยมีขั้นตอนการทำงาน ดังนี้

3.3.6.1 รับจุดศูนย์กลางของคลัสเตอร์เข้ามาทีละ 1 คลัสเตอร์

3.3.6.2 ทำการตรวจสอบจุดศูนย์กลางของคลัสเตอร์ของรอบปัจจุบัน ( $center_i^{new}$ ) กับรอบก่อนหน้า ( $center_i^{old}$ )

– ถ้า  $|center_i^{old} - center_i^{new}| < 0.0001$  แสดงว่าจุดศูนย์กลางของคลัสเตอร์ของรอบก่อนหน้ากับรอบปัจจุบันมีการเปลี่ยนแปลงน้อยมาก เราจะเริ่มทำการสุ่มค่าระยะทางตามขอบเขตที่กำหนดไว้ และนำไปรวมกับจุดศูนย์กลางของคลัสเตอร์ เพื่อหาจุดศูนย์กลางใกล้เคียงที่คิดว่าจุดเดิม โดยจะทำการสุ่มค่าระยะทางขนาด  $K$  ( $K =$  จำนวนคุณลักษณะทั้งหมด) และนำมารวมกับค่าจุดศูนย์กลางเดิมดังสมการที่ 3.10

$$center_{ik} = center_{ik} + testPath_k \quad (3.10)$$

โดย  $i$  คือ ลำดับของคลัสเตอร์

$k$  คือ ลำดับของคุณลักษณะ

$center_{ik}$  คือ จุดศูนย์กลางของคลัสเตอร์ที่เราจะทำการสุ่มระยะทางให้

$testPath$  คือ ค่าระยะทางที่ได้จากการสุ่มเพื่อนำไปรวมกับจุด  $center_i$

3.3.6.3 คำนวณหาค่าความถูกต้องได้จากระยะทางใหม่ โดยขั้นตอนการคำนวณหาค่าความถูกต้องมีขั้นตอนการทำงานเหมือนกับข้อ 3.3.2

3.3.6.4 เปรียบเทียบค่าความถูกต้องเก่า ( $correct_{old}$ ) กับค่าความถูกต้องที่ได้จากข้อ 3.3.6.3 ( $correct_{new}$ )

– ถ้า  $correct_{old} \geq correct_{new}$  แล้วจะใช้จุดศูนย์กลางเดิมและทำการสุ่มค่าใหม่จนครบจำนวนครั้งที่กำหนดไว้

– ถ้า  $correct_{old} < correct_{new}$  แล้วเราจะแทนที่จุดศูนย์กลางเก่าด้วยจุดศูนย์กลางใหม่ที่ได้และหยุดการสุ่มค่าระยะทางให้กับ  $center_i$  ทำการเก็บค่า  $correct_{new}$  ไว้ใช้เปรียบเทียบในการสุ่มระยะทางให้จุดศูนย์กลางของคลัสเตอร์ถัดไป

3.3.6.5 วงกลับไปทำข้อ 3.3.6.1 จนกระทั่งครบทุกคลัสเตอร์

### 3.3.7 ขั้นตอนการวนซ้ำ

วงกลับไปทำงานข้อ 3.3.1 จนกระทั่งครบรอบการเรียนรู้ที่กำหนดไว้หรือค่าความผิดพลาดที่เกิดจากการเรียนรู้มีค่าเป็น 0

## 3.4 กระบวนการทดสอบโมเดล

เมื่อผ่านกระบวนการเรียนรู้โมเดลแล้วสิ่งที่ได้จากการเรียนรู้ และนำมาใช้ประโยชน์ต่อได้แก่ จำนวนคลัสเตอร์และจุดศูนย์กลางของคลัสเตอร์ โดยในขั้นตอนนี้เราจะทำการทดสอบว่าเมื่อนำข้อมูลทดลองมาผ่านกระบวนการเรียนรู้แล้วเราได้จำนวนคลัสเตอร์ และจุดศูนย์กลางของคลัสเตอร์ที่เหมาะสมกับชุดข้อมูลนั้นๆ มากแค่ไหน โดยนำจำนวนคลัสเตอร์ และจุดศูนย์กลางของแต่ละคลัสเตอร์ที่ได้มาใช้กับข้อมูลทดสอบ (Testing Data) ซึ่งมีขั้นตอนการทำงานดังนี้

3.4.1 รับรูปแบบอินพุตจากชุดข้อมูลทดสอบมาครั้งละ 1 รูปแบบอินพุต

3.4.2 ใช้ Euclidean Distance ในการคำนวณหาค่าระยะห่างระหว่างรูปแบบอินพุตกับแต่ละคลัสเตอร์

$$dist_{ij} = \sqrt{\sum_{k=1}^K (x_{ik} - center_{jk})^2} \quad (3.11)$$

โดย  $k$  คือ คุณลักษณะลำดับที่  $k$

$K$  คือ จำนวนคุณลักษณะ

$dist_j$  คือ ระยะห่างระหว่างรูปแบบอินพุตกับคลัสเตอร์ลำดับที่  $j$

$x_{ik}$  คือ รูปแบบอินพุตลำดับที่  $i$  คุณลักษณะที่  $k$

$center_{jk}$  คือ จุดศูนย์กลางของคลัสเตอร์ลำดับที่  $j$  คุณลักษณะที่  $k$

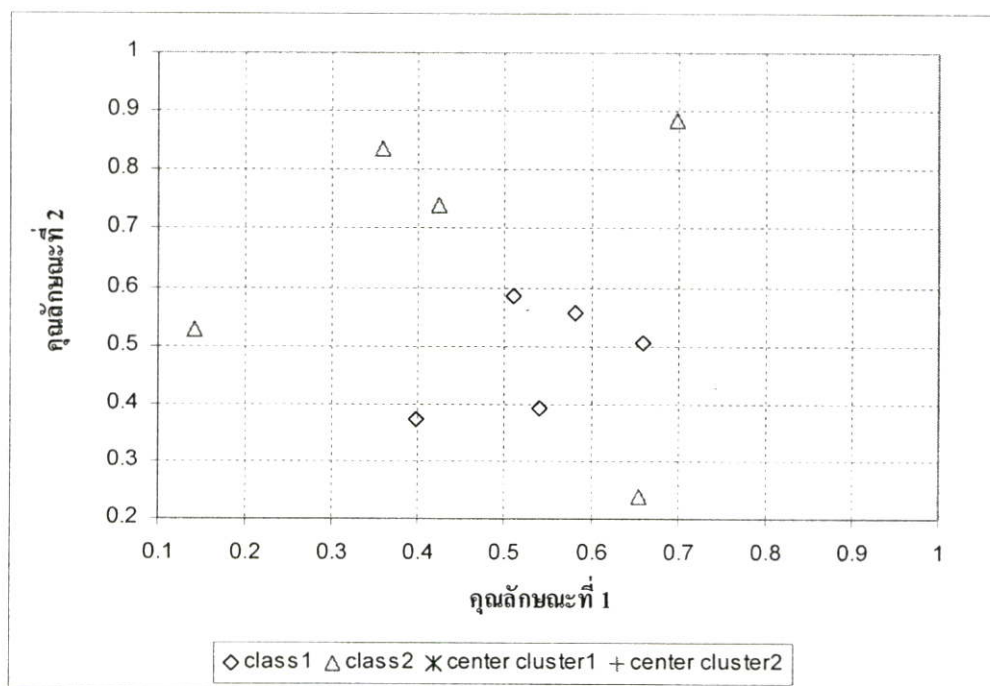
3.4.3 เลือกคลัสเตอร์ที่มีระยะห่างจากรูปแบบอินพุตน้อยที่สุดมาเป็นคลัสเตอร์ที่ชนะ

3.4.4 เปรียบเทียบคลาสของรูปแบบอินพุตกับคลาสของโหนดที่ชนะ ถ้าเป็นคลาสเดียวกันถือว่าทำนายถูก ถ้าเป็นคนละคลาสถือว่าทำนายผิด

3.4.5 วงกลับไปทำข้อ 3.4.1 จนกระทั่งครบทุกรูปแบบอินพุตของชุดข้อมูลทดสอบ

### 3.5 ตัวอย่างการทำงาน

- ข้อมูลที่ใช้มีจำนวน 10 รูปแบบอินพุท
- จำนวนคุณลักษณะ = 2
- จำนวนคลาส = 2
  1. 0.3981 0.3750 คลาส 1
  2. 0.6597 0.5046 คลาส 1
  3. 0.5093 0.5856 คลาส 1
  4. 0.5394 0.3935 คลาส 1
  5. 0.5810 0.5556 คลาส 1
  6. 0.1412 0.5278 คลาส 2
  7. 0.3588 0.8356 คลาส 2
  8. 0.6991 0.8843 คลาส 2
  9. 0.4236 0.7407 คลาส 2
  10. 0.6551 0.2384 คลาส 2



รูปที่ 3.1 แสดงรูปแบบอินพุทที่ใช้ในการเรียนรู้โมเดล โดย ◇ แทนรูปแบบอินพุทในคลาสที่ 1 และ △ แทนรูปแบบอินพุทในคลาสที่ 2

กำหนดค่าตัวแปรต่างๆ ดังนี้

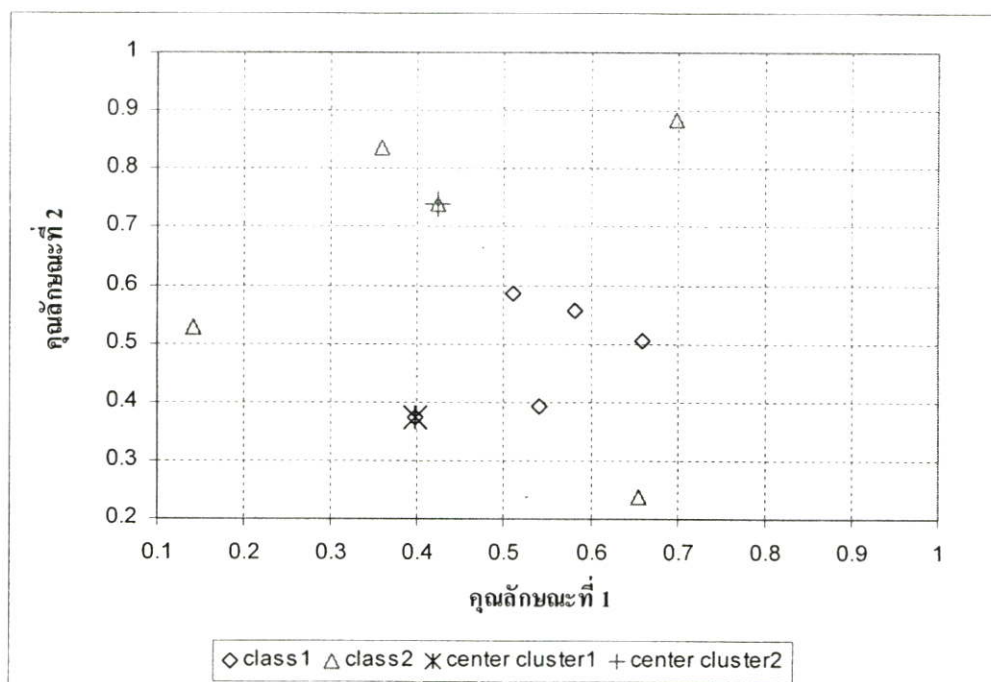
1. เปอร์เซนต์ความถูกต้องเริ่มต้น ( $percent_{correct}$ ) = 0
2. อัตราเร็วเริ่มต้นระหว่างคลัสเตอร์  $i$  และ คลัสเตอร์  $j$  ที่จะใช้เดินเข้าหากัน ( $\delta_{ij}$ ) = 1
3. อัตราเร็วเริ่มต้นของแต่ละคลัสเตอร์ ( $\beta_i$ ) = 1
4. ค่าที่ใช้เป็นเกณฑ์ในการลบคลัสเตอร์ (Vigilance :  $\rho$ ) = 0.02
5. ความกว้างของ Gaussian ( $\sigma$ ) = 0.1
6. ระยะห่างระหว่างจุดศูนย์กลางของ 2 คลัสเตอร์ที่อยู่ในคลาสเดียวกัน (DisConst) = 0.3
7. จำนวนรอบที่ใช้ตรวจสอบค่าเปอร์เซนต์ความถูกต้อง (MaxRound) = 30
8. จำนวนครั้งที่เราจะสุ่มค่าระยะทางให้กับแต่ละคลัสเตอร์ = 5
9. ขอบเขตของค่าระยะทางที่เราจะสุ่มให้กับแต่ละคลัสเตอร์ = -0.1 ถึง 0.1

### รอบการทำงานที่ 1

ทำการตรวจสอบว่ารูปแบบอินพุตที่เกิดความผิดพลาดอยู่ในคลาสใดบ้าง เนื่องจากเป็นรอบการทำงานแรก ดังนั้น ทั้ง 2 คลาสเป็นคลาสที่เกิดความผิดพลาด

1. สร้างคลัสเตอร์เพิ่มโดยการสุ่มจุดศูนย์กลางจากคลัสเตอร์ที่มีความผิดพลาดเกิดขึ้น ดังนั้นจึงทำการสุ่มจุดศูนย์กลางจากรูปแบบอินพุตที่อยู่ในคลัสเตอร์ที่ 1 และ 2 มาเป็นจุดศูนย์กลาง ได้ดังนี้

จุดศูนย์กลางที่ 1 คือ 0.3981 0.3750 คลาส 1 , จุดศูนย์กลางที่ 2 คือ 0.4236 0.7407 คลาส 2



รูปที่ 3.2 แสดงจุดศูนย์กลางเริ่มต้นของคลัสเตอร์ โดย \* แทนจุดศูนย์กลางของคลัสเตอร์ที่ 1 และ + แทนจุดศูนย์กลางของคลัสเตอร์ที่ 2

2. หาค่าความถูกต้องของรูปแบบอินพุตกับคลัสเตอร์ที่มีอยู่โดยใช้สมการ 3.1

รูปแบบอินพุตที่ 1 : 0.3981 0.3750 คลาส 1

$$dist_{11} = \sqrt{(0.3981 - 0.3981)^2 + (0.3750 - 0.3750)^2} = 0$$

$$dist_{12} = \sqrt{(0.3981 - 0.4236)^2 + (0.3750 - 0.7407)^2} = 0.3666$$

$\text{Min}(dist) = \text{Min}(0, 0.3666) = 0$  ดังนั้น คลัสเตอร์ที่ 1 เป็นผู้ชนะ

คลาสของคลัสเตอร์ 1 = คลาสของรูปแบบอินพุตที่ 1 ดังนั้น ทำนายถูก

รูปแบบอินพุตที่ 2 ถึง 10 ทำขั้นตอนเดียวกับรูปแบบอินพุตที่ 1 และนับจำนวนรูปแบบอินพุตที่ทำนายถูกคลาส ดังนั้นจะได้ค่าความถูกต้องหลังเพิ่มคลัสเตอร์ = 6 รูปแบบอินพุต

3. ปรับค่าจุดศูนย์กลางของทุกคลัสเตอร์

3.1 ปรับจุดศูนย์กลางของคลัสเตอร์ 1 โดยดูจากจุดศูนย์กลางของคลัสเตอร์ที่ 2

จุดศูนย์กลางของคลัสเตอร์ที่ 1 : 0.3981 0.3750 คลาส 1

จุดศูนย์กลางของคลัสเตอร์ที่ 2 : 0.4236 0.7407 คลาส 2

3.1.1 ทำการหาค่า  $\delta$  และ  $\eta$  เพื่อแทนค่าลงในสมการที่ 3.6

- ตรวจสอบค่า  $\delta_{12}^1 \cdot \delta_{12}^2 \rightarrow$  คลัสเตอร์ที่ 2 อยู่คนละคลาสกับคลัสเตอร์ที่ 1 ดังนั้น

$$\delta_{12}^1 \cdot \delta_{12}^2 = 1 \text{ (ไม่มีการเปลี่ยนแปลงค่า)}$$

- คำนวณหาค่า  $\eta_{12}^1 \cdot \eta_{12}^2$  โดยใช้สมการที่ 3.7 ได้ดังนี้

$$\eta_{12}^1 = e^{-\left(\frac{0.3981 - 0.4236}{0.1}\right)^2} = 0.9370 \text{ และ } \eta_{12}^2 = e^{-\left(\frac{0.3750 - 0.7407}{0.1}\right)^2} = 0$$

3.1.2 ปรับค่าจุดศูนย์กลางของคลัสเตอร์ที่ 1

3.1.2.1 คุณลักษณะที่ 1

$$\begin{aligned} center_{11} &= center_{11} - \beta_1 \delta_{12}^1 \eta_{12}^1 (center_{21} - center_{11}) \\ &= 0.3981 - [(1 * 1 * 0.9370) (0.4236 - 0.3981)] \\ &= 0.3981 - 0.0239 = 0.3742 \end{aligned}$$

3.1.2.2 คุณลักษณะที่ 2

$$\begin{aligned} center_{12} &= center_{12} - \beta_1 \delta_{12}^2 \eta_{12}^2 (center_{22} - center_{12}) \\ &= 0.3750 - [(1 * 1 * 0) (0.7407 - 0.3750)] \\ &= 0.3750 - 0 = 0.3750 \end{aligned}$$

3.1.3 หาค่าความถูกต้องหลังจากปรับจุดศูนย์กลางของคลัสเตอร์ที่ 1 ได้ค่าความถูกต้องคือ 6 รูปแบบอินพุต

3.1.4 ค่าความถูกต้องใหม่ = ค่าความถูกต้องเดิม ดังนั้นปรับค่า  $\beta_1$  โดยใช้สมการที่ 3.8 ได้ดังนี้  $\beta_1 = 1 * 0.6 = 0.6$

3.2 ปรับจุดศูนย์กลางของคลัสเตอร์ 2 โดยดูจากจุดศูนย์กลางของคลัสเตอร์ที่ 1

จุดศูนย์กลางของคลัสเตอร์ที่ 1 : 0.3742 0.3750 คลาส 1

จุดศูนย์กลางของคลัสเตอร์ที่ 2 : 0.4236 0.7407 คลาส 2

3.2.1 ทำการหาค่า  $\delta$  และ  $\eta$  เพื่อแทนค่าลงในสมการที่ 3.6

- ตรวจสอบค่า  $\delta_{21}^1, \delta_{21}^2 \rightarrow$  คลัสเตอร์ที่ 2 อยู่คนละคลาสกับคลัสเตอร์ที่ 1 ดังนั้น

$\delta_{21}^1, \delta_{21}^2 = 1$  (ไม่มีการเปลี่ยนแปลงค่า)

- กำหนดหาค่า  $\eta_{21}^1, \eta_{21}^2$  โดยใช้สมการที่ 3.7 ได้ดังนี้

$$\eta_{12}^1 = e^{-\left(\frac{0.3742-0.4236}{0.1}\right)^2} = 0.7835 \quad \text{และ} \quad \eta_{12}^2 = e^{-\left(\frac{0.3750-0.7407}{0.1}\right)^2} = 0$$

3.2.2 ปรับค่าจุดศูนย์กลางของคลัสเตอร์ที่ 2

3.2.2.1 คุณลักษณะที่ 1

$$\begin{aligned} center_{21} &= center_{21} - \beta_2 \delta_{21}^1 \eta_{21}^1 (center_{11} - center_{21}) \\ &= 0.4236 - [(1*1*0.7835) (0.3742 - 0.4236)] \\ &= 0.4236 - (-0.0387) = 0.4623 \end{aligned}$$

3.2.2.2 คุณลักษณะที่ 2

$$\begin{aligned} center_{22} &= center_{22} - \beta_2 \delta_{21}^2 \eta_{21}^2 (center_{12} - center_{22}) \\ &= 0.7407 - [(1*1*0) (0.3750 - 0.7407)] \\ &= 0.7407 - 0 = 0.7407 \end{aligned}$$

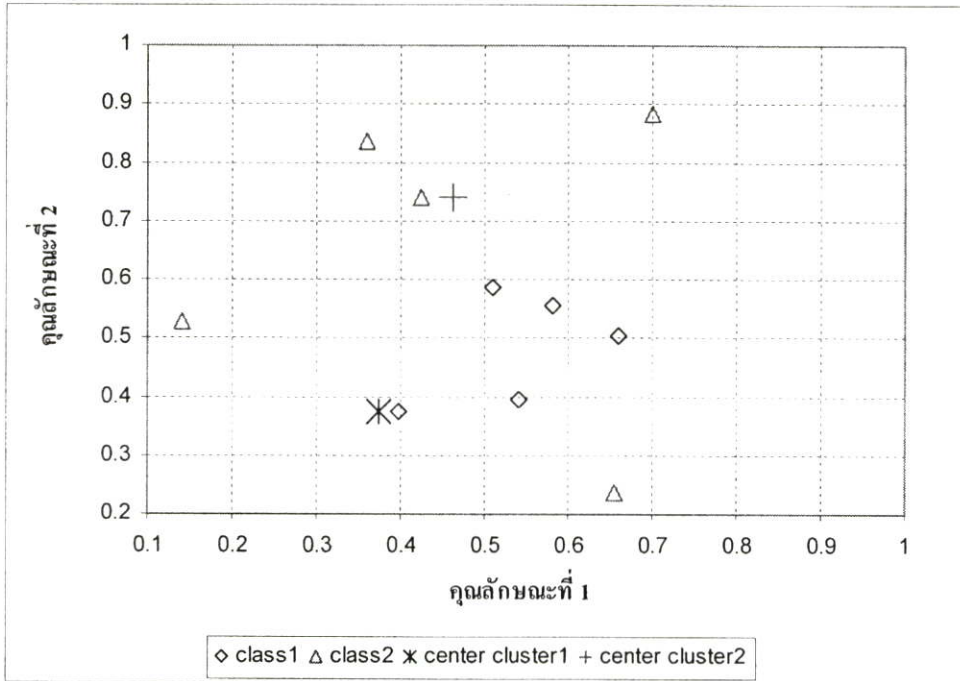
3.2.3 หาค่าความถูกต้องหลังจากปรับจุดศูนย์กลางของคลัสเตอร์ที่ 2 ได้ค่าความถูกต้องคือ 5 รูปแบบอินพุท

3.2.4 ค่าความถูกต้องใหม่ < ค่าความถูกต้องเดิม ดังนั้น กลับไปใช้จุดศูนย์กลางเดิมก่อนปรับค่า

หลังผ่านขั้นตอนการปรับค่าจุดศูนย์กลาง ได้ค่าจุดศูนย์กลางใหม่ดังนี้

จุดศูนย์กลางของคลัสเตอร์ที่ 1 :  $x = 0.3742$  และ  $y = 0.3750$

จุดศูนย์กลางของคลัสเตอร์ที่ 2 :  $x = 0.4623$  และ  $y = 0.7407$



รูปที่ 3.3 แสดงจุดศูนย์กลางของคลัสเตอร์ 1 และ 2 หลังการปรับค่าจุดศูนย์กลาง

4. การลบคลัสเตอร์

ทำการลบคลัสเตอร์โดยเริ่มจากการนำรูปแบบอินพุตทั้งหมดมาคำนวณหาคลัสเตอร์ที่ชนะ ซึ่งได้ผลลัพธ์ดังตารางที่ 3.1

ตารางที่ 3.1 แสดงคลัสเตอร์ที่ชนะเมื่อใช้คลัสเตอร์ 1 และ 2 ที่ผ่านการปรับค่าจุดศูนย์กลาง

รูปแบบอินพุต	x	y	Cluster 1		Cluster 2		คลัสเตอร์ที่ชนะ
			0.3742	0.3750	0.4623	0.7407	
1	0.3981	0.375	0.0239	0.3713	1		
2	0.6597	0.5046	0.3135	0.3077	2		
3	0.5093	0.5856	0.2502	0.1621	2		
4	0.5394	0.3935	0.1662	0.3557	1		
5	0.581	0.5556	0.2746	0.2199	2		
6	0.1412	0.5278	0.2786	0.3853	1		
7	0.3588	0.8356	0.4609	0.1404	2		
8	0.6991	0.8843	0.6041	0.2769	2		
9	0.4236	0.7407	0.3690	0.0387	2		
10	0.6551	0.2384	0.3124	0.5380	1		

4.1 จากตารางที่ 3.1 ทำการเก็บค่าจำนวนรูปแบบอินพุตที่ทำนายถูกและทำนายผิดของแต่ละคลัสเตอร์ ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 แสดงจำนวนรูปแบบอินพุตที่เป็นสมาชิกของแต่ละคลัสเตอร์

ลำดับคลัสเตอร์	จำนวนอินพุต		
	ถูกคลาส	ผิดคลาส	ผลรวม
1	2	2	4
2	3	3	6

4.2 คำนวณหาประสิทธิภาพของแต่ละคลัสเตอร์โดยใช้สมการที่ 3.9 ได้ดังนี้

4.2.1 คลัสเตอร์ 1

4.2.1.1 หาค่าเฉลี่ยของผลรวมของระยะทาง

$$\frac{0.0239 + 0.1662 + 0.2786 + 0.3124}{4} = \frac{0.7811}{4} = 0.1953$$

4.2.1.2 หาประสิทธิภาพของแต่ละคลัสเตอร์

$$rate_1 = \left[ -\frac{1}{2 \log_2 \frac{4}{10}} \right] * \left[ \frac{2}{4} * 0.1953 \right] = 0.0369$$

$0.0369 > 0.02$  (ค่า vigilance ที่กำหนดไว้) ดังนั้น คลัสเตอร์ 1 ไม่ถูกลบทิ้ง

4.2.2 คลัสเตอร์ 2

4.2.1.1 หาค่าเฉลี่ยของผลรวมของระยะทาง

$$\frac{0.3077 + 0.1621 + 0.2199 + 0.1404 + 0.2769 + 0.0387}{6} = \frac{1.1458}{6} = 0.2864$$

4.2.1.2 หาประสิทธิภาพของแต่ละคลัสเตอร์

$$rate_1 = \left[ -\frac{1}{2 \log_2 \frac{6}{10}} \right] * \left[ \frac{3}{6} * 0.2864 \right] = 0.0972$$

$0.0972 > 0.02$  (ค่า vigilance ที่กำหนดไว้) ดังนั้น คลัสเตอร์ 2 ไม่ถูกลบทิ้ง

5. ค่าความถูกต้องหลังจากผ่านขั้นตอนลบคลัสเตอร์ คือ 5 รูปแบบอินพุต คิดเป็นค่าความถูกต้อง 50%

6. ตรวจสอบค่าเปอร์เซ็นต์ความถูกต้องโดยค่าความถูกต้องเดิม คือ 0% และค่าความถูกต้องใหม่ คือ 50% แสดงว่าความถูกต้องยังคงเพิ่มขึ้น

7. ทำการสุ่มค่าระยะทางเพื่อหาจุดศูนย์กลางใหม่

7.1 สุ่มค่าระยะทางให้คลัสเตอร์ที่ 1 ซึ่งมีค่าจุดศูนย์กลางเดิม คือ  $x = 0.3742$  และ  $y = 0.3750$  คลาสที่ 1 ได้ค่าดังตารางที่ 3.3

ตารางที่ 3.3 แสดงการสุ่มระยะทางให้กับจุดศูนย์กลางของคลัสเตอร์ที่ 1

ลำดับการสุ่ม	ตัวเลขที่สุ่มได้		ค่าจุดศูนย์กลางใหม่		ค่าความถูกต้อง	
	x	y	x	y	เดิม	ใหม่
1	-0.0572	0.0814	0.3170	0.4564	5	5
2	-0.0304	0.0942	0.3438	0.4692	5	5
3	-0.0865	0.0217	0.2877	0.3967	5	5
4	-0.0531	0.0917	0.3211	0.4667	5	5
5	0.0454	-0.0268	0.4196	0.3482	5	6

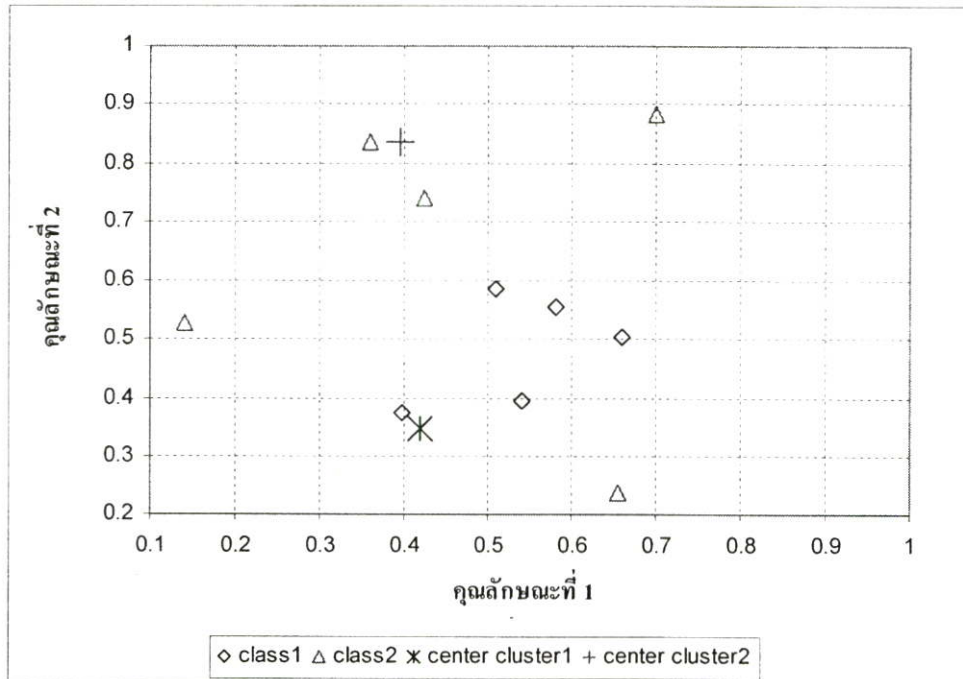
จากการสุ่มครั้งที่ 5 พบว่าจุดศูนย์กลางใหม่ที่ได้ทำให้ความถูกต้องมีค่าเพิ่มขึ้น ดังนั้น จุดศูนย์กลางของคลัสเตอร์ที่ 1 จุดใหม่ คือ  $x=0.4196$  และ  $y = 0.3482$

7.2 สุ่มค่าระยะทางให้คลัสเตอร์ที่ 2 ซึ่งมีค่าจุดศูนย์กลางเดิม คือ  $x = 0.4623$  และ  $y = 0.7407$  คลาส 2 ได้ค่าดังตารางที่ 3.4

ตารางที่ 3.4 แสดงการสุ่มระยะทางให้กับจุดศูนย์กลางของคลัสเตอร์ที่ 2

ลำดับการสุ่ม	ตัวเลขที่สุ่มได้		ค่าจุดศูนย์กลางใหม่		ค่าความถูกต้อง	
	x	y	x	y	เดิม	ใหม่
1	0.0494	-0.0685	0.4730	0.6722	6	5
2	0.0566	-0.0730	0.4802	0.6677	6	5
3	-0.0673	0.0969	0.3950	0.8376	6	8

จากการสุ่มครั้งที่ 3 พบว่าจุดศูนย์กลางใหม่ที่ได้ทำให้ความถูกต้องมีค่าเพิ่มขึ้น ดังนั้น จุดศูนย์กลางของคลัสเตอร์ที่ 2 จุดใหม่ คือ  $x = 0.3950$  และ  $y = 0.8376$



รูปที่ 3.4 แสดงจุดศูนย์กลางของคลัสเตอร์ 1 และ 2 หลังการสุ่มระยะทาง

## บทที่ 4

### การทดสอบการจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้าน

### พฤติกรรมทางสังคมของมนุษย์

ในบทนี้จะกล่าวถึงค่าพารามิเตอร์ที่ใช้ในการทดลองและผลที่ได้จากการทดลอง โดยข้อมูลที่นำมาใช้ในการทดลองมี 2 ชุด คือ 1) ข้อมูลมาตรฐานซึ่งประกอบด้วย 5 ชุดข้อมูล คือ Iris, Wine, Ionosphere, Haberman และ Heart-Statlog 2) ข้อมูลที่ผู้วิจัยสร้างขึ้นเองซึ่งประกอบด้วย 4 ชุดข้อมูล คือ Flower1, Flower2, Sawtooth และ Fan ในการทดลองนี้จะทำการเปรียบเทียบค่าเปอร์เซ็นต์ความถูกต้องและจำนวนคลัสเตอร์ที่ถูกสร้าง โดยทำการเปรียบเทียบกับ Fuzzy ARTMAP, Learning Vector Quantization (LVQ) และ Radial Basis Function (RBF) ซึ่งเป็นโมเดลทางด้านการจัดกลุ่มข้อมูลที่ได้รับการยอมรับและถูกนำไปประยุกต์ใช้ในหลายๆ ด้าน และในที่นี่ได้ทำการทดลองเกี่ยวกับงานวิจัยที่นำเสนออีก 2 กรณี คือ 1) งานวิจัยที่นำเสนอโดยไม่มีการสุ่มค่าระยะทางให้กับคลัสเตอร์ และ 2) งานวิจัยที่นำเสนอโดยทำการสุ่มค่าระยะทางให้ทุกคลัสเตอร์พร้อมกัน

#### 4.1 ข้อมูลที่ใช้ในการทดลอง

##### 4.1.1 ข้อมูลมาตรฐาน

###### 1) Iris

ข้อมูล Iris ถูกเก็บโดย Anderson ในปี 1935 ชุดของข้อมูลประกอบด้วยข้อมูลที่ได้จากการสุ่มตัวอย่างจำนวน 150 ตัวอย่าง โดยได้จำแนกดอก Iris ออกเป็น 3 ประเภท คือ Iris Setosa, Iris Virginica และ Iris Versicolour โดยแต่ละประเภทประกอบด้วยข้อมูลตัวอย่างจำนวน 50 ตัวอย่าง



iris setosa



Iris versicolor



Iris-virginica

รูปที่ 4.1 แสดงลักษณะของดอก Iris ทั้ง 3 ประเภท

ในชุดข้อมูลดอก Iris แต่ละข้อมูลตัวอย่างจะประกอบด้วย 4 คุณลักษณะ ดังนี้

1. Sepal length คือ ความยาวของกลีบเลี้ยงของดอก
2. Sepal width คือ ความกว้างของกลีบเลี้ยงของดอก
3. Petal length คือ ความยาวของกลีบดอกไม้
4. Petal width คือ ความกว้างของกลีบดอกไม้

## 2) Wine

ข้อมูล Wine ถูกเก็บโดย Stefan Aeberhard สถาบัน Institute of Pharmaceutical and Food Analysis and Technologies ในเดือนกรกฎาคม ปี 1991 โดยข้อมูล Wine เป็นข้อมูลที่ได้จากการวิเคราะห์ทางเคมีของเหล้าไวน์ในประเทศอิตาลีที่อยู่ในบริเวณเดียวกันแต่มาจาก 3 แหล่งการเพาะปลูกที่แตกต่างกัน

ชุดของข้อมูลประกอบด้วยข้อมูลที่ได้จากการสุ่มตัวอย่างจำนวน 178 ตัวอย่าง แต่ละข้อมูลตัวอย่างมีจำนวนคุณลักษณะ 13 คุณลักษณะ ไม่มี Missing Values ชุดของข้อมูลถูกแบ่งออกเป็น 3 ประเภทตามแหล่งการเพาะปลูกที่แตกต่างกัน โดยแต่ละประเภทประกอบด้วยข้อมูลตัวอย่างจำนวน 59, 71 และ 48 ตัวอย่างตามลำดับ

## 3) Ionosphere

ข้อมูล Ionosphere ถูกเก็บโดย Vince Sigillito จาก Applied Physics Laboratory Johns Hopkins University ในปี 1989 โดยทำการสะสมข้อมูลเรดาร์จากระบบใน Goose Bay, Labrador เพื่อต้องการการปลดปล่อยอิเล็กตรอนของชั้นบรรยากาศ

ชุดของข้อมูลประกอบด้วยข้อมูลตัวอย่างจำนวน 351 ตัวอย่าง ไม่มี Missing Values โดยได้จำแนกข้อมูลออกเป็น 2 ประเภท คือ Good (เรดาร์จะตอบกลับมาโดยแสดงให้เห็นโครงสร้างในชั้นบรรยากาศ) และ Bad (เรดาร์ไม่ตอบอะไรกลับมา) ข้อมูลแต่ละประเภทประกอบด้วยข้อมูลตัวอย่างจำนวน 225 และ 126 ตัวอย่างตามลำดับ มีจำนวนคุณลักษณะทั้งสิ้น 34 คุณลักษณะ ที่มีค่าเป็น continuous

## 4) Haberman

Haberman เป็นข้อมูลกรณีศึกษาของการอยู่รอดของคนไข้ที่เข้ารับการผ่าตัดด้วยโรคมะเร็งทรวงอกระหว่างปี 1958 – 1970 ของ University of Chicago's Billings Hospital ข้อมูลถูกเก็บโดย Tjen-Sien Lim ในปี 1999 ชุดข้อมูลประกอบด้วย 306 รูปแบบอินพุต โดยจำแนกข้อมูลออกเป็น 2 ประเภทตามสถานะคนไข้ คือ 1) คนไข้มีชีวิตอยู่เกิน 5 ปี และ 2) คนไข้ตายภายใน 5 ปี

ข้อมูลถูกแบ่งเป็นข้อมูลเรียนรู้ 214 รูปแบบอินพุต และข้อมูลทดสอบ 92 รูปแบบอินพุต ไม่มี Missing Values แต่ละรูปแบบอินพุตประกอบด้วย 3 คุณลักษณะ ดังนี้

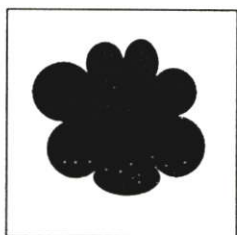
1. Age of patient at time of operation
2. Patient's year of operation (year - 1900)
3. Number of positive auxiliary nodes detected

### 5) Heart Statlog

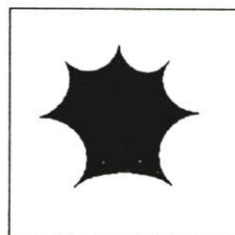
ข้อมูล Heart-statlog ถูกเก็บโดย University of California, San Diego medical Center ข้อมูล Heart-statlog มีจำนวนทั้งสิ้น 270 รูปแบบอินพุต แบ่งข้อมูลออกเป็น 2 ประเภทคือ 1) Absent (ไม่เป็นมะเร็ง) จำนวน 150 ข้อมูล และ 2) Presence (เป็นมะเร็ง) จำนวน 120 ข้อมูล ไม่มี Missing Values แต่ละรูปแบบอินพุตประกอบด้วย 14 คุณลักษณะ

#### 4.1.2 ข้อมูลที่ผู้วิจัยสร้างขึ้นเอง

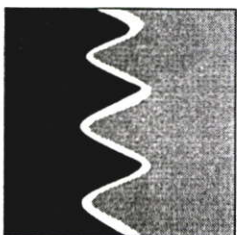
ข้อมูลที่ทำวิจัยสร้างขึ้นเองมีทั้งหมด 4 ชุดข้อมูลดังรูปที่ 4.2 ข้อมูลแต่ละรูปมีขนาด 500 x 500 pixel โดยแต่ละรูปแบบอินพุตได้จากการสุ่มคู่อันดับ (x,y) จากรูปที่สร้างขึ้น โดยมีการสุ่มรูปแบบอินพุตขึ้นมาทั้งหมด 4000 รูปแบบอินพุต และนำชุดข้อมูลที่ได้มาผ่านขั้นตอนการ normalize ให้เป็นเลขจำนวนจริงที่อยู่ในช่วง 0 ถึง 1 โดยข้อมูลจะถูกแบ่งเป็นข้อมูลที่ใช้ในการเรียนรู้โมเดล 2000 รูปแบบอินพุต และข้อมูลที่ใช้ในการทดสอบ โมเดล 2000 รูปแบบอินพุต



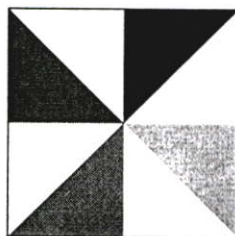
ก) ชุดข้อมูล Flower1



ข) ชุดข้อมูล Flower2



ค) ชุดข้อมูล Sawtooth



ง) ชุดข้อมูล Fan

รูปที่ 4.2 แสดงลักษณะของชุดข้อมูลที่ผู้ทำวิจัยสร้างขึ้นเอง

## 4.2 วิธีทำการทดลอง

แต่ละชุดข้อมูลมาตรฐานที่นำมาใช้ในการทดลองจะถูกแบ่งออกเป็น 2 ส่วน คือ ข้อมูลที่ใช้เรียนรู้โมเดล และข้อมูลที่ใช้ทดสอบโมเดล โดยแต่ละคลาสมีการแบ่งข้อมูลดังนี้

### 4.2.1 Iris

จำนวนข้อมูลทั้งหมด 150 รูปแบบอินพุต โดยถูกแบ่งเป็นข้อมูลที่ใช้ในการเรียนรู้โมเดล 120 รูปแบบอินพุต และข้อมูลที่ใช้ในการทดสอบโมเดล 30 รูปแบบอินพุต แบ่งออกเป็น 3 คลาส คือ Iris Setosa, Iris Virginica และ Iris Versicolour โดยแต่ละคลาสมีรายละเอียดดังตารางที่ 4.1

ตารางที่ 4.1 แสดงการแบ่งข้อมูลของชุดข้อมูล Iris

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1 (Iris Setosa)	40	10
2 (Iris Virginica)	40	10
3 (Iris Versicolour)	40	10

### 4.2.2 Wine

จำนวนข้อมูลทั้งหมด 178 รูปแบบอินพุต โดยถูกแบ่งเป็นข้อมูลที่ใช้ในการเรียนรู้โมเดล 90 รูปแบบอินพุต และข้อมูลที่ใช้ในการทดสอบโมเดล 88 รูปแบบอินพุต แบ่งออกเป็น 3 คลาสตามแหล่งการเพาะปลูก โดยแต่ละคลาสมีรายละเอียดดังตารางที่ 4.2

ตารางที่ 4.2 แสดงการแบ่งข้อมูลของชุดข้อมูล Wine

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1	30	29
2	36	35
3	24	24

### 4.2.3 Ionosphere

จำนวนข้อมูลทั้งหมด 351 รูปแบบอินพุต โดยถูกแบ่งเป็นข้อมูลที่ใช้ในการเรียนรู้โมเดล 200 รูปแบบอินพุต และข้อมูลที่ใช้ในการทดสอบโมเดล 151 รูปแบบอินพุต แบ่งออกเป็น 2 คลาส คือ good และ bad โดยแต่ละคลาสมีรายละเอียดดังตารางที่ 4.3

ตารางที่ 4.3 แสดงการแบ่งข้อมูลของชุดข้อมูล Ionosphere

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1 (good)	101	124
2 (bad)	99	27

#### 4.2.4 Haberman

จำนวนข้อมูลทั้งหมด 306 รูปแบบอินพุต โดยถูกแบ่งเป็นข้อมูลที่ใช้ในการเรียนรู้โมเดล 214 รูปแบบอินพุต และข้อมูลที่ใช้ในการทดสอบโมเดล 92 รูปแบบอินพุต แบ่งออกเป็น 2 คลาส คือ คนไข้มีชีวิตอยู่เกิน 5 ปี และคนไข้ตายภายใน 5 ปี โดยแต่ละคลาสมีรายละเอียดดังตารางที่ 4.4

ตารางที่ 4.4 แสดงการแบ่งข้อมูลของชุดข้อมูล Haberman

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1 (คนไข้มีชีวิตอยู่เกิน 5 ปี)	157	68
2 (คนไข้ตายภายใน 5 ปี)	57	24

#### 4.2.5 Heart-Statlog

จำนวนข้อมูลทั้งหมด 270 รูปแบบอินพุต โดยถูกแบ่งเป็นข้อมูลที่ใช้ในการเรียนรู้โมเดล 135 รูปแบบอินพุต และข้อมูลที่ใช้ในการทดสอบโมเดล 135 รูปแบบอินพุต แบ่งออกเป็น 2 คลาส คือ Absent และ Presence โดยแต่ละคลาสมีรายละเอียดดังตารางที่ 4.5

ตารางที่ 4.5 แสดงการแบ่งข้อมูลของชุดข้อมูล Heart-Statlog

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1 (Absent)	75	75
2 (Presence)	60	60

ชุดข้อมูลที่ผู้ทำวิจัยสร้างขึ้นเองแต่ละชุดจะประกอบด้วยข้อมูลที่ใช้ในการเรียนรู้ 2000 รูปแบบอินพุต และข้อมูลที่ใช้ทดสอบ 2000 รูปแบบอินพุต โดยแต่ละคลาสมีรายละเอียดดังตารางที่ 4.6 และ 4.7

ตารางที่ 4.6 แสดงการแบ่งข้อมูลของชุดข้อมูล Flower1, Flower2 และ Sawtooth

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1	1000	1000
2	1000	1000

ตารางที่ 4.7 แสดงการแบ่งข้อมูลของชุดข้อมูล Fan

คลาส	ข้อมูลที่ใช้ในการเรียนรู้โมเดล	ข้อมูลที่ใช้ในการทดสอบโมเดล
1	500	500
2	500	500
3	500	500
4	500	500

### 4.3 การกำหนดพารามิเตอร์ของระบบ

1. จำนวนรอบที่ใช้ในการเรียนรู้โมเดล 500 รอบ
2. ค่าเปอร์เซ็นต์ความถูกต้องเริ่มต้น ( $percent_{correct}$ ) = 0
3. อัตราความเร็วเริ่มต้นระหว่างคลัสเตอร์  $i$  และ คลัสเตอร์  $j$  ที่ใช้เดินเข้าหากัน ( $\delta_{ij}$ ) = 1
4. อัตราเร็วเริ่มต้นของแต่ละคลัสเตอร์ ( $\beta_i$ ) = 1
5. ความกว้างของ Gaussian ( $\sigma$ ) = 0.1
6. จำนวนรอบที่ใช้ดูเปอร์เซ็นต์ความถูกต้องว่ายังเพิ่มขึ้นอยู่หรือไม่ (MaxRound) = 30
7. ขอบเขตของค่าระยะทางที่จะสุ่มให้กับแต่ละคลัสเตอร์ = -0.1 ถึง 0.1
8. จำนวนครั้งที่สุ่มค่าระยะทางให้แต่ละคลัสเตอร์ = 5
9. ระยะห่างระหว่างจุดศูนย์กลางของ 2 คลัสเตอร์ที่อยู่ในคลาสเดียวกัน (DisConst)
10. ค่าที่ใช้เป็นเกณฑ์ในการลบคลัสเตอร์ (Vigilance :  $\rho$ )

### 4.4 ผลการทดลอง

#### 4.4.1 การทดลองที่ 1

ในงานวิจัยนี้ได้ทำการทดลองโดยเปรียบเทียบกับโมเดล Fuzzy ARTMAP และ LVQ โดยโมเดลที่นำเสนอจะแบ่งการทดลองออกเป็น 3 รูปแบบ คือ

โมเดล 1: งานวิจัยที่นำเสนอโดยทำการสุ่มค่าระยะทางให้กับคลัสเตอร์ทีละคลัสเตอร์

โมเดล 2: งานวิจัยที่นำเสนอโดยไม่มีการสุ่มค่าระยะทางให้กับคลัสเตอร์

โมเดล 3: งานวิจัยที่นำเสนอโดยทำการสุ่มค่าระยะทางให้ทุกคลัสเตอร์พร้อมกัน

ผลที่ได้จากการทดลองจะถูกแสดงใน 4 ลักษณะ ดังนี้

- 1) ค่าเปอร์เซ็นต์ความถูกต้องเฉลี่ย ซึ่งเป็นค่าเฉลี่ยที่ได้จากการทดลองที่มีการกำหนดค่าพารามิเตอร์และชุดข้อมูลที่เหมือนกันทั้งหมด 5 ครั้ง
- 2) ค่าเปอร์เซ็นต์ความถูกต้องต่ำสุดของชุดการทดลองในข้อ 1
- 3) ค่าเปอร์เซ็นต์ความถูกต้องสูงสุดของชุดการทดลองในข้อ 1

4) เวลาที่ใช้ในการเรียนรู้โมเดลของชุดการทดลองในข้อ 1 โดยแสดงในรูปแบบ  
ชั่วโมง:นาท:วินาที

#### 4.4.1.1 ชุดข้อมูลมาตรฐาน

##### 1) ชุดข้อมูล Iris

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.04, DistConst=0.05 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.55, Learning rate = 1 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.8

ตารางที่ 4.8 แสดงผลการทดลองของข้อมูล Iris

โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	100	3.2	100	4	100	3	0:01:00
Model 2	100	5	100	5	100	5	0:00:33
Model 3	99.33	4.6	96.67	3	100	5	0:00:40
Fuzzy ARTMAP	100	4	100	4	100	4	0:00:006
LVQ	100	3	100	3	100	3	0:00:38

##### 2) ชุดข้อมูล Wine

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.05, DistConst=0.15 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.8, Learning rate = 1 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.9

ตารางที่ 4.9 แสดงผลการทดลองของข้อมูล Wine

โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	97.04	9.8	96.59	12	98.86	9	0:03:16
Model 2	95.45	6.60	94.32	7	97.73	5	0:02:12
Model 3	95.91	21.6	90.91	26	98.86	21	0:00:23
Fuzzy ARTMAP	96.36	33	95.45	33	96.59	32	0:00:002
LVQ	93.18	3	93.18	3	93.18	3	0:00:29

### 3) ชุดข้อมูล Ionosphere

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.04, DistConst=0.05 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.9, Learning rate = 0.1 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.10

ตารางที่ 4.10 แสดงผลการทดลองของข้อมูล Ionosphere

โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	94.70	39.6	92.72	39	95.36	37	2:05:47
Model 2	94.70	50.40	90.73	50	97.35	51	3:21:50
Model 3	94.57	34.6	92.72	37	96.03	33	0:21:12
Fuzzy ARTMAP	97.75	130.6	97.35	133	98.01	124	0:00:31
LVQ	67.55	2	67.55	2	67.55	2	0:01:08

### 4) ชุดข้อมูล Haberman

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.01, DistConst=0.02 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.35, Learning rate = 0.9 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.11

ตารางที่ 4.11 แสดงผลการทดลองของข้อมูล Haberman

โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	74.35	16.4	73.91	16	75.00	16	0:30:58
Model 2	75.22	16	73.91	17	76.09	14	0:30:36
Model 3	75.65	6.2	72.83	5	78.26	7	0:03:28
Fuzzy ARTMAP	72.39	9.2	66.57	9	75.00	9	0:00:03
LVQ	76.09	2	76.09	2	76.09	2	0:01:03

### 5) ชุดข้อมูล Heart-Statlog

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.07, DistConst=0.2 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.8, Learning rate = 0.4 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.12

ตารางที่ 4.12 แสดงผลการทดลองของข้อมูล Heart-Statlog

โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	84.89	14.8	80.74	14	88.15	15	0:11:40
Model 2	84.00	13.60	82.22	15	85.93	13	0:09:30
Model 3	84.15	11.8	80.74	14	85.93	11	0:01:58
Fuzzy ARTMAP	82.07	64.6	77.78	61	85.19	67	0:00:07
LVQ	84.44	2	84.44	2	84.44	2	0:00:40

#### 4.4.1.2 ชุดข้อมูลที่ผู้ทำวิจัยสร้างขึ้นเอง

##### 1) ชุดข้อมูล Flower1

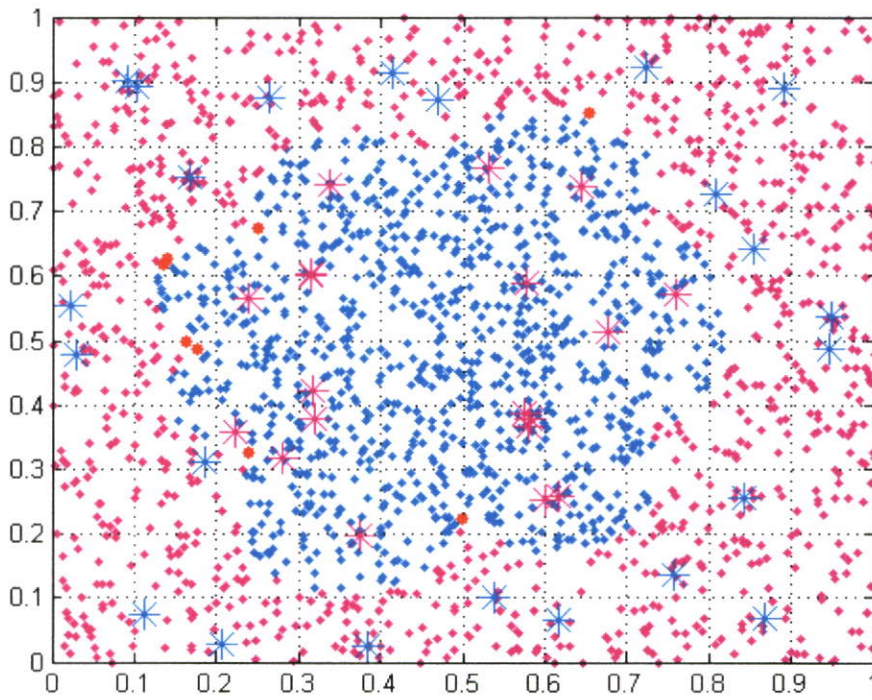
ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.004, DistConst=0.05 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.9, Learning rate = 0.1 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.13

ตารางที่ 4.13 แสดงผลการทดลองของข้อมูล Flower1

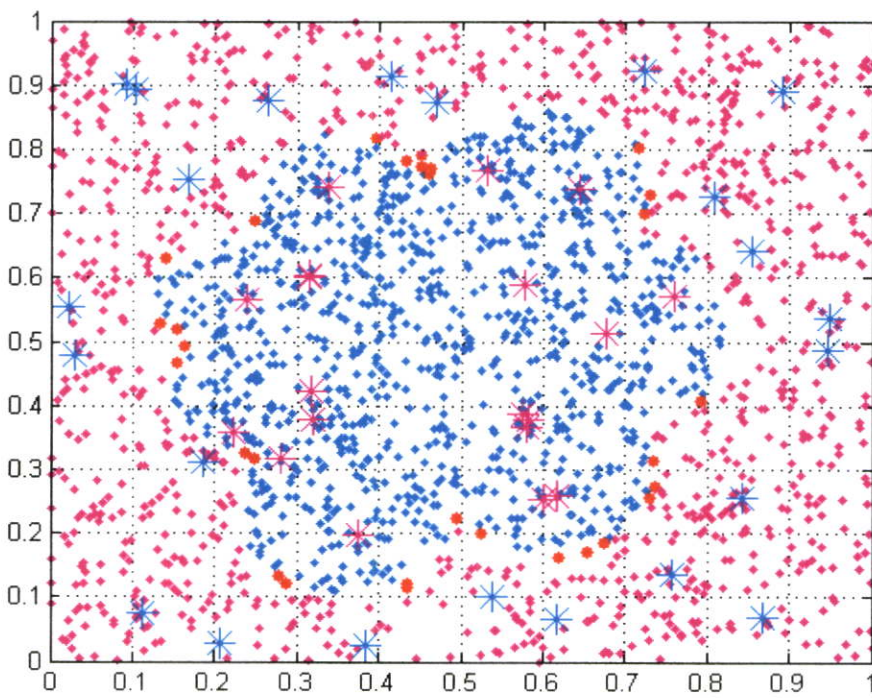
โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	98.44	40.4	98.20	39	98.80	24	12:40:48
Model 2	97.05	15.40	97.05	15	97.2	15	4:56:20
Model 3	96.54	44	95.45	33	97.50	49	4:25:50
Fuzzy ARTMAP	97.69	144.2	97.45	139	97.85	142	0:03:30
LVQ	61.2	2	61.2	2	61.2	2	0:09:48

จากผลการทดลองที่ได้สามารถแสดง Scatter plot ของข้อมูล Flower ได้ดังรูปที่ 4.3 และ 4.4 โดย

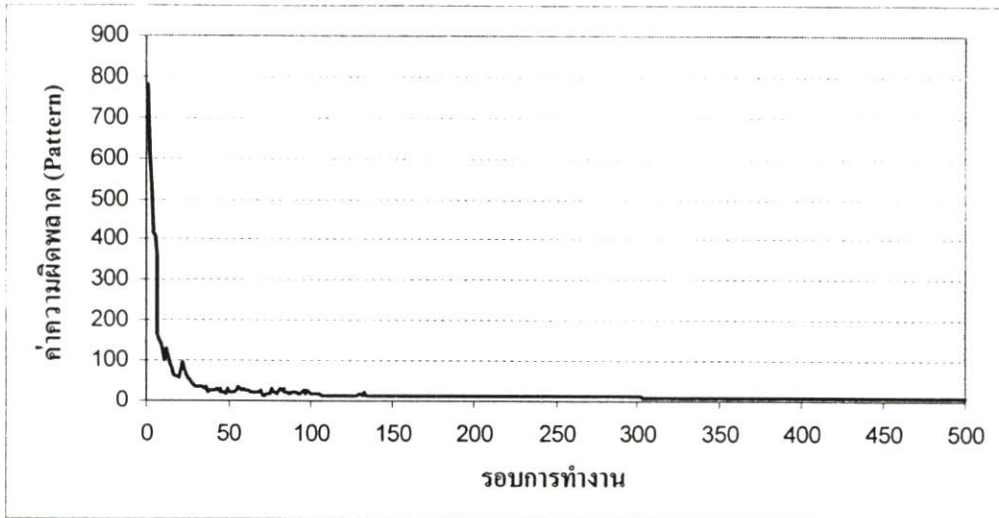
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 1      \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 1
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 2      \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 2
- แสดงถึงรูปแบบอินพุทที่ทำนายผิด



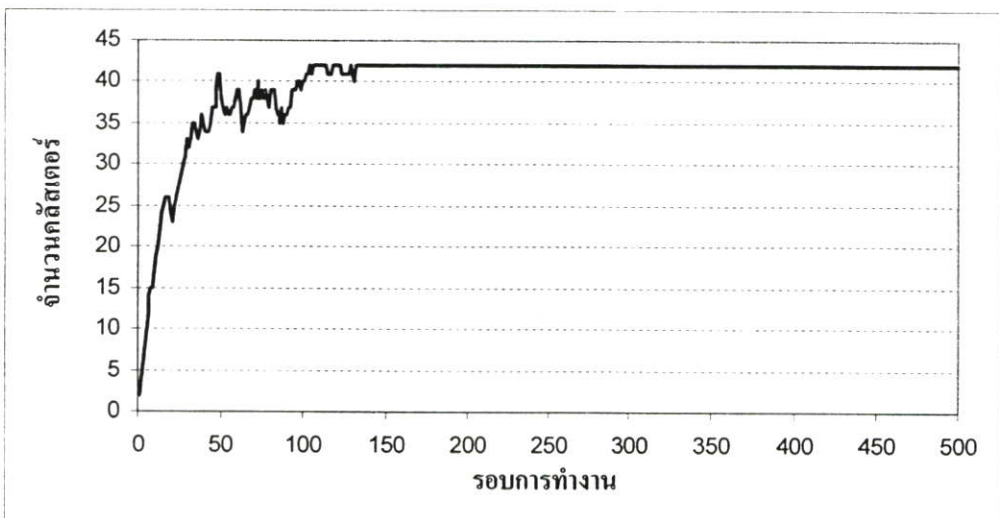
รูปที่ 4.3 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Flower1



รูปที่ 4.4 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบโมเดลของชุดข้อมูล Flower1



รูปที่ 4.5 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงานของชุดข้อมูล Flower1



รูปที่ 4.6 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงานของชุดข้อมูล Flower1

## 2) ชุดข้อมูล Flower2

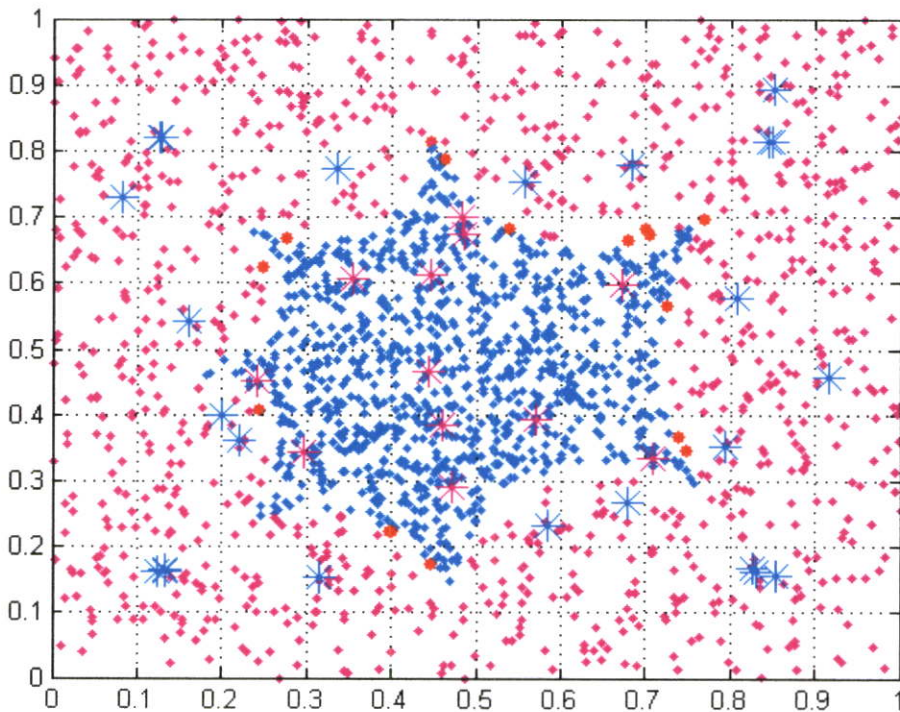
ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.004, DistConst=0.1 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.9, Learning rate = 0.3 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.14

ตารางที่ 4.14 แสดงผลการทดลองของข้อมูล Flower2

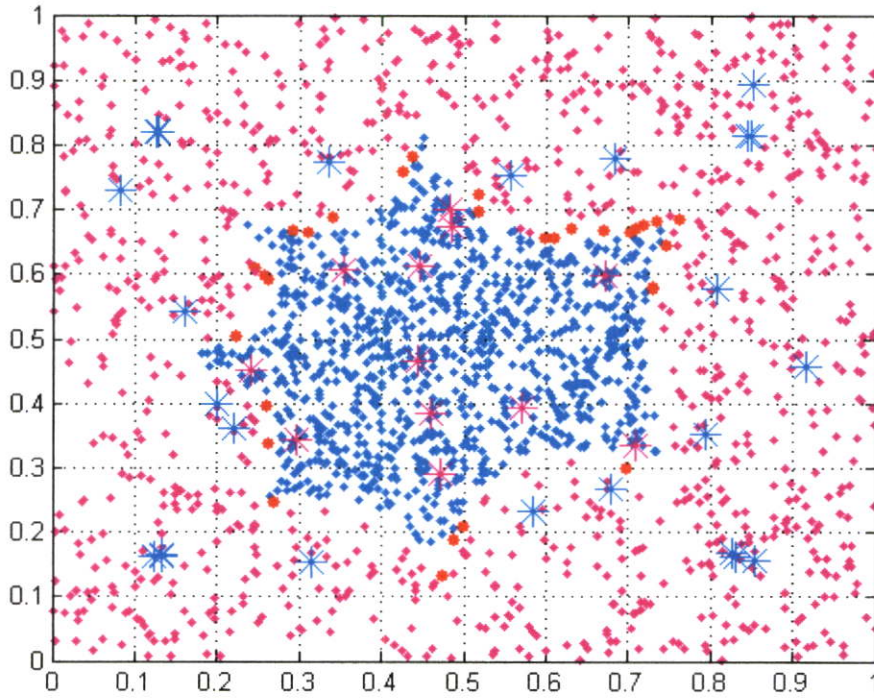
โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	98.06	43.2	97.60	43	98.40	45	12:26:39
Model 2	97.49	16.40	96.55	15	98.05	16	5:26:36
Model 3	95.56	17	94.75	16	96.30	16	1:48:37
Fuzzy ARTMAP	97.83	145.6	97.55	148	98.20	142	0:03:45
LVQ	60.35	2	60.35	2	60.35	2	0:09:47

จากผลการทดลองที่ได้สามารถแสดง Scatter plot ของข้อมูล Flower2 ได้ดังรูปที่ 4.7 และ 4.8 โดย

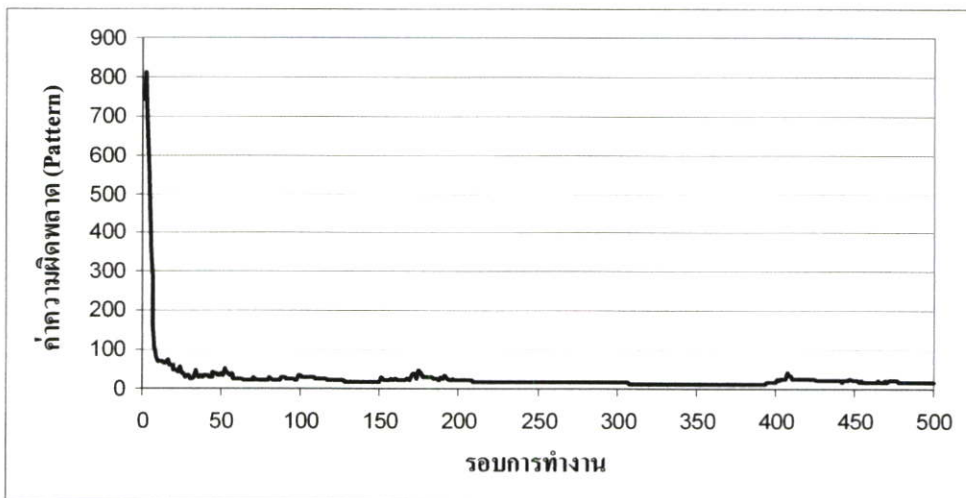
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 1      \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 1
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 2      \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 2
- แสดงถึงรูปแบบอินพุทที่ทำนายผิด



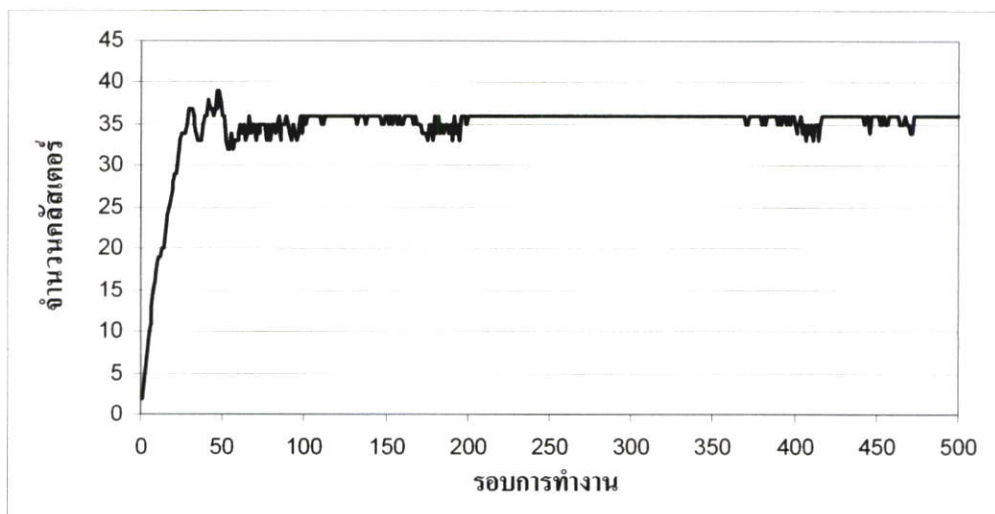
รูปที่ 4.7 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Flower2



รูปที่ 4.8 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบโมเดลของชุดข้อมูล Flower2



รูปที่ 4.9 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงานของชุดข้อมูล Flower2



รูปที่ 4.10 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน  
ของชุดข้อมูล Flower2

### 3) ชุดข้อมูล Sawtooth

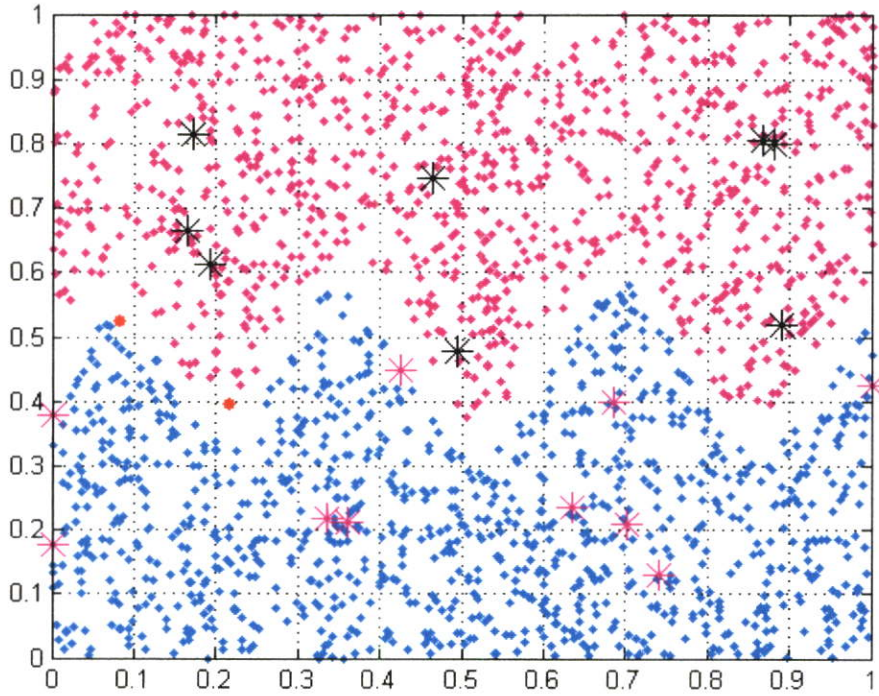
ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.008, DistConst = 0.1 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.9, Learning rate = 0.1 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.15

ตารางที่ 4.15 แสดงผลการทดลองของข้อมูล Sawtooth

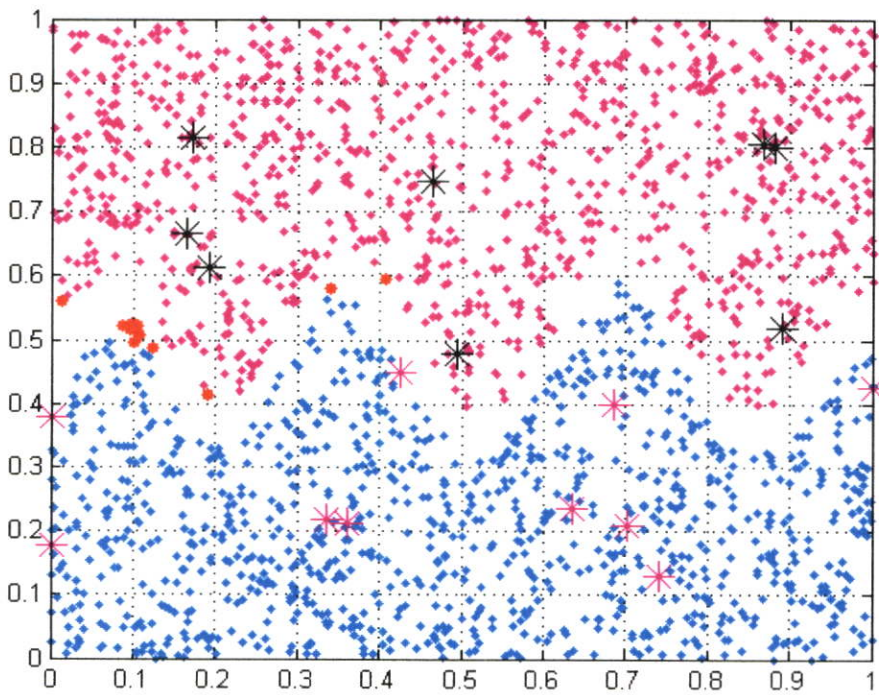
โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	99.25	19.20	99.05	22	99.55	19	3:58:45
Model 2	96.56	15.40	96	15	97.2	15	1:35:05
Model 3	97.67	26.2	96.85	24	98.10	26	3:33:33
Fuzzy ARTMAP	99.76	132.6	99.70	136	99.80	111	0:02:20
LVQ	95.55	2	95.55	2	95.55	2	0:09:48

จากผลการทดลองที่ได้สามารถแสดง Scatter plot ของข้อมูล Sawtooth ได้ดังรูปที่ 4.11 และ 4.12 โดย

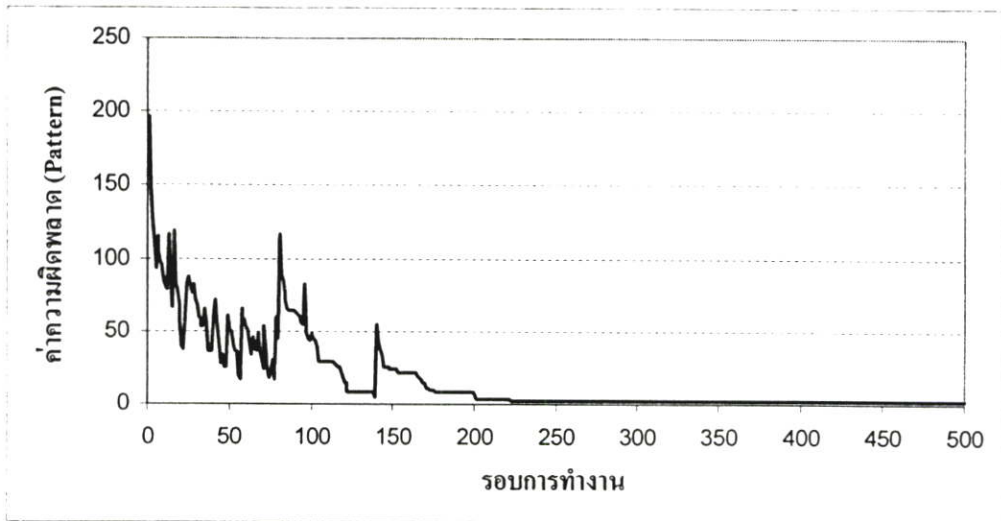
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 1      \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 1
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 2      \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 2
- แสดงถึงรูปแบบอินพุทที่ทำนายผิด



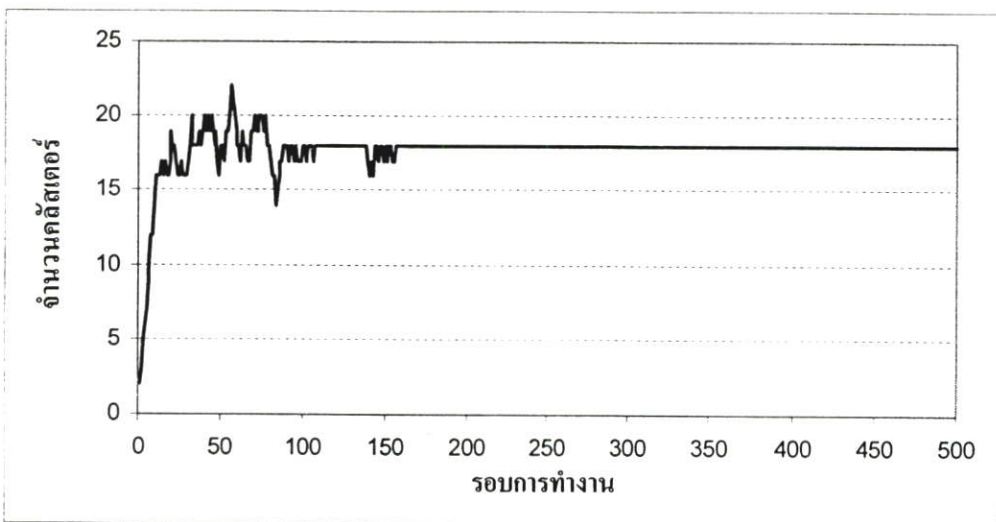
รูปที่ 4.11 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Sawtooth



รูปที่ 4.12 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบโมเดลของชุดข้อมูล Sawtooth



รูปที่ 4.13 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงานของชุดข้อมูล Sawtooth



รูปที่ 4.14 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงานของชุดข้อมูล Sawtooth

#### 4) ชุดข้อมูล Fan

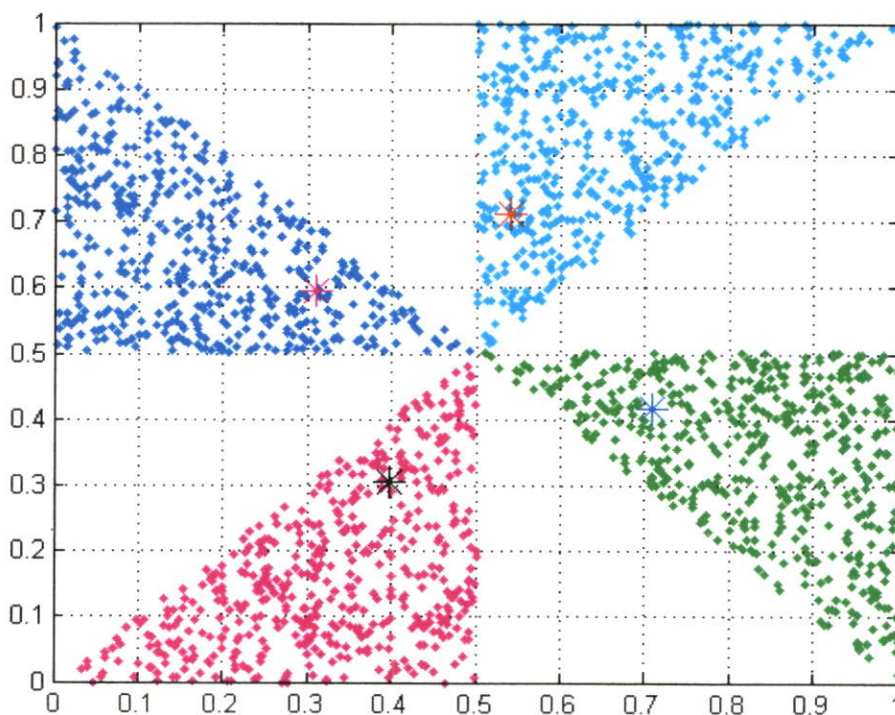
ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.05, DistConst=0.1 ค่าของตัวแปรที่ใช้ใน Fuzzy ARTMAP คือ Vigilance = 0.5, Learning rate = 0.9 และค่าของตัวแปรที่ใช้ใน LVQ คือ Learning rate = 0.001 โดยผลการทดลองแสดงได้ดังตารางที่ 4.16

ตารางที่ 4.16 แสดงผลการทดลองของข้อมูล Fan

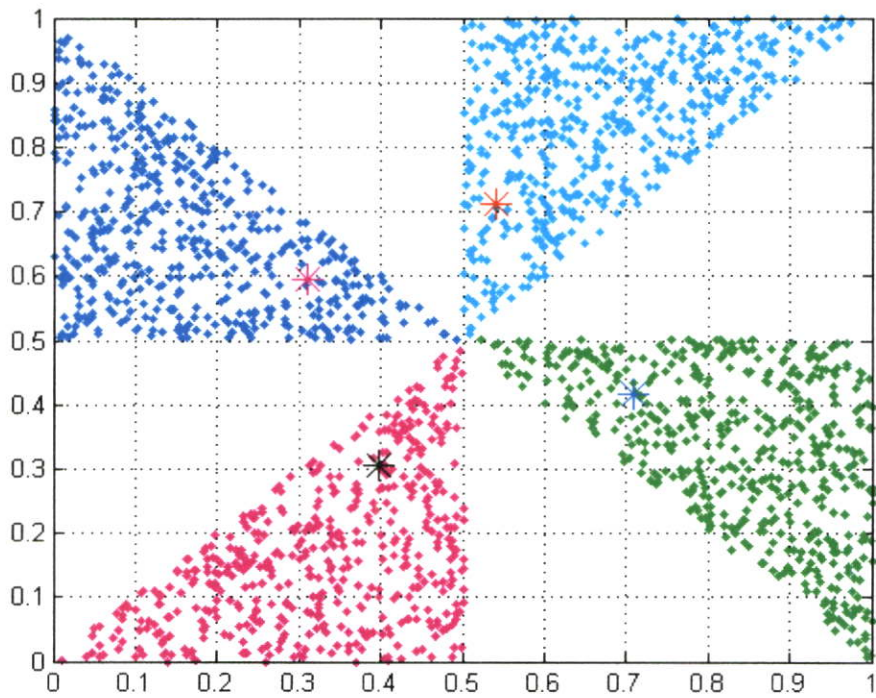
โมเดล	ความถูกต้องเฉลี่ย (%)		ความถูกต้องต่ำสุด (%)		ความถูกต้องสูงสุด (%)		เวลาที่ใช้ในการเรียนรู้
	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	ถูกต้อง	คลัสเตอร์	
Model 1	100	4	100	4	100	4	0:03:12
Model 2	99.98	4	99.90	4	100	4	0:03:46
Model 3	99.99	6	99.95	5	100	5	0:00:40
Fuzzy ARTMAP	100	4	100	4	100	4	0:00:002
LVQ	100	4	100	4	100	4	0:10:47

จากผลการทดลองที่ได้สามารถแสดง Scatter plot ของข้อมูล Fan ได้ดังรูปที่ 4.15 และ 4.16 โดย

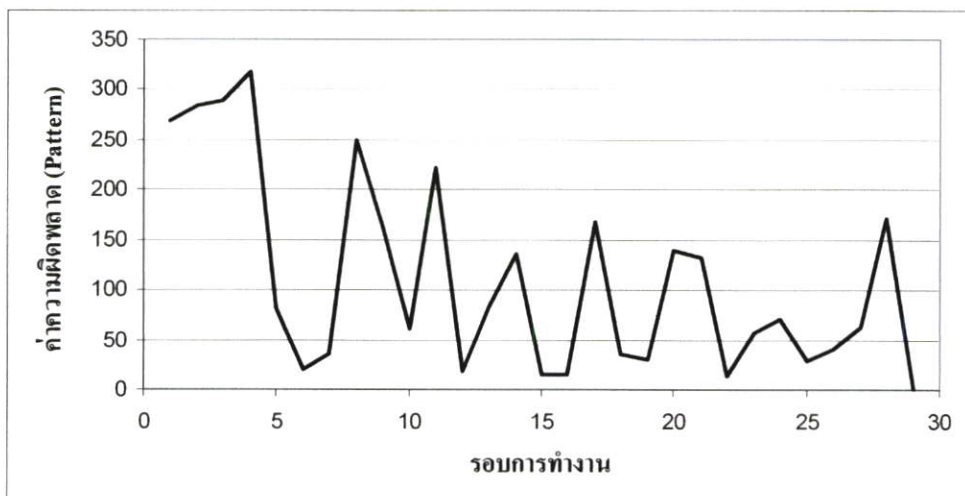
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 1
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 2
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 3
- ◆ แสดงถึงรูปแบบอินพุทในคลาสที่ 4
- แสดงถึงรูปแบบอินพุทที่ทำนายผิด
- \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 1
- \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 2
- \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 3
- \* แสดงถึงจุดศูนย์กลางของคลัสเตอร์ในคลาสที่ 4



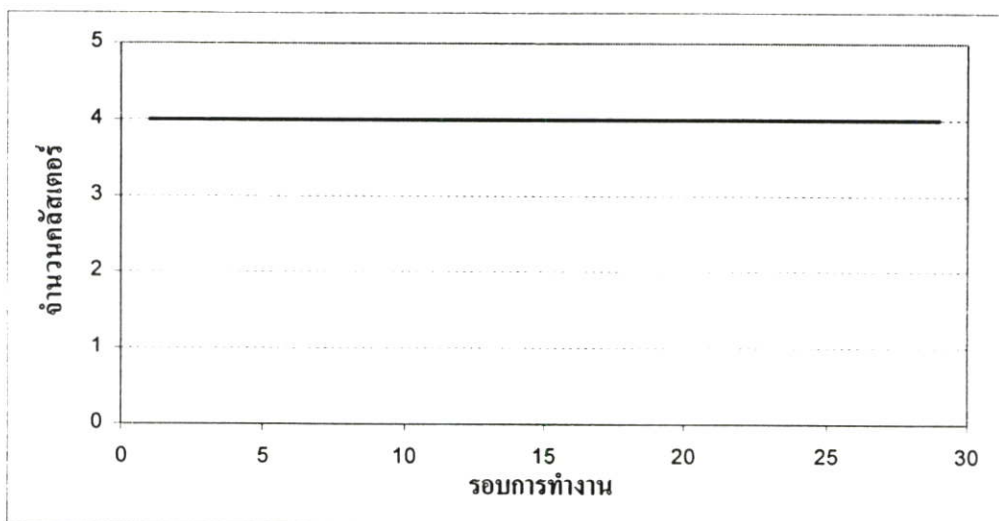
รูปที่ 4.15 แสดง Scatter Plot ของข้อมูลที่ใช้ในการเรียนรู้โมเดลของชุดข้อมูล Fan



รูปที่ 4.16 แสดง Scatter Plot ของข้อมูลที่ใช้ในการทดสอบ โมเดลของชุดข้อมูล Fan



รูปที่ 4.17 กราฟแสดงค่าความผิดพลาดในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงานของชุดข้อมูล Fan



รูปที่ 4.18 กราฟแสดงจำนวนคลัสเตอร์ที่ใช้ในการเรียนรู้เปรียบเทียบกับจำนวนรอบการทำงาน  
ของชุดข้อมูล Fan

#### 4.4.2 การทดลองที่ 2

ในการทดลองนี้ได้ทำการเปรียบเทียบงานวิจัยที่นำเสนอกับโมเดล Radial Basis Function (RBF) ซึ่งมีการกำหนดค่าอัตราการเรียนรู้ของโมเดล RBF เท่ากับ 0.01 โดยในการทดลองนี้จะทำการเปรียบเทียบค่าเปอร์เซ็นต์ความถูกต้องสูงสุด และจำนวนคลัสเตอร์ที่ถูกสร้าง ซึ่งได้ผลการทดลองดังนี้

##### 4.4.2.1 ชุดข้อมูลมาตรฐาน

###### 1) ชุดข้อมูล Iris

ค่าของตัวแปรที่ใช้ในโมเดล 1 คือ Vigilance = 0.04 และ DistConst = 0.05 โมเดล 2 คือ Vigilance = 0.07 และ DistConst = 0.2 และ โมเดล 3 คือ Vigilance = 0.09 และ DistConst = 0.2 โดยผลการทดลองแสดงได้ดังตารางที่ 4.17

ตารางที่ 4.17 แสดงผลการทดลองของข้อมูล Iris เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	100	3	0:01:00
Model 2	100	3	0:00:38
Model 3	100	3	0:00:32
Radial Basis Function	96.61	3	0:00:01

## 2) ชุดข้อมูล Wine

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.005 และ DistConst = 0.15 โมเดล 2 คือ Vigilance = 0.01 และ DistConst = 0.15 และ โมเดล 3 คือ Vigilance = 0.1 และ DistConst = 0.05 โดยผลการทดลองแสดงได้ดังตารางที่ 4.18

ตารางที่ 4.18 แสดงผลการทดลองของข้อมูล Wine เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	100	18	0:12:07
Model 2	100	22	0:00:24
Model 3	98.86	4	0:00:09
Radial Basis Function	96.59	3	0:00:15

## 3) ชุดข้อมูล Ionosphere

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.06 และ DistConst = 0.05 โมเดล 2 คือ Vigilance = 0.09 และ DistConst = 0.15 และ โมเดล 3 คือ Vigilance = 0.05 และ DistConst = 0.2 โดยผลการทดลองแสดงได้ดังตารางที่ 4.19

ตารางที่ 4.19 แสดงผลการทดลองของข้อมูล Ionosphere เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	97.35	28	0:58:01
Model 2	97.35	17	0:28:33
Model 3	96.69	40	0:27:12
Radial Basis Function	98.68	7	0:00:19

## 4) ชุดข้อมูล Haberman

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.1 และ DistConst = 0.1 โมเดล 2 คือ Vigilance = 0.09 และ DistConst = 0.1 และ โมเดล 3 คือ Vigilance = 0.05 และ DistConst = 0.2 โดยผลการทดลองแสดงได้ดังตารางที่ 4.20

ตารางที่ 4.20 แสดงผลการทดลองของข้อมูล Haberman เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	77.17	2	0:01:40
Model 2	78.26	2	0:02:00
Model 3	78.26	3	0:01:55
Radial Basis Function	73.91	2	0:00:10

#### 5) ชุดข้อมูล Heart-Statlog

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.07 และ DistConst = 0.2 โมเดล 2 คือ Vigilance = 0.07 และ DistConst = 0.05 และโมเดล 3 คือ Vigilance = 0.04 และ DistConst = 0.1 โดยผลการทดลองแสดงได้ดังตารางที่ 4.21

ตารางที่ 4.21 แสดงผลการทดลองของข้อมูล Heart-Statlog เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	88.15	15	0:12:40
Model 2	86.67	11	0:07:03
Model 3	87.41	43	0:16:57
Radial Basis Function	87.41	11	0:00:38

#### 4.4.2.2 ชุดข้อมูลที่ผู้วิจัยสร้างขึ้นเอง

##### 1) ชุดข้อมูล Flower1

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.004 และ DistConst = 0.15 โมเดล 2 คือ Vigilance = 0.01 และ DistConst = 0.2 และโมเดล 3 คือ Vigilance = 0.003 และ DistConst = 0.15 โดยผลการทดลองแสดงได้ดังตารางที่ 4.22

ตารางที่ 4.22 แสดงผลการทดลองของข้อมูล Flower1 เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	98.90	38	11:20:47
Model 2	97.25	16	5:10:23
Model 3	97.50	49	5:19:18
Radial Basis Function	62.05	27	0:10:13

### 2) ชุดข้อมูล Flower2

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.004 และ DistConst = 0.05 โมเดล 2 คือ Vigilance = 0.007 และ DistConst = 0.1 และโมเดล 3 คือ Vigilance = 0.009 และ DistConst = 0.2 โดยผลการทดลองแสดงได้ดังตารางที่ 4.23

ตารางที่ 4.23 แสดงผลการทดลองของข้อมูล Flower2 เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	98.70	43	14:05:53
Model 2	98.10	22	9:06:00
Model 3	96.60	16	1:29:26
Radial Basis Function	50	2	0:03:58

### 3) ชุดข้อมูล Sawtooth

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอ คือ Vigilance = 0.001 และ DistConst = 0.15 โมเดล 2 คือ Vigilance = 0.01 และ DistConst = 0.15 และโมเดล 3 คือ Vigilance = 0.006 และ DistConst = 0.05 โดยผลการทดลองแสดงได้ดังตารางที่ 4.24

ตารางที่ 4.24 แสดงผลการทดลองของข้อมูล Sawtooth เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	99.80	23	1:36:26
Model 2	97.20	15	1:25:39
Model 3	98.45	27	3:43:13
Radial Basis Function	92.60	26	0:08:08

#### 4) ชุดข้อมูล Fan

ค่าของตัวแปรที่ใช้ในโมเดลที่นำเสนอคือ Vigilance = 0.05 และ DistConst = 0.1 โมเดล 2 คือ Vigilance = 0.08 และ DistConst = 0.15 และโมเดล 3 คือ Vigilance = 0.07 และ DistConst = 0.2 โดยผลการทดลองแสดงได้ดังตารางที่ 4.25

ตารางที่ 4.25 แสดงผลการทดลองของข้อมูล Fan เมื่อเปรียบเทียบกับ RBF

โมเดล	ความถูกต้อง (%)		เวลาในการเรียนรู้โมเดล
	ความถูกต้อง	คลัสเตอร์	
Model 1	100	4	0:01:46
Model 2	100	4	0:05:32
Model 3	100	4	0:03:43
Radial Basis Function	99.9	20	0:14:04

### 4.5 สรุปผลการทดลอง

#### 4.5.1 สรุปผลการทดลองของชุดข้อมูลมาตรฐาน

1) **ข้อมูล Iris** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าเทียบเท่ากับค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP และ LVQ และสูงกว่าค่าความถูกต้องของ RBF 3.33% และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์เฉลี่ยแล้วจำนวนคลัสเตอร์ของ Fuzzy ARTMAP มีจำนวนมากกว่าโมเดลที่นำเสนอ 1.25 เท่า

2) **ข้อมูล Wine** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าสูงกว่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 0.68% สูงกว่าค่าความถูกต้องเฉลี่ยของ LVQ 3.86% และสูงกว่าค่าความถูกต้องของ RBF 3.41% และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์เฉลี่ยแล้ว Fuzzy ARTMAP มีจำนวนคลัสเตอร์มากกว่าโมเดลที่นำเสนอ 3.37 เท่า

3) **ข้อมูล Ionosphere** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าต่ำกว่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 3.05% สูงกว่าค่าความถูกต้องเฉลี่ยของ LVQ 1.74% และต่ำกว่าค่าความถูกต้องของ RBF 1.33%

4) **ข้อมูล Haberman** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าสูงกว่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 1.74% ต่ำกว่าค่าความถูกต้องเฉลี่ยของ LVQ 1.74% และสูงกว่าค่าความถูกต้องของ RBF 3.26% และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์แล้ว RBF มีจำนวนคลัสเตอร์มากกว่าโมเดลที่นำเสนอ 12 เท่า

5) **ข้อมูล Heart-Statlog** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าสูงกว่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 2.82% สูงกว่าค่าความถูกต้องเฉลี่ยของ LVQ 0.45% และสูง

กว่าค่าความถูกต้องของ RBF 0.74% และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์เฉลี่ยแล้ว Fuzzy ARTMAP มีจำนวนคลัสเตอร์มากกว่าโมเดลที่นำเสนอ 4.36 เท่า

#### 4.5.2 สรุปผลการทดลองของชุดข้อมูลที่ทำวิจัยสร้างขึ้นมาเอง

1) **ข้อมูล Flower1** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าสูงกว่ากับค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 0.75% สูงกว่าค่าความถูกต้องเฉลี่ยของ LVQ 37.24% และสูงกว่าค่าความถูกต้องของ RBF 36.85% และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์เฉลี่ยแล้วจำนวนคลัสเตอร์ของ Fuzzy ARTMAP มีจำนวนมากกว่าโมเดลที่นำเสนอ 3.57 เท่า

2) **ข้อมูล Flower2** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าสูงกว่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 0.23% สูงกว่าค่าความถูกต้องเฉลี่ยของ LVQ 37.71% และสูงกว่าค่าความถูกต้องของ RBF 48.7% และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์เฉลี่ยแล้ว Fuzzy ARTMAP มีจำนวนคลัสเตอร์มากกว่าโมเดลที่นำเสนอ 3.37 เท่า

3) **ข้อมูล Sawtooth** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าต่ำกว่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP 0.51% สูงกว่าค่าความถูกต้องเฉลี่ยของ LVQ 3.7% และสูงกว่าค่าความถูกต้องของ RBF 7.2 % และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์แล้ว RBF มีจำนวนคลัสเตอร์มากกว่าโมเดลที่นำเสนอ 1.13 เท่า

4) **ข้อมูล Fan** ค่าความถูกต้องเฉลี่ยของโมเดลที่นำเสนอมีค่าเทียบเท่าค่าความถูกต้องเฉลี่ยของ Fuzzy ARTMAP และ LVQ และสูงกว่าค่าความถูกต้องของ RBF 0.1 % และเมื่อเปรียบเทียบกับจำนวนคลัสเตอร์แล้ว RBF มีจำนวนคลัสเตอร์มากกว่าโมเดลที่นำเสนอ 5 เท่า

ดังนั้นเมื่อทำการเปรียบเทียบค่าเปอร์เซ็นต์ความถูกต้องพบว่าโมเดลที่นำเสนอมีค่าเปอร์เซ็นต์ความถูกต้องเมื่อเปรียบเทียบกับ Fuzzy ARTMAP, LVQ และ RBF ดังนี้

- สูงกว่า Fuzzy ARTMAP และ LVQ จำนวน 4 ชุดข้อมูล คือ Wine, Heart-Statlog, Flower1 และ Flower2 และสูงกว่า RBF จำนวน 8 ชุดข้อมูล คือ Iris, Wine, Haberman, Heart-Statlog, Flower1, Flower2, Sawtooth และ Fan

- เทียบเท่า Fuzzy ARTMAP และ LVQ จำนวน 2 ชุดข้อมูล คือ Iris และ Fan

- ต่ำกว่า Fuzzy ARTMAP จำนวน 2 ชุดข้อมูล คือ Ionosphere และ Sawtooth ต่ำกว่า LVQ จำนวน 1 ชุดข้อมูล คือ Haberman และต่ำกว่า RBF จำนวน 1 ชุดข้อมูล คือ Ionosphere

เมื่อเปรียบเทียบทางด้านจำนวนคลัสเตอร์จะพบว่าโมเดลที่นำเสนอสร้างคลัสเตอร์น้อยกว่า Fuzzy ARTMAP จำนวน 6 ชุดข้อมูล คือ Wine, Ionosphere, Heart-Statlog, Flower1, Flower2 และ Sawtooth และสร้างคลัสเตอร์น้อยกว่า RBF จำนวน 3 ชุดข้อมูล คือ Haberman, Sawtooth และ Fan

เมื่อทำการเปรียบเทียบผลการทดลองของโมเดลที่นำเสนอทั้ง 3 รูปแบบพบว่าค่าเปอร์เซ็นต์ความถูกต้องเฉลี่ยของโมเดล 1 เมื่อเปรียบเทียบกับโมเดล 2 และ โมเดล 3 เป็นดังนี้

- สูงกว่าโมเดลที่ 2 จำนวน 6 ชุดข้อมูล คือ Wine, Heart-Statlog, Flower1, Flower2, Sawtooth และ Fan และสูงกว่าโมเดลที่ 3 จำนวน 8 ชุดข้อมูล คือ Iris, Wine, Heart-Statlog, Ionosphere, Flower1, Flower2, Sawtooth และ Fan

- เทียบเท่าโมเดลที่ 2 จำนวน 2 ชุดข้อมูล คือ Iris และ Ionosphere และเทียบเท่าโมเดลที่ 3 จำนวน 1 ชุดข้อมูล คือ Fan

- ต่ำกว่าโมเดลที่ 2 จำนวน 1 ชุดข้อมูล คือ Haberman

เมื่อเปรียบเทียบทางด้านจำนวนคลัสเตอร์จะพบว่าโมเดล 1 มีการสร้างคลัสเตอร์น้อยกว่าโมเดล 2 จำนวน 1 ชุดข้อมูล คือ Ionosphere และสร้างคลัสเตอร์น้อยกว่าโมเดล 3 จำนวน 8 ชุดข้อมูล คือ Iris, Wine, Flower1, Sawtooth และ Fan

## บทที่ 5

# สรุปผลการวิจัย และข้อเสนอแนะ

### 5.1 สรุปผลงานวิจัย

งานวิจัยนี้ได้นำเสนอวิธีการจัดกลุ่มข้อมูล (Classification) แบบใหม่ที่ใช้หลักการที่เลียนแบบวิวัฒนาการของสิ่งมีชีวิต (Evolutionary Algorithms) โดยงานวิจัยที่นำเสนอจะพยายามสร้างโมเดลที่เลียนแบบพฤติกรรมทางสังคมของมนุษย์ที่มีการสร้างวงเขตของกลุ่ม (group boundary) ทำให้เกิดกลุ่มวงใน (in-group) และ กลุ่มวงนอก (out-group) โดยสมาชิกจะมีทัศนคติในแง่บวกกับผู้ที่อยู่ในกลุ่มเดียวกัน และจะมีทัศนคติในแง่ลบกับผู้ที่อยู่ต่างกลุ่ม และสมาชิกของกลุ่มวงในจะมีการรักษาระยะห่างทางสังคม (Social distance) กับกลุ่มวงนอกโดยการสร้างความรู้สึกอคติกับกลุ่มอื่น

หลักการทำงานของงานวิจัยที่นำเสนอเริ่มจากทำการสุ่มรูปแบบอินพุตจากแต่ละคลาสมาเป็นจุดศูนย์กลางเริ่มต้น ทำการปรับค่าจุดศูนย์กลางของคลัสเตอร์ ( $center_i$ ) โดยดูจากจุดศูนย์กลางของคลัสเตอร์อื่นๆ ( $center_j$ ) ซึ่งตัวแปรที่มีผลในการปรับค่าจุดศูนย์กลางมี 3 ตัวแปร คือ 1) ค่าอัตราความเร็วระหว่าง 2 คลัสเตอร์ ( $\delta_{ij}$ ) โดยถ้าคลัสเตอร์  $i$  และ  $j$  อยู่ในคลาสเดียวกันและมีค่าระยะห่างต่ำกว่าค่าที่กำหนดไว้เราจะทำการปรับลดค่าอัตราความเร็วที่ 2 คลัสเตอร์มีต่อกันเพื่อให้ถ้าคลัสเตอร์  $i$  เดินเข้าหาคลัสเตอร์  $j$  ด้วยอัตราเร็วที่ช้าลง 2) ค่าความเป็นสมาชิกของจุดศูนย์กลางของ 2 คลัสเตอร์มีต่อกัน ( $\eta_{ij}$ ) โดยถ้าคลัสเตอร์  $j$  อยู่ใกล้จะมีผลต่อการปรับจุดศูนย์กลางของคลัสเตอร์  $i$  มาก แต่ถ้าคลัสเตอร์  $j$  อยู่ไกลจะมีผลต่อการปรับจุดศูนย์กลางของคลัสเตอร์  $i$  น้อย 3) ค่าอัตราเร็วในการเดินของคลัสเตอร์  $i$  ( $\beta_i$ ) หลังจากทีคลัสเตอร์  $i$  ทำการปรับค่าจุดศูนย์กลางแล้วจะทำการเปรียบเทียบค่าความผิดพลาดระหว่างก่อนปรับกับหลังปรับค่าจุดศูนย์กลางของคลัสเตอร์  $i$  โดยถ้าค่าความถูกต้องลดลงเราจะกลับไปใช้จุดศูนย์กลางก่อนปรับค่า ถ้าความถูกต้องเท่าเดิมเราจะปรับลดค่า  $\beta$  เพื่อให้จุดศูนย์กลางของคลัสเตอร์  $i$  เดินด้วยอัตราเร็วที่ช้าลง และถ้าค่าความถูกต้องมากขึ้นเราจะใช้จุดศูนย์กลางใหม่ที่ได้จากการปรับค่า เมื่อทำการปรับค่าจุดศูนย์กลางครบทุกคลัสเตอร์แล้วเราจะทำการตรวจสอบการลบบคลัสเตอร์โดยดูจากขนาดของคลัสเตอร์ จำนวนรูปแบบอินพุตที่เป็นสมาชิก และจำนวนสมาชิกที่ทำนายผิด โดยถ้าค่าที่คำนวณได้มีค่าต่ำกว่าเกณฑ์ที่ตั้งไว้แล้วคลัสเตอร์นั้นๆ จะถูกลบทิ้ง ทำการตรวจสอบค่าความถูกต้องหลังจากการลบบคลัสเตอร์เพื่อนำไปใช้เป็นเงื่อนไขในการสร้างคลัสเตอร์ในรอบถัดไป โดยถ้าค่าความถูกต้องยังเพิ่มขึ้นเรื่อยๆ ในรอบถัดไปเราจะยังคงสร้างคลัสเตอร์เพิ่ม โดยการสุ่มรูปแบบอินพุตจากคลาสที่มีความผิดพลาดเกิดขึ้นมาเป็นจุดศูนย์กลางใหม่ แต่ถ้าค่าความถูกต้องหลังการลบบคลัสเตอร์ไม่เพิ่มขึ้นครบตามจำนวนรอบที่กำหนดไว้เราจะหยุดการสร้างคลัสเตอร์

เพิ่มเติมจะทำการสร้างคลัสเตอร์ทดแทนคลัสเตอร์ที่ถูกลบไปในรอบก่อนหน้า หลังจากนั้นเราจะตรวจสอบแต่ละจุดศูนย์กลางว่าจุดศูนย์กลางใดมีการเปลี่ยนแปลงค่าน้อยมากจุดศูนย์กลางนั้นจะถูกสุ่มค่าระยะทางให้กับจุดศูนย์กลางตามจำนวนครั้งและขอบเขตของค่าที่กำหนดไว้ โดยถ้าการสุ่มค่าระยะทางทำให้ค่าความถูกต้องโดยรวมเพิ่มขึ้นเราจะใช้จุดศูนย์กลางใหม่ที่ได้อันตอนทั้งหมดจะถูกทำซ้ำจนกระทั่งได้ค่าความถูกต้องในการเรียนรู้เป็น 100% หรือครบตามจำนวนรอบที่กำหนดไว้

ข้อมูลที่นำมาใช้ในขั้นตอนการเรียนรู้และขั้นตอนการทดสอบโมเดลนำมาจากฐานข้อมูล UCI จำนวน 5 ชุดข้อมูล และเป็นข้อมูลที่ผู้ทำวิจัยสร้างขึ้นเองจำนวน 4 ชุดข้อมูล โดยข้อมูลที่นำมาใช้เป็นข้อมูลตัวเลขที่นำมาผ่านขั้นตอนการ Normalize ให้เป็นข้อมูลเลขจำนวนจริงที่อยู่ในช่วง 0-1

## 5.2 ข้อดีของงานวิจัย

1. จากข้อมูลทำการทดลองจำนวน 9 ชุดข้อมูล พบว่าโมเดลที่นำเสนอมีการสร้างคลัสเตอร์น้อยกว่า Fuzzy ARTMAP จำนวน 5 ชุดข้อมูล ซึ่งจาก 5 ชุดข้อมูลนี้โมเดลที่นำเสนอมีการสร้างคลัสเตอร์น้อยกว่าคิดเป็นค่าเฉลี่ยเท่ากับ 3.182 เท่า และโมเดลที่นำเสนอมีการสร้างคลัสเตอร์น้อยกว่า RBF จำนวน 2 ชุดข้อมูล ซึ่งจาก 2 ชุดข้อมูลนี้โมเดลที่นำเสนอมีการสร้างคลัสเตอร์น้อยกว่าคิดเป็นค่าเฉลี่ยเท่ากับ 3.06 เท่า

2. จากข้อมูลทำการทดลองจำนวน 9 ชุดข้อมูล พบว่าโมเดลที่นำเสนอมีความถูกต้องในการทำนายสูงกว่า Fuzzy ARTMAP จำนวน 5 ชุดข้อมูล ซึ่งจาก 5 ชุดข้อมูลนี้โมเดลที่นำเสนอมีความถูกต้องในการทำนายสูงกว่าคิดเป็นค่าเฉลี่ยเท่ากับ 1.152%, สูงกว่า LVQ จำนวน 5 ชุดข้อมูล ซึ่งจาก 5 ชุดข้อมูลนี้โมเดลที่นำเสนอมีความถูกต้องในการทำนายสูงกว่าคิดเป็นค่าเฉลี่ยเท่ากับ 6.596% และสูงกว่า RBF จำนวน 8 ชุดข้อมูล ซึ่งจาก 8 ชุดข้อมูลนี้โมเดลที่นำเสนอมีความถูกต้องในการทำนายสูงกว่าคิดเป็นค่าเฉลี่ยเท่ากับ 12.95%

## 5.3 ปัญหาที่พบในงานวิจัย

1. การปรับค่าอัตราเร็วของคลัสเตอร์ ( $\beta_i$ ) จะดูจากค่าความถูกต้องโดยหลังจากที่ทุกคลัสเตอร์ทำการปรับค่าจุดศูนย์กลางแล้วทุกคลัสเตอร์จะต้องทำขั้นตอนในการคำนวณหาค่าความถูกต้อง ซึ่งในขั้นตอนนี้ทำให้โมเดลต้องเสียเวลาในการประมวลผลค่อนข้างมาก

2. พารามิเตอร์ที่ใช้ในการประมวลผลมีจำนวนมาก ดังนั้นจึงยากต่อการหาค่าพารามิเตอร์ที่เหมาะสมสำหรับแต่ละชุดข้อมูล

## 5.4 แนวทางการพัฒนาในอนาคต

1. พัฒนาเงื่อนไขในการปรับค่าอัตราเร็วของคลัสเตอร์ ( $\beta_i$ ) เพื่อปรับปรุงโมเดลในด้านเวลาที่ใช้ในการประมวลผล
2. พัฒนาโมเดลที่ใช้ในการจัดกลุ่มข้อมูล โดยอ้างอิงจากทฤษฎีทางด้านพฤติกรรมทางสังคมของมนุษย์ด้านอื่นๆ ที่สามารถนำมาประยุกต์ใช้ในการจัดกลุ่มข้อมูลได้

## บรรณานุกรม

- [1] Jiawei Han and Micheline Kamper. **Data Mining : concepts and techniques**. San Francisco : Morgan Kaufmann Publishers. 2001.
- [2] David E.Goldberg. **Genetic Algorithms in search, optimization, and machine learning**. United States of America : Addison Wesley Longman, Inc. 1989.
- [3] Zbigniew Michalewicz. **Genetic Algorithms+Data Structures = Evolution Programs**. New York : Springer-Verlag. 1996.
- [4] Randy L.Haupt and Sue Ellen Haupt. **Practical Genetic Algorithms**. United States of America : John Wiley & Sons, Inc. 1998.
- [5] Matthew Settles. 2005. **An Introduction to Particle Swarm Optimization**. [Online]. Available : [www.cs.uidaho.edu/~tsoule/cs472fall05/particleswarm.pdf](http://www.cs.uidaho.edu/~tsoule/cs472fall05/particleswarm.pdf).
- [6] Xiaohui Hu. 2006. **PSO Tutorial**. [Online]. Available : [www.swarmintelligence.org/tutorials.php](http://www.swarmintelligence.org/tutorials.php).
- [7] Falkenauer, E “The worth of the uniform [uniform crossover].” in **Proceeding of IEEE Evolutionary Computation (CEC 99)**,. vol. 1, 1999. pp. 277-282.
- [8] Oguzhan Hasancebia and Fuat Erbatur. “Evaluation of crossover techniques in genetic algorithm based optimum structural design.” in **Computers and Structure**. vol 78, no 1, Nov 2000. pp. 435-448.
- [9] Marilyn B. Brewer. 1999. **The Psychology of Prejudice: Ingroup Love or Outgroup Hate?**. [Online]. Available : [www.findarticles.com/p/articles/mi\\_m0341/is\\_3\\_55/ai\\_58549254](http://www.findarticles.com/p/articles/mi_m0341/is_3_55/ai_58549254).
- [10] Michele G. Alexander and Shana Levin. 1998. **Theoretical, empirical, and practical approaches to intergroup conflict**. [Online]. Available : [www.findarticles.com/p/articles/mi\\_m0341/is\\_4\\_54/ai\\_54422530](http://www.findarticles.com/p/articles/mi_m0341/is_4_54/ai_54422530).
- [11] ภาควิชาสังคมวิทยาและมานุษยวิทยา คณะสังคมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์. **สังคมวิทยา Sociology**. กรุงเทพมหานคร : สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์. 2549.
- [12] Ta-Cheng Chen and Tung-Chou Hsu. “A Gas based approach for mining breast cancer pattern.” in **Expert Systems With Applications**,. vol. 30, no 4, May 2006. pp. 674-681.

- [13] Seyed-Hamid Zahiri and Seyed-Alreza Seyedin. "Swarm intelligence based classifiers." in **Journal of the Franklin Institute.**, 2005. pp. 1-15.
- [14] XiaoWang and HanWang. "Classification by evolutionary ensembles." in **Journal of the Pattern Recognition society.**, vol 39, no 4, Apr 2006. pp. 595-607.
- [15] Wei Gao. "New evolutionary neural network." in **IEEE International Conference on Neural Interface and Control proceedings.**, 2005. pp. 167-171.
- [16] A.Thammano and T.Supin. "Isan Dhamma Character Recognition Using Genetic Algorithm." in **Proceeding of the Third International Conference on Natural Computation (ICNC2007).**, vol.3, 2007. pp. 566-570.

**ภาคผนวก**

### ภาคผนวก ก.

#### ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่

- A. Thammano and J.Moolwong. "Classification Algorithm based on Human Social Behavior." **Proceedings of the 7<sup>th</sup> IEEE International Conference on Computer and Information Technology (CIT2007)**. Fukushima, Japan, 2007.



## Call for Papers and Announcement IEEE 7th International Conference on Computer and Information Technology 2007 (CIT2007)

University of Aizu, Aizu-Wakamatsu, Fukushima, Japan

October 16 to 19, 2007

<http://www.u-aizu.ac.jp/conference/cit07/>

Jointly held with

- 6th International Conference on Bioelectromagnetism (ICBEM2007)
- Workshop on Databases in Networked Information Systems (DNIS2007)
- Eighth International Symposium on Spatial Media (ISSM'07'08)



### Welcome Message from General Chair

You are cordially invited to IEEE CIT2007, the 7th International Conference on Computer and Information Technology, to be held in a beautiful resort of Japan, Oct. 16-19, 2007. CIT has become a major venue for researchers from different fields of computer and information technology. Each year, the CIT attendees appreciate and benefit from the multidisciplinary exchanges in computer and information technology. (Please find CIT publications in the digital library of IEEE computer Society at <http://csdl2.computer.org/persagen/DLPublication.jsp?pubtype=p&acronym=CIT>.) CIT2007 will be held jointly with the official meeting of the International Society for Bioelectromagnetism (<http://www.u-aizu.ac.jp/conference/icbem07/>). This highlights the CIT2007 and gives a unique opportunity, through keynotes and scientific sessions, for computer scientists to exchange information technology in biomedicine. CIT2007 will be held under the most beautiful season in a famous area for its beautiful nature, tradition of Japanese culture and history. We will make every effort to make your stay and participation fruitful and pleasant, and look forward to meeting you in October 2007.

### Scope and Topics

Topics of particular interest include but are

not limited to the following tracks:

- o Databases and Data Mining
- o Internet and Web Applications
- o High Performance Computing
- o Networking and Mobile Computing
- o Data Processing and Multimedia Systems
- o Information Security
- o Embedded Systems
- o Software Engineering
- o Ubiquitous and Sensor Network
- o Service Oriented Computing and E-Technologies
- o Information Technology in Biomedicine
- o VLSI and Design Methodology
- o Computer Graphics and Virtual/Mixed Reality

Sincerely,  
Prof. Daming Wei  
General Chair of CIT 2007

### Important dates:

Submission Due

~~April 27, 2007~~ Extended

o May 16, 2007

Acceptance Notification

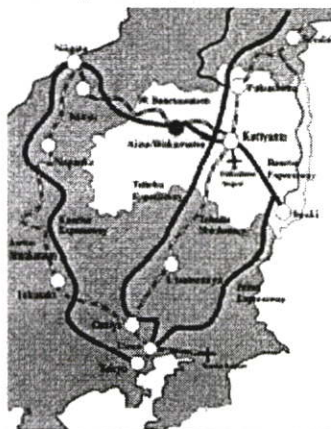
o June 21, 2007

Camera Ready Due

o July 20, 2007

Author Registration

o July 20, 2007



### Special journal issues

International Journal of Pervasive Computing and Communications

E-Commerce and Intelligent Services Computing

Journal of Virtual Reality and Broadcasting

#### Access Guide

##### International transportations

Flight to Tokyo (Nanta airport) then use domestic transportation.  
For attendees from China and Korea, there are direct flights from Shanghai or Seoul

##### Domestic transportations

Express and local trains from Tokyo: about 2.5 hours  
Bus from Fukushima airport: about 1.5 hours

## Classification Algorithm Based on Human Social Behavior

Arit Thammano\* and Jittrapor Moolwong\*\*

*Computational Intelligence Laboratory*

*Faculty of Information Technology*

*King Mongkut's Institute of Technology Ladkrabang,*

*Bangkok, 10520 Thailand*

*E-mail: arit@it.kmitl.ac.th\* and j.moolwong@hotmail.com\*\**

### Abstract

*This paper proposes the new approach to deal with the classification problems by employing the sociological concept of "in-group" and "out-group." The main idea of the concept is about the behavior of in-group members that try to unite with their own group as much as possible, and at the same time maintain social distance from the out-group members. The performance of the proposed model was compared with the fuzzy ARTMAP neural network. The results on five benchmark problems are very encouraging.*

### 1. Introduction

Classification is the process of finding a set of models that describe and distinguish data classes or concepts, for the purpose of being able to use the model to predict the class of objects whose class label is unknown [1]. A variety of techniques have been applied to deal with the classification problems, such as artificial neural networks, a decision tree, and statistical methods. Among a number of classification methods in use, artificial neural networks are the most widely known. Neural networks are a family of naturally inspired models. Many previous research works [2, 3, 4, 5, 6] show that neural network classifiers have better performance, lower classification error rate, and more robust to noise than other classification methods. However, it is not without disadvantage. The disadvantage of the artificial neural networks is that their efficiency depends highly on the user's ability to specify appropriate neural network architecture, e.g. the number of hidden neurons in the backpropagation neural network, the number of basis functions in the radial basis function neural network, and the vigilance parameter in the fuzzy ARTMAP neural network.

Other naturally inspired models such as the genetic algorithm and the particle swarm optimization are also widely used by the research community to solve various types of problems. However, on their own, they are not intended to be used as the classifier. In this paper, therefore, the new algorithm, which is inspired by the human social behavior, is proposed. This proposed algorithm is designed to be a classification algorithm. The performance of the proposed method is evaluated against the fuzzy ARTMAP neural network.

This paper is divided into 4 sections. Following this introduction, section 2 describes the theoretical background of the human social behavior and the learning algorithm of the proposed approach. A brief description of the experimental data and the experimental results are given in section 3. Finally, section 4 is the conclusion.

### 2. Proposed Algorithm

Sociologists have created the concept of "in-group" and "out-group" to help identify the group boundary. Members of the in-group are people who are accepted by the group as belonging to the group, while the out-group members are those whom in-group members consider not belonging to their group. In-group members, even when far away from the group, try to unite with their own group as much as possible, and at the same time maintain social distance from the non-members by creating negative attitude or taking a hostile stance against them. However, among the in-group, each member tends to have their own territory instead of staying tightly close to one another.

According to the above theory of human social behavior, the proposed algorithm is developed. The detailed procedure of the proposed algorithm is as follows:

A. Arbitrarily select one input pattern from the training data set of each class, and assign them to

be the initial cluster centers. At first, there are a total of  $Q$  clusters, which is equal to the number of target output classes.

B. Calculate the accuracy of the model as follows:

$$\text{Accuracy} = \frac{\sum_{i=1}^P A_i}{P} \quad (1)$$

$$A_i = \begin{cases} 1, & \text{if } J \in Y_i \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$J = \arg \min_j (d_j(X_i)) \quad (3)$$

$$d_j(X_i) = \|X_i - z_j\| \quad (4)$$

where  $P$  is the total number of patterns in the training data set.

$J$  is the index of a cluster whose reference pattern is the closest match to the incoming input pattern  $X_i$ .

$Y_i$  is the target output of the  $i^{\text{th}}$  input pattern.

$z_j$  is the center of the  $j^{\text{th}}$  cluster.

$d_j(X_i)$  is the Euclidean distance between the input pattern  $X_i$  and the center of the  $j^{\text{th}}$  cluster.

C. Update the center value of each cluster ( $z_j$ ) by using the following formula:

$$z_{jk}^{\text{new}} = z_{jk}^{\text{old}} + \Delta z_{jk} \quad (5)$$

$$\Delta z_{jk} = \sum_{m \in q} \eta_{jm} \beta_j \delta_{jm} (z_{mk} - z_{jk}) - \sum_{n \notin q} \eta_{jn} \beta_j \delta_{jn} (z_{nk} - z_{jk}) \quad (6)$$

where  $k = 1, 2, 3, \dots, K$ .  $K$  is the number of features in the input pattern.

$q$  is the class to which the  $j^{\text{th}}$  cluster belongs.

$$\eta_{jm} = e^{-\left(\frac{z_{jk} - z_{mk}}{\sigma}\right)^2} \text{ and } \eta_{jn} = e^{-\left(\frac{z_{jk} - z_{nk}}{\sigma}\right)^2}$$

$\beta_j$  is the velocity of the  $j^{\text{th}}$  cluster with respect to its own ability to move in the search space.

$\delta_{jm}$  is the velocity of the  $j^{\text{th}}$  cluster with respect to 2 factors: (1) the distance between the  $j^{\text{th}}$  cluster and the  $m^{\text{th}}$  cluster, and (2) the territorial boundary of the clusters ( $T$ ). If the distance between the  $j^{\text{th}}$  cluster and the  $m^{\text{th}}$  cluster is less than  $T$ , the value of  $\delta_{jm}$  will be decreased by a predefined amount.

After each center is updated, recalculate the accuracy of the model according to equation 1 to 4. If the accuracy is higher, save this new center value and then continue updating the next cluster center. If the accuracy of the model is lower, discard the

new center value and return to the previous center. However, if the accuracy does not change, save the new center value and decrease the value of  $\beta_j$  by a predefined amount. As long as the accuracy remains unchanged, the values of  $\beta_j$  are continuously decreased. In doing this, the algorithm allows the cluster center to move around the search space, however with a slower and slower pace, even if its move does not do any good at the moment.

D. This step is the cluster reduction step. The cluster which satisfies the following equation will be deleted.

$$\left(\frac{n_j}{P}\right) \left(\frac{n_j^q}{n_j}\right) < \rho \quad (7)$$

where  $n_j$  is the number of input patterns in the  $j^{\text{th}}$  cluster.

$n_j^q$  is the number of input patterns in the  $j^{\text{th}}$

cluster whose target outputs ( $Y$ ) are  $q$ .

$q$  is the class to which the  $j^{\text{th}}$  cluster belongs.

$P$  is the total number of patterns in the training data set.

$\rho$  is the vigilance parameter.

E. Recalculate the accuracy of the model according to equation 1 to 4.

F. For each cluster, if the distance between the new center updated in step C and the previous center is less than  $0.0001$  ( $\|z_{jk}^{\text{new}} - z_{jk}^{\text{old}}\| < 0.0001$ ), randomly pick  $K$  small numbers between  $-0.1$  and  $0.1$  and then add them to the center value of the cluster.

$$\begin{bmatrix} z_{j1}^{\text{new}} \\ z_{j2}^{\text{new}} \\ \vdots \\ z_{jk}^{\text{new}} \end{bmatrix} = \begin{bmatrix} z_{j1}^{\text{old}} \\ z_{j2}^{\text{old}} \\ \vdots \\ z_{jk}^{\text{old}} \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_K \end{bmatrix} \quad (8)$$

If the above operation makes the accuracy higher, save this new center value and then continue with the next cluster center. However, if the accuracy does not improve, discard the new center value and repeat this step until the accuracy improves or a predetermined number of iteration is reached.

G. If a predetermined number of iteration is reached or the end condition is satisfied, stop the loop. If not, add new clusters to the model according to the following conditions:

- G.1. If the accuracy of the model improves over the previous iteration, randomly select one input pattern from the training data set of each target output class that still has error, and assign them to be the new clusters. Then go to step B.
- G.2. However, if the accuracy does not improve, randomly select the input patterns, a number equal to the number of clusters deleted in step D, from the training data set of each target output class, and assign them to be the new clusters. Then go to step B.

### 3. Experimental Results

To test the performance, the proposed algorithm was benchmarked against the fuzzy ARTMAP neural network. The experiments were conducted on 5 benchmark data sets:

1. The first data set is the well-known iris data (Figure 1). The iris data [7] has been widely used in the classification problem. The sepal length, sepal width, petal length, and petal width of 150 iris flowers from 3 species (Iris-setosa, Iris-versicolor, and Iris-virginica) are measured in centimeters, and are used as the input of the problem. The training set contains 120 records, while the testing set contains 30 records.
2. The second data set is the Haberman's survival data (Figure 2). It is publicly available from the UCI machine learning database repository [8]. This dataset contains 306 cases from a study that was conducted between 1958 and 1970 at the University of Chicago's Billings Hospital on the survival of patients who had undergone surgery for breast cancer. Three numerical attributes are used to predict the output class (1 or 2). The class value of 1 is corresponded to "the patient survived 5 years or longer" and the value of 2 is corresponded to "the patient died within 5 years." There is no standard split between the training and testing sets. In this paper, the data was divided into a training set of 214 examples and a testing set of 92 examples.
3. The third data set (Figure 3) is the ionosphere data created by the Space Physics Group at Johns Hopkins University. The data is taken from the UCI machine learning database repository [8]. It contains the radar data collected by a system in Goose Bay, Labrador. This data set has 200 instances in the training set and 151 instances in the testing set. Each instance is described by 34 continuous attributes and belongs to one of two classes ("good" or "bad"). "Good" radar returns are those showing evidence of some type of structure in

the ionosphere. "Bad" radar returns are those that do not.

4. The fourth data set (Figure 4) is the heart disease problem. It is retrieved from Statlog Project Datasets [9]. The problem concerns the prediction of the absence (1) or presence (2) of heart disease given the results of various medical tests carried out on a patient. This data set contains 13 attributes and 270 records. There are 150 records of class 1 and 120 records of class 2. In this paper, the 270 records in the database were randomly divided into a training set of 135 records and a testing set of 135 records.
5. The fifth data set is the wine recognition data (Figure 5). It is retrieved from the UCI machine learning database repository [8]. This data is the result of a chemical analysis of wines grown in the same region but from three different cultivars. Thirteen continuous attributes are used to determine the type of wine (class 1, 2, or 3). There are 59 instances of class 1, 71 instances of class 2, and 48 instances of class 3. In this paper, the 178 instances in the database were randomly divided into a training set of 90 instances and a testing set of 88 instances.

5.1	3.5	1.4	0.2	Iris-setosa
4.9	3.0	1.4	0.2	Iris-setosa
7.0	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.3	3.3	6.0	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica

Figure 1. Samples of the iris data

33,58,10,1  
33,60,0,1  
34,59,0,2  
34,66,9,2

Figure 2. Samples of the Haberman's survival data

1,0,0.99539,-0.05889,0.85243,0.02306,0.83398,  
-0.37708,1,0.03760,0.85243,-0.17755,0.59755,  
-0.44945,0.60536,-0.38223,0.84356,-0.38542,  
0.58212,-0.32192,0.56971,-0.29674,0.36946,-0.47357,  
0.56811,-0.51171,0.41078,-0.46168,0.21266,-0.34090,  
0.42267,-0.54487,0.18641,-0.45300,g  
1,0,1,-0.18829,0.93035,-0.36156,-0.10868,-0.93597,1,  
-0.04549,0.50874,-0.67743,0.34432,-0.69707,  
-0.51685,-0.97515,0.05499,-0.62237,0.33109,-1,  
-0.13151,-0.45300,-0.18056,-0.35734,-0.20332,  
-0.26569,-0.20468,-0.18401,-0.19040,-0.11593,  
-0.16626,-0.06288,-0.13738,-0.02447,b

1,0,1,-0.45161,1,1,0.71216,-1,0,0,0,0,0,-1,0.14516,  
0.54094,-0.39330,-1,-0.54467,-0.69975,1,0,0,1,  
0.90695,0.51613,1,1,-0.20099,0.25682,1,-0.32382,1,b

Figure 3. Samples of the Johns Hopkins university ionosphere database

70.0,1.0,4.0,130.0,322.0,0.0,2.0,109.0,0.0,2.4,2.0,3.0,  
3.0,2  
67.0,0.0,3.0,115.0,564.0,0.0,2.0,160.0,0.0,1.6,2.0,0.0,  
7.0,1  
57.0,1.0,2.0,124.0,261.0,0.0,0.0,141.0,0.0,0.3,1.0,0.0,  
7.0,2  
64.0,1.0,4.0,128.0,263.0,0.0,0.0,105.0,1.0,0.2,2.0,1.0,  
7.0,1  
74.0,0.0,2.0,120.0,269.0,0.0,2.0,121.0,1.0,0.2,1.0,1.0,  
3.0,1

Figure 4. Samples of the heart disease problem

14.23,1.71,2.43,15.6,127,2.8,3.06,0.28,2.29,5.64,1.04,  
3.92,1065,1  
12.37,0.94,1.36,10.6,88,1.98,0.57,0.28,0.42,1.95,1.05,  
1.82,520,2  
12.86,1.35,2.32,18,122,1.51,1.25,0.21,0.94,4.1,0.76,  
1.29,630,3

Figure 5. Samples of the wine recognition data

The performance is measured in terms of the classification accuracy and the number of clusters created. In each experiment, the parameters of both models –  $\rho_a$  and  $\beta$  for the fuzzy ARTMAP, and  $\rho$  and  $T$  for the proposed model – are varied throughout their ranges in order to get the best out of the two models. Results of the experiments are shown in Table 1. Considering the number of clusters created during the training process, the proposed model creates much fewer clusters than the fuzzy ARTMAP while having equal or better classification accuracy.

For the iris data (data set 1), both models – the fuzzy ARTMAP neural network and the proposed model – give a perfect accuracy (100%). These best prediction performance are obtained from the fuzzy ARTMAP neural network using the vigilance parameter ( $\rho_a$ ) = 0.5 and the learning rate ( $\beta$ ) = 0.1, and from the proposed model using  $\rho$  = 0.06 and  $T$  = 0.1.

For the data set 2, the best performance of the fuzzy ARTMAP (75%) is obtained when the vigilance parameter is equal to 0.65 and  $\beta$  = 0.4, while the best performance of the proposed model (77.17%) is obtained when  $\rho$  = 0.03 and  $T$  = 0.2.

For the data set 3, the fuzzy ARTMAP produces 2 misclassifications out of a total of 151 examples; this corresponds to an accuracy of 98.68%. The vigilance

parameter and the learning rate are set at 0.8 and 0.6 respectively. On the other hand, the proposed model with  $\rho$  = 0.005 and  $T$  = 0.1 yields accuracy of 96.03%.

For the data set 4, the proposed model with  $\rho$  = 0.05 and  $T$  = 0.2 marginally outperforms the fuzzy ARTMAP with  $\rho_a$  = 0.5 and  $\beta$  = 0.5 (85.93% versus 84.44%).

For the data set 5, the two models obtain the same accuracy of 97.73%. The vigilance parameter and the learning rate of the fuzzy ARTMAP are set at 0.7 and 1 respectively, while  $\rho$  and  $T$  of the proposed model are set at 0.18 and 0.15 respectively.

Table 1. Experimental results

Data set	Fuzzy ARTMAP		Proposed Algorithm	
	Correct (%)	Cluster	Correct (%)	Cluster
Iris	100	4	100	5
Haberman	75	15	77.17	12
Ionosphere	98.68	94	96.03	48
Heart	84.44	35	85.93	9
Wine	97.73	18	97.73	3

#### 4. Conclusion

An approach which employs the concept of the human social behavior is proposed and its performance is compared with the Fuzzy ARTMAP neural network. In comparison to the fuzzy ARTMAP, the time used for training the proposed algorithm is larger. However, its classification accuracy compared with fuzzy ARTMAP is found to be superior. In addition, the number of reference patterns of the proposed approach is also less.

#### 5. References

- [1] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Academic Press, 2001.
- [2] J. R. Quinlan, "Comparing Connectionist and Symbolic Learning Methods," *Proceedings of a Workshop on Computational Learning Theory and Natural Learning Systems*, Vol. 1, The MIT Press, 1994, pp. 445-456.
- [3] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 1995.
- [4] J. W. Shavlik, R. J. Mooney, and G. G. Towell, "Symbolic and Neural Learning Algorithms: An Experimental Comparison," *Machine Learning*, Vol. 6, No. 2, 1991, pp. 111-143.
- [5] S. Danaher, G. Herries, T. Selige, and M. Mac Suintan, "A Comparison of the Characterisation of Agricultural Land Using Singular Value Decomposition and Neural Networks," *Neurocomputation in Remote Sensing Data Analysis*, Springer, 1997.

- [6] J. Lee, R. C. Weger, S. K. Sengupta, and R. M. Welch, "A Neural Network Approach to Cloud Classification," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 5, 1990, pp. 846-855.
- [7] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems" *Annual Eugenics*, Vol. 7, Part II, 1936, pp. 179-188.
- [8] C. L. Blake and C. J. Merz, *UCI Repository of machine learning databases* [<http://www.ics.uci.edu/~mllearn/MLRepository.html>], University of California, Department of Information and Computer Science, 1998.
- [9] Statlog Project Datasets, Retrieved March 28, 2003, From <http://www.liacc.up.pt/ML/statlog/datasets/heart/>.

## ประวัติผู้เขียน

ชื่อ-นามสกุล	นางสาว จิตรภรณ์ มุลวงศ์
วัน เดือนปีเกิด	28 ตุลาคม พ.ศ.2522 ที่จังหวัดนครราชสีมา
ประวัติการศึกษา	2541-2545 วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยกรุงเทพ
ประสบการณ์การทำงาน	
พ.ศ. 2545-2547	อาจารย์ประจำคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยกรุงเทพ