

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

การจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม

DATA CLASSIFICATION BASED ON GENETIC ALGORITHM



T131423

โดย

ธิดา จุลศักดิ์

THIDA JULASAK

อาจารย์ที่ปรึกษา

รศ.ดร.อาริต ธรรมโน

๘๗
ค ๕๘/๗
๒๕๕๕

เลขหมู่.....
เลขทะเบียน 131423
วัน,เดือน,ปี... 2 ส. ค. 2557

b. 12610394
i.

รายงานนี้เป็นส่วนหนึ่งของวิชาการศึกษาระดับ 2

หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 2 ปีการศึกษา 2555

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

DATA CLASSIFICATION BASED ON GENETIC ALGORITHM



**A REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT OF THE COURSE**

INDEPENDENT STUDY2

MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
2/2012



COPYRIGHT 2013

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อ	การจำแนกประเภทของข้อมูล โดยใช้เจเนติกอัลกอริทึม
นักศึกษา	นางสาวธิดา จุลศักดิ์
รหัสนักศึกษา	53660510
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	เทคโนโลยีระบบสารสนเทศ
ปีการศึกษา	2555
อาจารย์ที่ปรึกษา	รศ.ดร.อาริต ธรรมโน

บทคัดย่อ

รายงานฉบับนี้นำเสนอวิธีการจำแนกประเภทของข้อมูล โดยใช้เจเนติกอัลกอริทึม และได้นำหลักการในการปรับปรุงคุณภาพของภาพ โดยการกำจัดสิ่งรบกวนในภาพในโดเมนตำแหน่ง ด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน มาประยุกต์ใช้ในการปรับปรุงชุดข้อมูลเรียนรู้ ในขั้นตอนการเรียนรู้ ของการจำแนกประเภทข้อมูล เพื่อให้สามารถสร้างโครง โมโซมต้นแบบที่ดี แล้วจะส่งผลให้มีความสามารถในการจำแนกประเภทข้อมูลได้มีประสิทธิภาพมากขึ้น ซึ่งในวิทยานิพนธ์เล่มนี้ ได้แสดงผลของการจำแนกประเภทของข้อมูลแบบที่ใช้ชุดข้อมูลเรียนรู้ปกติ เพื่อนำมาเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลแบบที่ใช้ชุดข้อมูลเรียนรู้ที่ผ่านกระบวนการปรับปรุงคุณภาพของภาพ โดยการกำจัดสิ่งรบกวนในภาพในโดเมนตำแหน่ง ด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

Title	Data Classification Based on Genetic Algorithm
Student	Ms. Thida Julasak
Student ID	53660510
Degree	Master of Science
Program	Information Technology
Major	Information System Technology
Academic Year	2012
Advisor	Assoc. Prof. Dr. Arit Thammano

ABSTRACT

This report proposes a method for data classification based on genetic algorithm. And apply image enhance in spatial domain using median filter technique on training dataset. After Genetic algorithm learning in training dataset it will provide result in the ability to classify data efficiently. The report show result of classify from original training data set and enhance training data set. And compare efficiency between training data set.

กิตติกรรมประกาศ

รายงานฉบับนี้สำเร็จลงได้ด้วยความรู้ และสนับสนุนจากอาจารย์ที่ปรึกษา รศ.ดร.อา
ริต ธรรมโน ที่ให้คำปรึกษา คำแนะนำต่าง ๆ แก่ข้าพเจ้า ตลอดจนช่วยแก้ไขข้อบกพร่อง และข้อคิด
ที่เป็นประโยชน์อย่างยิ่งต่อโครงการ ข้าพเจ้าขอขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณอาจารย์ทุกท่าน ที่อบรม สั่งสอน ให้ความรู้ ให้คำแนะนำมาตลอดตั้งแต่
อนุบาล ประถมศึกษา มัธยมศึกษา จนถึงระดับอุดมศึกษา

ขอขอบคุณเพื่อนร่วมงาน ที่ธนาคาร เพื่อน ๆ ที่เรียนร่วมกันทั้งที่คณะเทคโนโลยี
สารสนเทศ และคณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่
คอยให้คำปรึกษา คำแนะนำ กำลังใจ และแรงกระตุ้นในการจัดทำรายงาน

เหนือสิ่งอื่นใด ขอขอบพระคุณบิดา มารดา และบุคคลในครอบครัวที่ให้ความสนับสนุน
ความรัก ความห่วงใย และคอยเป็นกำลังใจที่สำคัญที่สุดอย่างดีเสมอมา

สำหรับคุณงามความดี และประโยชน์อันพึงได้มาจาก โครงการนี้ ข้าพเจ้าขอขอบแต่ผู้มี
พระคุณทุกท่าน



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	I
สารบัญ.....	III
สารบัญตาราง.....	V
สารบัญรูป.....	XV
บทที่ 1 บทนำ	
1.1 ความเป็นมาและความสำคัญของปัญหา.....	
1.2 วัตถุประสงค์.....	1
1.3 ขอบเขตของปัญหา.....	1
1.4 ขั้นตอนการศึกษา.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	
2.1 การจำแนกประเภทของข้อมูล (Data Classification).....	3
2.2 เจเนติกอัลกอริทึม (Genetic Algorithm).....	4
2.3 การวัดความแตกต่างสำหรับการจำแนกประเภทของข้อมูล.....	14
2.4 การปรับปรุงคุณภาพของภาพ.....	15
2.5 บทความงานวิจัยที่เกี่ยวข้อง.....	23
บทที่ 3 การจำแนกประเภทของข้อมูล โดยใช้เจเนติกอัลกอริทึม	
3.1 การนำเจเนติกอัลกอริทึมใช้ในการจำแนกประเภทของข้อมูล.....	26
3.2 การนำวิธีการปรับปรุงคุณภาพของภาพมาประยุกต์ใช้กับการจำแนกประเภท ของข้อมูล.....	30
3.3 การทดสอบประสิทธิภาพของโครโมโซมต้นแบบในการจำแนกประเภทของข้อมูล.....	35
บทที่ 4 ผลการทดลอง	
4.1 ตัวแปรที่เกี่ยวข้องในการทดลองการจำแนกประเภทของข้อมูลโดยใช้ เจเนติกอัลกอริทึม.....	36
4.2 การทดลองเพื่อทดสอบการจำแนกประเภทของข้อมูล โดยใช้เจเนติกอัลกอริทึม.....	40
4.3 ผลการทดลองการทดสอบการจำแนกประเภทของข้อมูล โดยใช้เจเนติกอัลกอริทึม.....	42
4.4 ค่าเฉลี่ยของผลการทดลองของแต่ละวิธีทดลอง.....	82

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา IV นี้ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
4.5 วิจารณ์ผลการทดลอง.....	87
บทที่ 5 สรุปผลการทดลอง และข้อเสนอแนะ	
5.1 สรุปผลการทดลอง.....	89
5.2 ข้อเสนอแนะ.....	89
บรรณานุกรม.....	90



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 ตัวอย่างการจัดเรียงโครโมโซมตามค่าความเหมาะสมจากมากไปน้อย	6
2.2 ตัวเลขที่สุ่มได้ และลำดับของโครโมโซมที่ได้จากการคำนวณ	7
2.3 ค่าความน่าจะเป็นของแต่ละโครโมโซม โดยใช้วิธี Rank weighting	8
2.4 ค่าความน่าจะเป็นของแต่ละโครโมโซมโดยใช้วิธี Roulette wheelghting	9
2.5 รายละเอียดการแข่งขันของโครโมโซมที่ได้จากการสุ่ม	10
3.1 ตัวอย่างชุดข้อมูลทดสอบ สำหรับสร้างโครโมโซม	27
3.2 ตัวอย่างชุดข้อมูลเรียนรู้ ที่เตรียมไปปรับปรุงโดยวิธีการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	31
3.3 ตัวอย่างชุดข้อมูลประเภท A ที่เตรียมเข้าสู่การปรับปรุงฯ	32
3.4 ตัวอย่างชุดข้อมูลประเภท B ที่เตรียมเข้าสู่การปรับปรุงฯ	32
3.5 ตัวอย่างชุดข้อมูลที่ผ่านการปรับปรุงฯ และพร้อมเข้าสู่ตารางฐานข้อมูล	35
4.1 สรุปคุณลักษณะที่สำคัญของชุดข้อมูลทดสอบ	39
4.2 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ	42
4.3 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	42
4.4 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ	43
4.5 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	43
4.6 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ	44

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา VI6 ต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.16 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	49
4.17 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	49
4.18 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	50
4.19 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	50
4.20 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	51
4.21 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	51
4.22 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	52
4.23 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	52
4.24 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	53

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.25 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	53
4.26 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	54
4.27 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	54
4.28 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	55
4.29 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	55
4.30 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	56
4.31 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	56
4.32 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	57
4.33 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	57

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.34 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	58
4.35 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	58
4.36 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	59
4.37 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	59
4.38 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	60
4.39 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	60
4.40 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	61
4.41 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	61
4.42 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	62

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.43 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	62
4.44 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	63
4.45 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	63
4.46 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	64
4.47 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	64
4.48 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	65
4.49 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	65
4.50 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	66
4.51 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	66

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.52 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ	67
4.53 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	67
4.54 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ	68
4.55 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	68
4.56 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ	69
4.57 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	69
4.58 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ	70
4.59 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน	70
4.60 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ	71

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.61 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	71
4.62 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	72
4.63 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	72
4.64 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	73
4.65 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	73
4.66 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	74
4.67 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	74
4.68 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	75
4.69 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	75

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.70 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	76
4.71 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	76
4.72 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	77
4.73 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	77
4.74 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	78
4.75 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	78
4.76 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	79
4.77 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่าน การปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	79
4.78 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	80

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
4.79 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	80
4.80 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ.....	81
4.81 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน.....	81
4.82 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ปกติ เพื่อเปรียบเทียบประสิทธิภาพ ในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูล ทั้ง 4 สมการ.....	82
4.83 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ปกติ เพื่อเปรียบเทียบประสิทธิภาพ ในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูล ทั้ง 4 สมการ (ต่อ).....	83
4.84 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงฯ เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการ ที่ใช้ในการจำแนกประเภทของข้อมูลทั้ง 4 สมการ.....	84
4.85 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงฯ เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการ ที่ใช้ในการจำแนกประเภทของข้อมูลทั้ง 4 สมการ (ต่อ).....	85
4.86 ค่าเฉลี่ยผลการทดลอง เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูล ระหว่างชุดข้อมูลเรียนรู้ปกติ และชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงฯ.....	86

สารบัญรูป

รูปที่	หน้า
2.1	โครงสร้างการทำงานของเจเนติกอัลกอริทึม.....5
2.2	วงล้อรูเลทท์ที่ได้โดยใช้วิธี Rank weighting.....8
2.3	วงล้อรูเลทท์ที่ได้โดยใช้วิธี Roulette wheel.....14
2.4	การไขว้เปลี่ยนด้วยวิธี Single – Point Crossover.....11
2.5	การไขว้เปลี่ยนด้วยวิธี 2 – Point Crossover.....12
2.6	การไขว้เปลี่ยนด้วยวิธี Multi – Point Crossover.....12
2.7	การไขว้เปลี่ยนด้วยวิธี Variable to Variable Crossover.....13
2.8	การไขว้เปลี่ยนด้วยวิธี Uniform Crossover.....14
2.9	ลักษณะของยีน และลักษณะของโครโมโซม.....24
3.1	ตัวอย่างจำลองแบบโครโมโซมแบบเลขฐาน 10.....27
3.2	ตัวอย่างแบบจำลองโครโมโซมแบบเลขฐาน 2.....28
3.3	โปรแกรม Matlab ที่ทำการปรับปรุงชุดข้อมูลประเภท A.....33
3.4	โปรแกรม Matlab ที่ทำการปรับปรุงชุดข้อมูลประเภท B.....34
4.1	ขั้นตอนในการดำเนินการทดลอง.....41

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ข้อมูลเป็นสิ่งที่มียู้อย่างมากมายในโลกยุคปัจจุบัน หากมีการนำข้อมูลมาวิเคราะห์ อธิบาย หรือนำเสนอถึงรูปแบบ ที่ซ่อนอยู่ในชุด หรือกลุ่มของข้อมูลเหล่านั้นได้ ย่อมก่อให้เกิดประโยชน์ ต่อการดำเนินงาน การจำแนกประเภทของข้อมูล ก็เป็นส่วนหนึ่งที่จะช่วยสร้างมูลค่าให้กับข้อมูล ซึ่งมีหลายวิธีที่นำมาใช้ในการจำแนกประเภทของข้อมูล หนึ่งในนั้นคือ เจเนติกอัลกอริทึม (Genetic Algorithm) ซึ่งเป็นทฤษฎีที่เกิดขึ้นจากแนวความคิดเรื่อง ทฤษฎีวิวัฒนาการของสิ่งมีชีวิต ที่ชาร์ล ดา วิน ได้นำเสนอไว้ ซึ่งได้อธิบายกระบวนการ ขั้นตอนในการกำเนิดของสิ่งมีชีวิตในโลก ซึ่งมีการ วิวัฒนาการตามธรรมชาติมาเป็นระยะเวลาหลายล้านปีเพื่อให้เกิดสายพันธุ์ที่ดี ทฤษฎีเจเนติก อัลกอริทึม จึงใช้วิธีสร้างตัวแปรให้อยู่ในรูปของโครโมโซม และให้กระบวนการตามทฤษฎี วิวัฒนาการเพื่อให้เกิดโครโมโซมที่ดีที่สุด หรือผลลัพธ์ที่ดีที่สุดออกมา

อย่างไรก็ดี หากมีการพัฒนาการให้เจเนติกอัลกอริทึม สามารถเรียนรู้ชุดข้อมูลตั้งต้นที่มี ประสิทธิภาพ จะสามารถนำมาสร้างให้เกิด โครโมโซมต้นแบบที่ดีที่สุด เพื่อให้สามารถนำมาใช้ ใน การจำแนกประเภทของข้อมูล ได้อย่างแม่นยำมากยิ่งขึ้น

1.2 วัตถุประสงค์

เพื่อศึกษาทฤษฎีเจเนติกอัลกอริทึม ในส่วนที่นำมาใช้ในการจำแนกประเภทของข้อมูล และ นำทฤษฎีอื่นมาประยุกต์ใช้ร่วมกับเจเนติกอัลกอริทึม เพื่อช่วยเพิ่มประสิทธิภาพในการจำแนก ประเภทของข้อมูล

1.3 ขอบเขตของปัญหา

1.3.1 สร้างวิธีการในการจำแนกประเภทของข้อมูล โดยการนำเจเนติกอัลกอริทึม มา ประยุกต์ใช้งานร่วมกับทฤษฎีอื่น

1.3.2 ใช้ชุดข้อมูลสำหรับจำแนกประเภท ที่เป็นตัวเลขเท่านั้น

1.3.3 ใช้การทดลองแบบออฟไลน์

1.4 ขั้นตอนการศึกษา

1.4.1 กำหนดวัตถุประสงค์

1.4.2 กำหนดขอบเขตของการจัดทำโครงการ

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่อนุญาติให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 1.4.3 กำหนดขั้นตอนการศึกษา
- 1.4.4 ศึกษาทฤษฎีเจเนติกอัลกอริทึม
- 1.4.5 ศึกษาทฤษฎีที่คาดว่าจะนำมาใช้งานร่วมกับเจเนติกอัลกอริทึม
- 1.4.6 ออกแบบแบบจำลอง โดยนำเจเนติกอัลกอริทึม มาทำงานร่วมกับทฤษฎีอื่น
- 1.4.7 สร้างแบบจำลอง
- 1.4.8 ทดสอบแบบจำลอง กับชุดข้อมูลลักษณะต่าง ๆ
- 1.4.9 สรุปผลการทดลอง

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1.5.1 ศึกษาและมีความเข้าใจในทฤษฎีของเจเนติกอัลกอริทึม
- 1.5.2 สามารถนำเจเนติกอัลกอริทึม มาใช้ในการจำแนกประเภทของข้อมูลได้
- 1.5.3 สามารถสร้างแบบจำลองในการจำแนกประเภทของข้อมูล โดยใช้เจเนติกอัลกอริทึมมาทำงานร่วมกับทฤษฎีอื่นได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

2.1 การจำแนกประเภทของข้อมูล (Data Classification)

การจำแนกประเภทของข้อมูล เป็นกระบวนการสร้างโมเดลจัดการชุดข้อมูล (Dataset) เพื่อจัดให้ข้อมูลที่มีความคล้ายคลึงกันอยู่ในประเภทหรือกลุ่ม (Class) เดียวกัน หรือจัดข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ เพื่อแสดงให้เห็นความแตกต่างระหว่างกลุ่มของข้อมูลได้ และเพื่อให้สามารถเข้าสู่กระบวนการทำนายชุดข้อมูลใหม่ ๆ ได้ว่าควรจัดอยู่ในกลุ่มข้อมูลใด โดยใช้โมเดลในจำแนกข้อมูลออกเป็นกลุ่มตามที่ได้กำหนดไว้ ซึ่งขึ้นอยู่กับการวิเคราะห์ชุดข้อมูลว่าสามารถแบ่งออกเป็นกลุ่มใดบ้าง

การจำแนกประเภทของข้อมูล จะนำชุดข้อมูลเรียนรู้ (Training dataset) มาทำการสอนให้ระบบเรียนรู้ลักษณะของกลุ่มข้อมูล เช่น จำนวนกลุ่ม ลักษณะโดยส่วนใหญ่ของแต่ละกลุ่ม เป็นต้น ซึ่งผลลัพธ์ที่ได้จากการเรียนรู้ คือ โมเดลต้นแบบในการจัดประเภทข้อมูล และโมเดลนี้ สามารถแทนได้ในหลายรูปแบบ เช่น Classification (IF-THEN) rules, ต้นไม้ตัดสินใจ (Decision Tree), Mathematical formulae หรือระบบโครงข่ายประสาทเทียม (Neural Network) และจะนำชุดข้อมูลอีกส่วนหนึ่งมาทำการทดสอบ (Testing dataset) เพื่อทดสอบค่าความถูกต้องของโมเดล โดยเปรียบเทียบกลุ่มข้อมูล ซึ่งเป็นผลที่ได้จากโมเดลต้นแบบในการจัดประเภทข้อมูล กับกลุ่มข้อมูลที่แท้จริงของชุดข้อมูล และจะมีกระบวนการปรับปรุง โมเดลจนกว่าจะได้ค่าความถูกต้องในระดับที่น่าพอใจ ภายหลังจากที่มีชุดข้อมูลใหม่เข้ามา เราจะนำชุดข้อมูลนั้นผ่าน โมเดลต้นแบบในการจำแนกประเภทของข้อมูล เพื่อนำมาทำนายกลุ่มของข้อมูลได้

ขั้นตอนการจำแนกประเภทของข้อมูล มี 2 ขั้นตอน ดังนี้

2.1.1 การสร้างโมเดลต้นแบบ (Classifier Model)

เป็นการนำชุดข้อมูล ผ่านกระบวนการของอัลกอริทึมการจำแนกประเภท (Classification Algorithm) ซึ่งผลลัพธ์ที่ได้จะอยู่ในรูปของโมเดลการจำแนก เช่น ต้นไม้ตัดสินใจ เป็นต้น จากนั้นจึงจะสามารถนำมาสร้างเป็นกฎต่อไปได้

2.1.2 การใช้โมเดลเพื่อการทำนาย (Prediction)

เป็นการนำโมเดลต้นแบบที่สร้างได้ มาทำการทำนายกลุ่มของข้อมูล เมื่อได้รับชุดข้อมูลใหม่เข้ามา

ชุดข้อมูลใหม่ที่ได้รับมา จะถูกนำเข้าสู่โมเดลการจำแนกประเภท เพื่อวิเคราะห์และจัดกลุ่มข้อมูล ให้กับชุดข้อมูลนั้น ๆ

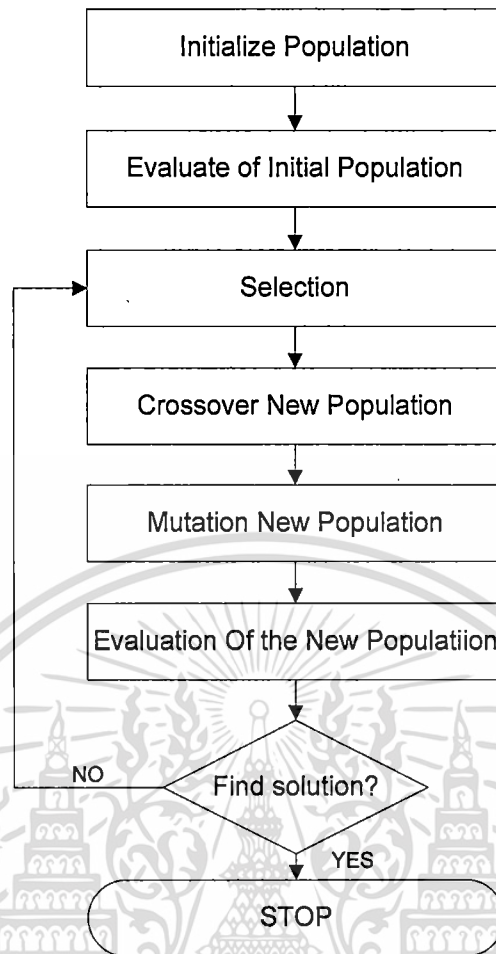
2.2 เจเนติกอัลกอริทึม (Genetic Algorithm)

สิ่งมีชีวิตในธรรมชาตินั้น ก่อกำเนิดขึ้นมาโดยมีชุดของโครโมโซมที่แตกต่างกันตามแต่ละเผ่าพันธุ์ของสิ่งมีชีวิต โดยโครโมโซม (Chromosome) คือ โครงสร้างที่บรรจุสารพันธุกรรมหรือยีนในเซลล์ของสิ่งมีชีวิต ซึ่งยีนจะมีหน้าที่ในการควบคุม และถ่ายทอดลักษณะทางพันธุกรรมของสิ่งมีชีวิตจากพ่อแม่ไปสู่ลูก เช่น ลักษณะของเส้นผม ลักษณะดวงตา เพศ หมูเหือด และผิว เป็นต้น

เจเนติกอัลกอริทึมได้นำแนวความคิดเกี่ยวกับพันธุศาสตร์ทางชีววิทยา และวิวัฒนาการทางธรรมชาติของสิ่งมีชีวิตมาประยุกต์ใช้ในการหาคำตอบให้กับปัญหาโดยนำคำตอบมาสร้างเป็นโครโมโซมซึ่งแต่ละโครโมโซมจะประกอบไปด้วยยีนหลาย ๆ ยีน ที่แสดงถึงคุณลักษณะ (Variable) ของโครโมโซม โดยโครโมโซมที่ดีจะถูกคัดเลือกมาเพื่อเป็นต้นแบบในการสร้างชุดคำตอบใหม่ ๆ ซึ่งเหมือนกับระบบธรรมชาติที่ทำการคัดกรองสิ่งมีชีวิต มีการแพร่พันธุ์ และมีการถ่ายทอดพันธุกรรมที่ดี ตามหลักพันธุศาสตร์ให้กับประชากรรุ่นถัดไป

เจเนติกอัลกอริทึมมีลักษณะการทำงานเหมือนกับ Directed search algorithm หรือ Heuristic algorithm คือ ทำหน้าที่ค้นหาเขตของคำตอบที่เป็นไปได้ โดยเจเนติกอัลกอริทึมจะใช้ฟังก์ชันความเหมาะสม (Fitness Function) เพื่อคัดเลือกโครโมโซมที่มีความเหมาะสมมาเพื่อมาเป็นโครโมโซมพ่อ-แม่ (parents) โครโมโซมพ่อ-แม่ จะถูกทำการไขว้เปลี่ยนคุณลักษณะ (crossover) เพื่อสร้างเป็นโครโมโซมลูก (offspring) และโครโมโซมลูกจะถูกกลายพันธุ์ (mutation) เพื่อให้ได้โครโมโซมที่มีความหลากหลาย หลังจากนั้นจะหาค่าความเหมาะสมของโครโมโซมพ่อ-แม่ และโครโมโซมลูกทั้งหมด เพื่อมาทำการเลือกโครโมโซมชุดที่ดีที่สุดสำหรับเป็นประชากรในรุ่นถัดไป ส่วนโครโมโซมที่ไม่ดีจะถูกตัดทิ้ง

โครโมโซมลูกจะถูกสร้างโดยการจับคู่กันของโครโมโซมพ่อแม่ โดยโครโมโซมพ่อแม่ 1 คู่ จะทำให้เกิดโครโมโซมลูก 2 โครโมโซม จากประชากรทั้งหมดเราจะนำโครโมโซมมาเรียงตามค่า cost function จากน้อยไปมาก โดยโครโมโซมที่มีค่า cost function น้อยจะเป็นโครโมโซมที่ดี ส่วนโครโมโซมที่มีค่า cost function มากจะเป็นโครโมโซมที่ไม่ดี หรือใช้วิธีการเรียงลำดับตามค่าความเหมาะสมจากมากไปน้อย โดยโครโมโซมที่มีค่าความเหมาะสมมากจะเป็นโครโมโซมที่ดี ส่วนโครโมโซมที่ค่าความเหมาะสมน้อยจะเป็นโครโมโซมที่ไม่ดี



รูปที่ 2.1 โครงสร้างการทำงานของเจเนติกอัลกอริทึม

ตัวอย่างเช่น กำหนดให้

จำนวนประชากร = 6

จำนวนบิต = 14

จำนวนยีน = 7

ฟังก์ชันความเหมาะสม = Max: $f(x,y) = x\sin(4x) + 1.1y\sin(2y)$

โครโมโซม คือ 00001101010110

จะได้

$x = 0000110 = 6$

$y = 1010110 = 86$

ค่าความเหมาะสม $f(x,y) = x\sin(4x) + 1.1y\sin(2y)$

$= \text{Max} (6\sin(4*6) + (1.1*86)\sin(2*86))$

$= 1.6057$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จัดเรียงโครโมโซมตามค่าความเหมาะสมจากมากไปน้อย ได้ดังตารางที่ 2.1

ตารางที่ 2.1 ตัวอย่างการจัดเรียงโครโมโซมตามค่าความเหมาะสมจากมากไปน้อย

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

2.2.1 วิธีการเลือกโครโมโซมพ่อแม่ (Selection)

2.2.1.1 Paring of top to bottom

ทำการจัดเรียงโครโมโซมจากโครโมโซมที่ดีที่สุดไปหาโครโมโซมที่แย่ที่สุด และทำการจับคู่โครโมโซมพ่อแม่จากบนลงล่าง

เช่น จากตารางที่ 2.1 สามารถจับคู่ของโครโมโซม ได้ดังนี้

คู่ที่ 1 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 2

คู่ที่ 2 : โครโมโซมที่ 3 คู่กับ โครโมโซมที่ 4

คู่ที่ 1 : โครโมโซมที่ 5 คู่กับ โครโมโซมที่ 6

2.2.1.2 Random paring

ทำการจับคู่โครโมโซมพ่อแม่โดยการสุ่มตัวเลขตามจำนวนประชากรที่กำหนดไว้ เพื่อหาลำดับของโครโมโซม

$$Parent = Roundup \times (N \times Random) \quad (2.1)$$

โดยที่ N คือ จำนวนประชากร

Random คือ ตัวเลขจำนวนจริงในช่วง 0-1 ที่ได้จากการสุ่มค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Roundup คือ ฟังก์ชันที่ใช้ในการปัดเลขจำนวนจริงขึ้นเป็นเลขจำนวนเต็ม

ตารางที่ 2.2 ตัวเลขที่สุ่มได้ และลำดับของโครโมโซมที่ได้จากการคำนวณ

เลขที่สุ่มได้	$6 \times$ เลขที่สุ่มได้	ลำดับของโครโมโซมที่ได้
0.1535	0.9210	1
0.6781	4.0686	5
0.0872	0.5232	1
0.1936	1.1616	2
0.7201	4.2126	5
0.3933	2.3598	3

จากตารางที่ 2.2 จะได้คู่ของโครโมโซม ดังนี้

คู่ที่ 1 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 5

คู่ที่ 2 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 2

คู่ที่ 3 : โครโมโซมที่ 5 คู่กับ โครโมโซมที่ 3

2.2.1.3 Weighted random paring

การหาค่าความน่าจะเป็นให้กับโครโมโซม โดยที่ค่าความน่าจะเป็นขึ้นอยู่กัค่าความเหมาะสมของโครโมโซม โดยโครโมโซมที่มีค่าความเหมาะสมมาก จะมีความน่าจะเป็นที่จะถูกเลือกสูง ส่วนโครโมโซมที่มีค่าความเหมาะสมน้อยจะมีความน่าจะเป็นที่จะถูกเลือกต่ำ บ่อยครั้งที่เราจะเรียกวิธีนี้ว่า Roulette wheel weighting ซึ่งจะนำเสนอ 2 เทคนิค ดังนี้

2.2.1.3.1 Rank weighting การหาค่าความน่าจะเป็นจากลำดับของโครโมโซม เมื่อชุดของโครโมโซมต้นแบบถูกเรียงจากโครโมโซมที่ดีที่สุดไปหาโครโมโซมที่แย่ที่สุด ดังนั้นลำดับของโครโมโซมจึงสามารถบอกได้ว่า ในแต่ละโครโมโซม มีความเหมาะสมที่จะเป็นคำตอบมากน้อยแค่ไหน

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

โดยที่ P_n คือ ความน่าจะเป็นของโครโมโซมลำดับที่ n

N คือ จำนวนประชากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

n คือ ลำดับของโครโมโซม
จากตารางที่ 2.1 ถ้า $N = 6$ แล้วจะได้ว่า

$$P_n = \frac{6-n+1}{1+2+3+4+5+6} = \frac{1-n}{21}$$

ตารางที่ 2.3 ค่าความน่าจะเป็นของแต่ละโครโมโซม โดยใช้วิธี Rank weighting

ลำดับโครโมโซม	โครโมโซม	$P_n = \frac{1-n}{21}$	$\sum_{i=1}^n P_i$
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	01011000011010	$\frac{7-4}{21} = 0.1429$	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 วงล้อรูเลตต์ที่ได้โดยใช้วิธี Rank weighting

เมื่อได้ค่าความน่าจะเป็นของแต่ละโครโมโซมแล้ว ทำการสุ่มตัวเลขทศนิยมที่มีค่าอยู่ในช่วง 0 - 1 ตามจำนวนประชากรที่กำหนดไว้ แล้วดูว่าเลขที่สุ่มได้แต่ละครั้งตกอยู่ในช่วงของโครโมโซมใด ให้ทำการเลือกโครโมโซมนั้นขึ้นมาเป็นโครโมโซมพ่อ - แม่

2.2.1.3.2 Roulette wheel การหาค่าความน่าจะเป็นโดยการนำค่าความเหมาะสมที่คำนวณได้มาหารออกด้วยผลรวมของค่าความเหมาะสมทั้งหมด

เอกสารนี้เป็นเอกสารลิขสิทธิ์ของสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

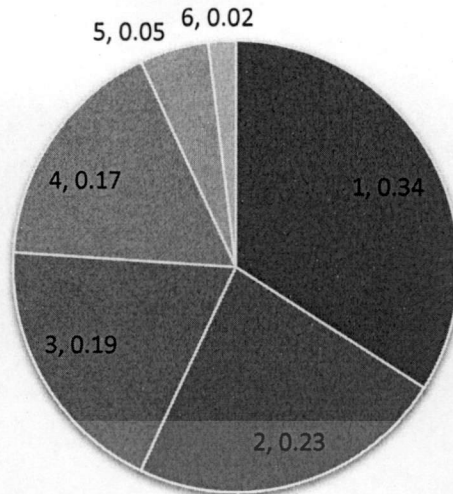
$$P_n = \frac{f_n}{\sum_{p=1}^N f_p} \quad (2.3)$$

โดยที่ P_n	คือ	ความน่าจะเป็นของโครโมโซมลำดับที่ n
f_n	คือ	ค่าความเหมาะสมของโครโมโซมลำดับที่ n
N	คือ	จำนวนประชากร
n	คือ	ลำดับของโครโมโซม

ตารางที่ 2.4 ค่าความน่าจะเป็นของแต่ละโครโมโซมโดยใช้วิธี 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	01011000011010	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	1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.3 วงล้อรูเลตที่ที่ได้โดยใช้วิธี Roulette wheel

เมื่อได้ค่าความน่าจะเป็นของแต่ละโครโมโซมแล้ว ทำการสุ่มตัวเลขทศนิยมที่มีค่าอยู่ในช่วง 0 – 1 ตามจำนวนประชากรที่กำหนดไว้ แล้วดูว่าเลขที่สุ่มได้แต่ละครั้งตกอยู่ในช่วงของโครโมโซมใด ให้ทำการเลือกโครโมโซมนั้นขึ้นมาเป็นโครโมโซมพ่อแม่

2.2.1.3.3 Tournament selection นำคู่ของโครโมโซมมาจัดการแข่งขัน โดยในแต่ละรอบของการแข่งขันจะทำการสุ่มลำดับของโครโมโซมขึ้นมา 2 ตัว และทำการเปรียบเทียบค่าความเหมาะสมของโครโมโซม โดยโครโมโซมที่มีค่าความเหมาะสมมากที่สุดจะเป็นผู้ชนะ

ตารางที่ 2.5 รายละเอียดการแข่งขันของโครโมโซมที่ได้จากการสุ่ม

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

นำโครโมโซมที่ชนะในแต่ละการแข่งขัน มาทำการจับคู่โครโมโซมจากบนลงล่าง ดังนั้นจะได้คู่ของโครโมโซมพ่อแม่ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คู่ที่ 1 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 5

คู่ที่ 2 : โครโมโซมที่ 3 คู่กับ โครโมโซมที่ 4

คู่ที่ 3 : โครโมโซมที่ 1 คู่กับ โครโมโซมที่ 4

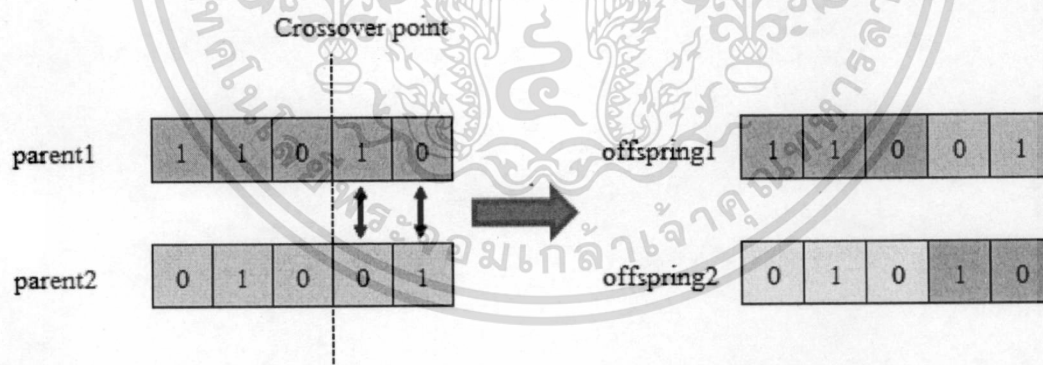
2.2.2 การไขว้เปลี่ยน (Crossover)

เกิดขึ้นภายหลังจากที่ได้คู่ของโครโมโซมพ่อ – แม่ ซึ่งจะนำโครโมโซมพ่อ – แม่มาทำการไขว้เปลี่ยนคุณลักษณะกันระหว่างโครโมโซมพ่อ – แม่ เพื่อสร้างเป็นโครโมโซมลูก โดยจะมีการกำหนดค่าคงที่เพื่อนำมาเป็นค่าความน่าจะเป็นที่จะไขว้เปลี่ยน (P_c) ซึ่งโดยทั่วไปเราจะกำหนด P_c ให้มีค่าสูง ๆ หรือประมาณ 0.9 หลังจากนั้นทำการสุ่มตัวเลขในช่วง 0 – 1 ซึ่งถ้าค่าที่สุ่มได้มีค่าน้อยกว่าค่าความน่าจะเป็นที่ตั้งไว้โครโมโซมนั้นจะถูกไขว้เปลี่ยน

ให้กำหนดตำแหน่งที่จะทำการไขว้เปลี่ยนโดยการสุ่มเลขจำนวนเต็มที่มีค่าอยู่ในช่วง 1 ถึง $m-1$ (m คือ จำนวนยีนในโครโมโซม) จำนวน n ตัว ตามแต่เทคนิคที่เลือกใช้ ในที่นี้จะนำเสนอเทคนิคการไขว้เปลี่ยน 5 เทคนิค ดังนี้

2.2.2.1 Single – Point Crossover

ทำการสุ่มตำแหน่งยีนที่ทำการไขว้เปลี่ยน 1 ตำแหน่ง และสลับค่ายีนพ่อที่อยู่หลังตำแหน่งที่สุ่มได้ทุกยีน

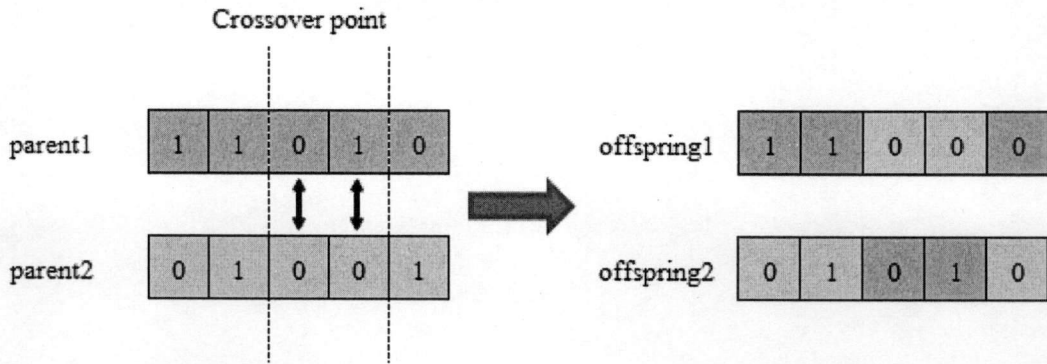


รูปที่ 2.4 การไขว้เปลี่ยนด้วยวิธี Single – Point Crossover

2.2.2.2 2-Point Crossover

สุ่มตำแหน่งยีนที่จะทำการไขว้เปลี่ยน 2 ตำแหน่ง และทำการสลับค่ายีนพ่อแม่ที่อยู่ภายในช่วงตำแหน่งที่สุ่มได้ หรือสลับค่ายีนพ่อแม่ที่อยู่ภายนอกช่วงตำแหน่งที่สุ่มได้

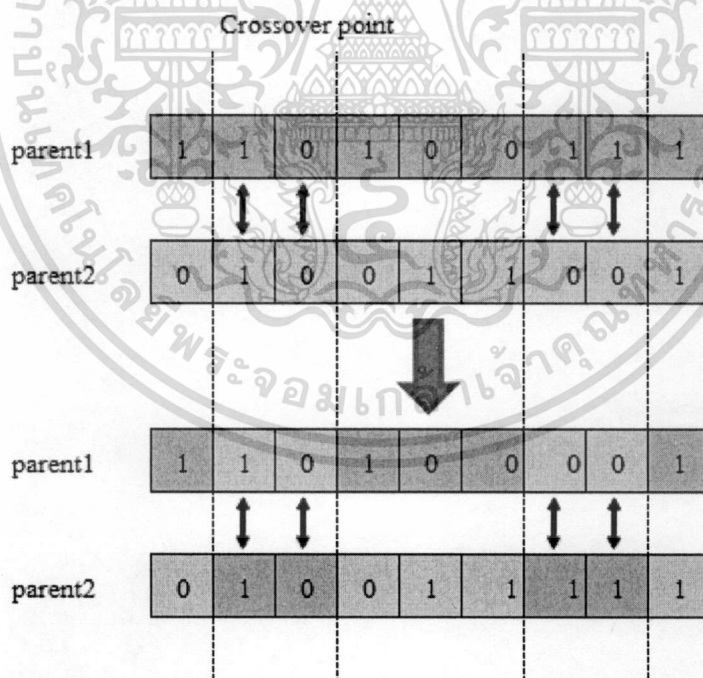
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.5 การไขว้เปลี่ยนด้วยวิธี 2 – Point Crossover

2.2.2.3 Multi – Point Crossover

มีลักษณะการทำคล้ายกับ Single – Point Crossover และ 2 – Point Crossover โดยทำการสุ่มตำแหน่งที่จะไขว้เปลี่ยน ($n_c \geq 3$) โดยจะทำการไขว้เปลี่ยนยีนที่อยู่ระหว่างจุดไขว้เปลี่ยนแต่ละจุด

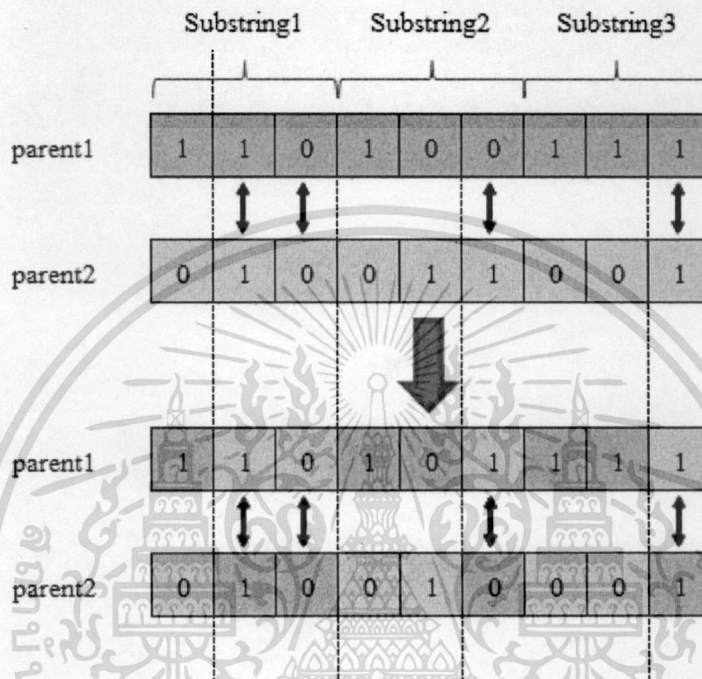


รูปที่ 2.6 การไขว้เปลี่ยนด้วยวิธี Multi – Point Crossover

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2.4 Variable to Variable Crossover

ใช้วิธีการไขว้เปลี่ยน โดยเริ่มจากการแบ่งกลุ่มของข้อมูล หรือกลุ่มของตัวเลข (string) ซึ่งแสดงถึงตัวแปรแต่ละตัวจะถูกแบ่งออกเป็นส่วนย่อย (substring) ในรูปของเลขฐานสอง จากนั้นนำวิธีการไขว้เปลี่ยนแบบ Single – point Crossover ให้กับแต่ละ substring หรือให้กับแต่ละตัวแปร



รูปที่ 2.7 การไขว้เปลี่ยนด้วยวิธี Variable to Variable Crossover

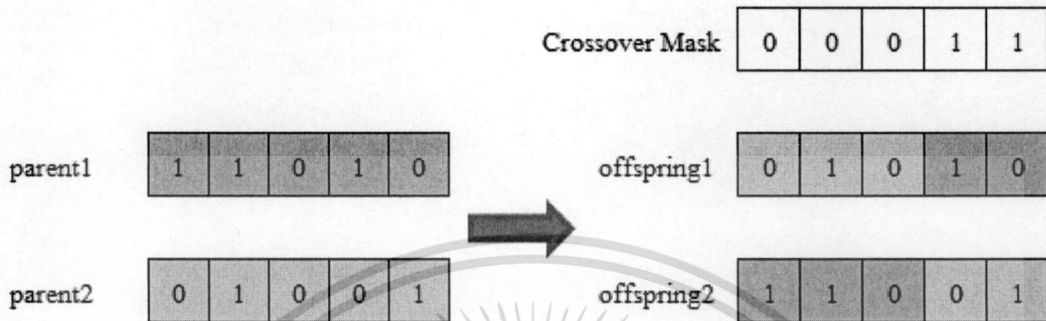
2.2.2.5 Uniform Crossover

ใช้การไขว้เปลี่ยน โดยการเปรียบเทียบกับ Crossover Mark การทำงาน เริ่มต้นจากการสร้าง Crossover mask ด้วยการสุ่มตัวเลข (0,1) ตามจำนวนยีนในโครโมโซม โดยตัวเลขที่อยู่ในแต่ละยีนจะเป็นตัวบอกว่าโครโมโซมลูกจะต้องทำการคัดลอกยีนจากโครโมโซมของพ่อ หรือ แม่ โดยโครโมโซมลูกตัวที่ 1 และ 2 จะมีขั้นตอนการคัดลอกยีนจากโครโมโซมพ่อแม่ดังนี้

- โครโมโซมลูกตัวที่ 1 จะดูว่ายีนใน Crossover mask เป็นเลข 1 โครโมโซมลูกจะทำการคัดลอกยีนจากโครโมโซมพ่อแม่ตัวที่ 1 แต่ถ้ายีนใน Crossover mask เป็นเลข 0 โครโมโซมลูกจะทำการคัดลอกยีนจากโครโมโซมพ่อแม่ตัวที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- โครโมโซมลูกตัวที่ 2 ทำการสร้าง Crossover mask ขึ้นมาใหม่อีก 1 ชุดแล้วทำกระบวนการเดียวกับโครโมโซมลูกตัวที่ 1 หรือใช้ Crossover mask ร่วมกับโครโมโซมลูกตัวที่ 1 หรือใช้กฎในลักษณะที่ตรงกันข้ามกับโครโมโซมลูกตัวที่ 1 ก็ได้



รูปที่ 2.8 การไขว้เปลี่ยนด้วยวิธี Uniform Crossover

2.2.3 การกลายพันธุ์ (Mutation)

เป็นการเพิ่มลักษณะพิเศษให้กับโครโมโซมลูก ซึ่งอาจจะเป็นลักษณะที่ไม่พบในลักษณะพื้นฐานของประชากรเริ่มต้นที่เรากำหนด โดยทั่ว ๆ ไปในแต่ละรอบจะมียีนที่ถูกกลายพันธุ์เพียง 1 – 5% การกลายพันธุ์ส่วนใหญ่จะทำให้ค่าความเหมาะสมของโครโมโซมลดลงแต่จะเป็นการเพิ่มความหลากหลาย และทำให้ประชากรมีความแข็งแรงมากขึ้น เพราะการกลายพันธุ์เป็นการหันเหเส้นทางมารอบรรจบกันของคำตอบช่วยให้คำตอบที่เหมาะสมมารอบรรจบกันช้าลง

ขั้นตอนการกลายพันธุ์เริ่มต้นด้วยการกำหนดค่าคงที่ซึ่งเป็นค่าความน่าจะเป็นที่แต่ละยีนจะถูกกลายพันธุ์ (P_M) ซึ่งโดยทั่วไปเราจะกำหนด P_M ให้มีค่าต่ำ ๆ หรือประมาณ 0.01 (มีการกลายพันธุ์ 1%) หลังจากนั้นแต่ละยีนจะทำการสุ่มค่าตัวเลขระหว่าง 0 – 1 โดยถ้าค่าที่สุ่มได้มีค่าน้อยกว่าความน่าจะเป็นที่ตั้งไว้ แต่ละยีนนั้น จะถูกกลายพันธุ์ ให้ทำการสุ่มตัวเลขที่อยู่ในช่วงที่กำหนดไว้มาแทนที่ค่าที่อยู่ในตำแหน่งนั้น ๆ หรือทำการสุ่มตัวเลขเพื่อนำมาใช้คำนวณร่วมกับค่าเดิมในยีนนั้น ๆ เพื่อให้ได้เป็นค่าใหม่

2.3 การวัดความแตกต่างสำหรับการจำแนกประเภทของข้อมูล

การวัดความแตกต่างสำหรับการจำแนกประเภทของข้อมูลนั้น ใช้วิธีการเปรียบเทียบเปรียบเทียบผลลัพธ์ระหว่างชุดข้อมูลตั้งต้น กับชุดข้อมูลทดลอง กลุ่มข้อมูลที่มีผลลัพธ์ของความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แตกต่างกันน้อย จะถูกจัดให้เป็นข้อมูลประเภทเดียวกัน โดยมีสมการที่ใช้ในการคำนวณหาค่าความแตกต่าง ดังนี้

2.3.1 ระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง (The Distance Formula)

$$dis = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + \dots} \quad (2.4)$$

2.3.2 Mean Absolute Deviation (MAD)

$$MAD = \frac{1}{n} \sum_{i=1}^n |g_i - f_i| \quad (2.5)$$

2.3.3 Mean Square Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (g_i - f_i)^2 \quad (2.6)$$

2.3.4 Mean Absolute Percent Error (MAPE)

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{g_i - f_i}{g_i} \times 100 \right|}{n} \quad (2.7)$$

2.4 การปรับปรุงคุณภาพของภาพ

รูปภาพโดยทั่วไปนั้นอาจมีข้อบกพร่องเกิดขึ้นในกระบวนการใดกระบวนการหนึ่ง เช่น กระบวนการรับภาพ เป็นต้น ทำให้ได้ภาพที่ไม่ได้คุณภาพตามที่ต้องการ การปรับปรุงคุณภาพของภาพจึงเกิดขึ้นมาเพื่อให้รูปภาพมีคุณภาพที่ดีขึ้น โดยหลักการของการปรับปรุงคุณภาพ คือ ขั้นตอนการประมวลผลเพื่อให้ได้ภาพผลลัพธ์ที่เหมาะสมกับการใช้งานเฉพาะทางได้ดีกว่าภาพต้นฉบับ เนื่องจากการปรับปรุงคุณภาพของภาพในแต่ละประเภท ย่อมใช้วิธีการปรับปรุงที่ต่างกันออกไปตามแต่ลักษณะของภาพ และสิ่งรบกวน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับปรุงคุณภาพของภาพนั้น มีวิธีการปรับปรุงหลายแบบ โดยหลัก ๆ จะสามารถแบ่งออกได้เป็น 2 แนวทาง คือ การปรับปรุงคุณภาพของภาพในโดเมนตำแหน่ง และการปรับปรุงคุณภาพของภาพในโดเมนความถี่

2.4.1 การปรับปรุงคุณภาพของภาพในโดเมนตำแหน่ง

การปรับปรุงคุณภาพของภาพในโดเมนตำแหน่งนั้น จะประมวลผลอยู่บนพื้นฐานของตำแหน่งจุดภาพ โดยจะกระทำกับข้อมูลในแต่ละพิกเซล ซึ่งเทคนิคในการปรับปรุงคุณภาพสามารถแบ่งออกเป็นหัวข้อหลัก ได้หลายหัวข้อ เช่น การปรับปรุงคุณภาพของภาพเพื่อกำจัดสิ่งรบกวนภาพในโดเมนตำแหน่ง การปรับเปลี่ยนหรือการแปลงค่าความเข้มแสงของภาพ การปรับความคมชัดของภาพ การปรับภาพด้วยฟังก์ชันทางคณิตศาสตร์และฟังก์ชันตรรกะ โดยในที่นี้จะกล่าวถึงเฉพาะการปรับปรุงคุณภาพเพื่อกำจัดสิ่งรบกวนภาพในโดเมนตำแหน่ง เท่านั้น

การปรับปรุงคุณภาพเพื่อกำจัดสิ่งรบกวนภาพในโดเมนตำแหน่ง (Spatial Image Filtering)

เมื่อมีสัญญาณรบกวนภาพ ทำให้ภาพสูญเสียรายละเอียดไป การกำจัดสัญญาณรบกวนให้หมดไปจากภาพนั้นจึงทำได้ยาก เนื่องจากตำแหน่งและขนาดของสัญญาณรบกวนที่เกิดการรบกวนนั้นไม่แน่นอน สิ่งที่ได้คือ การลดผลกระทบของสัญญาณรบกวน เพื่อช่วยให้ได้ภาพที่มีรายละเอียดดีขึ้น และมีคุณภาพดีพอที่จะนำไปใช้งานได้ เทคนิคในการลดผลกระทบของสัญญาณรบกวนที่เราเรียกว่า “การกรองสัญญาณรบกวน (Noise filtering)”

2.4.1.1 การกรองสัญญาณเชิงซ้อนหรือตัวกรองลำดับสถิติ

(Nonlinear Filtering หรือ Order-Statistics Filter)

ตัวกรองลำดับสถิติ เป็นเทคนิคแบบไม่เป็นเชิงเส้นซึ่งผลที่ได้ขึ้นอยู่กับการจัดเรียงค่าของพิกเซลที่อยู่ภายในพื้นที่ภายในเทมเพลตตัวกรอง จากนั้นจะทำการแทนค่าพิกเซลนั้นด้วยค่าผลลัพธ์ที่ได้จากการจัดเรียงค่าของพิกเซล ตัวอย่างของเทคนิคตัวกรองลำดับสถิติ ได้แก่ ตัวกรองค่ามัธยฐาน (Median Filter) ตัวกรองค่ามากที่สุดหรือตัวกรองค่าน้อยที่สุด (Max-Min Filter) ตัวกรองค่ากึ่งกลาง (Midpoint Filter) และตัวกรองค่าเฉลี่ยตัดขอบ (Alpha-trimmed Mean Filter)

2.4.1.1.1 ตัวกรองมัธยฐาน (Median Filter)

เป็นเทคนิคของตัวกรองลำดับสถิติ โดยเทคนิคนี้จะใช้วิธีการแทนที่ค่าของพิกเซลนั้นด้วยค่ามัธยฐาน (Median) ของค่าความเข้มแสงในพิกเซลที่อยู่ในพื้นที่ภายในเทมเพลตตัวกรองนั้น

$$g(x, y) = \underset{(s,t) \in S_{xy}}{\text{median}} \{ \hat{f}(s, t) \} \quad (2.8)$$

โดยที่ \hat{f} คือ ความเข้มแสงของพิกเซล
 s_{xy} คือ เทมเพลตตัวกรอง หรือวินโดว์
 $g(x, y)$ คือ ค่าพิกเซลที่ตำแหน่ง (x, y)

การหาค่ามัธยฐานในกรณีที่เทมเพลตตัวกรองของ $n \times n$ และ n เป็นเลขคี่นั้น ภายหลังจากจัดเรียงค่าของความเข้มแสงของพิกเซล ที่อยู่ในเทมเพลตตัวกรอง จากค่าความเข้มแสงน้อยไปหา มากแล้วแทนที่ค่าพิกเซลนั้น ด้วยค่ากลางที่ตำแหน่ง $\left[\frac{n^2}{2} \right] + 1$ จากค่าความเข้มแสงที่ถูกจัดเรียงไว้

การใช้เทคนิคนี้สามารถลดผลกระทบของสิ่งรบกวนได้ โดยเฉพาะกรณีของผลกระทบจากสัญญาณรบกวนแบบอิมพัลส์ หรือสัญญาณรบกวนแบบซอลท์แอนด์เปปเปอร์ อย่างไรก็ตามในกรณีที่สัญญาณรบกวนในภาพมีปริมาณมาก การกรองสัญญาณรบกวนในภาพเพียงครั้งเดียวไม่เพียงพอที่จะกำจัดสัญญาณรบกวนให้หมดไปได้ ดังนั้นเราสามารถนำตัวกรองค่ามัธยฐานของภาพ มาวนซ้ำจะช่วยทำให้ได้ผลลัพธ์จากการกรองมีสัญญาณรบกวนน้อยลงไปตามลำดับ

2.4.1.1.2 ตัวกรองตัวกรองค่ามากที่สุดหรือตัวกรองค่าน้อยที่สุด (Max-Min Filter)

เป็นเทคนิคของตัวกรองลำดับสถิติอีกวิธีหนึ่ง โดยผลที่ได้จากเทคนิคนี้เป็นประโยชน์ในการหาจุดที่สว่างที่สุดของภาพ และยังสามารถใช้ลดสัญญาณรบกวนแบบเปปเปอร์ ซึ่งมีค่าความเข้มแสงต่ำ โดยทำการแทนที่จุดมืดของเปปเปอร์ ด้วยค่าความเข้มแสงสูงสุดในเทมเพลตตัวกรอง

$$g(x, y) = \max_{(s,t) \in S_{xy}} \{ \hat{f}(s, t) \} \quad (2.9)$$

อีกเทคนิคหนึ่ง คือ ตัวกรองค่าน้อยที่สุด ซึ่งเลือกค่าความเข้มแสงต่ำสุดในพื้นที่ของเทมเพลตตัวกรอง เทคนิคนี้สามารถใช้ในการหาจุดมืดที่สุดของภาพ หรือสามารถลดผลกระทบจากสัญญาณรบกวนแบบซอลท์ ด้วยการแทนที่จุดสว่างด้วยค่าความเข้มแสงต่ำสุดในเทมเพลตตัวกรอง

$$g(x, y) = \min_{(s,t) \in S_{xy}} \{ \hat{f}(s, t) \} \quad (2.10)$$

2.4.1.1.3 ตัวกรองค่ากึ่งกลาง (Midpoint Filter)

เทคนิคนี้เป็นการผสมระหว่างเทคนิคของตัวกรองลำดับสถิติ กับการหาค่าเฉลี่ยของค่าความเข้มแสงสูงสุด เบาะความเข้มแสงต่ำสุดในพื้นที่ของตัวกรอง

ค่าของตัวกรองค่ากึ่งกลาง สามารถคำนวณได้โดยการหาค่าเฉลี่ยระหว่างค่ามากที่สุด และค่าน้อยที่สุดภายในพื้นที่ของเทมเพลตตัวกรอง ดังนี้

$$g(x, y) = \frac{1}{2} \left[\max_{(s,t) \in S_{xy}} \{ \hat{f}(s, t) \} + \min \{ \hat{f}(s, t) \} \right] \quad (2.11)$$

เทคนิคนี้ไม่เหมาะสมกับสัญญาณรบกวนแบบซอลท์แอนด์เปปเปอร์ที่มีค่าไปในทางค่าสูงสุด และค่าต่ำสุด ผลลัพธ์ที่ได้จะเป็นการเฉลี่ยสัญญาณรบกวน แต่ไม่ได้กำจัดหรือลดระดับลง แต่เทคนิคนี้จะให้ผลลัพธ์ที่ดีที่สุดในกรณีของสัญญาณรบกวนที่มีการกระจายตัวแบบสุ่ม เช่น สัญญาณรบกวนแบบเกาส์เซียน หรือสัญญาณรบกวนที่มีการกระจายตัวแบบสม่ำเสมอ

2.4.1.1.4 ตัวกรองค่าเฉลี่ยตัดขอบ (Alpha-trimmed Mean Filter)

เทคนิคนี้เป็นลำดับสถิติที่ตัดค่าต่ำสุด $(d/2)$ และค่าสูงสุด $(d/2)$ ที่อยู่ในพื้นที่ภายในเทมเพลตตัวกรองออกแล้วเฉลี่ยค่าที่เหลืออยู่ ถ้าเราให้ $f_r(s, t)$ แทนค่าความเข้มแสงที่เหลืออยู่ $mn - d$ พิกเซล ค่าผลลัพธ์จากตัวกรองค่าเฉลี่ยตัดขอบ หาได้ดังนี้

$$g(x, y) = \frac{1}{mn - d} \sum_{(s,t) \in S_{xy}} f_r(s, t) \quad (2.12)$$

ค่า d มีค่าในช่วง $[0, mn - 1]$ โดยที่เมื่อ ค่าจากการกรองสัญญาณจะลดรูปเป็นค่าตัวกรองเฉลี่ยเชิงเลขคณิต ถ้ากำหนดให้ $d = mn - 1$ แล้วค่าที่ได้จะเป็นค่าตัวกรองมัธยฐาน

ผลลัพธ์ที่ได้จากเทคนิคนี้จะเป็นประโยชน์ในการลดผลกระทบจากสัญญาณรบกวนหลายชนิดด้วยกัน เช่น สัญญาณรบกวนที่มีการกระจายผสมระหว่างสัญญาณรบกวนแบบซอลท์แอนด์เปปเปอร์ กับสัญญาณรบกวนแบบเกาส์เซียน

2.4.1.2 การกรองสัญญาณเชิงเส้น (Linear Filtering)

การกรองสัญญาณเชิงเส้น เป็นการกรองสัญญาณด้วยฟังก์ชันคณิตศาสตร์ที่มีความสัมพันธ์เชิงเส้น ซึ่งเหมาะสมกับสัญญาณรบกวนในลักษณะที่เป็นเชิงเส้น เช่น สัญญาณรบกวนแบบเกาส์เซียน ตัวกรองสัญญาณแบบเชิงเส้นมีหลายวิธี เช่น ตัวกรองค่าเฉลี่ยเชิงเลขคณิต ตัวกรองเฉลี่ยเชิง

เอกสารถ้าคุณคิดว่าการคัดลอกเนื้อหาโดยไม่ได้รับอนุญาตเป็นการละเมิด กรุณาแจ้งให้เราทราบที่ info@scs.kmutt.ac.th เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.1.2.1 ตัวกรองค่าเฉลี่ยเชิงเลขคณิต (Arithmetic Mean Filter)

ตัวกรองค่าเฉลี่ยเชิงเลขคณิตเป็นวิธีการทำตัวกรองค่าเฉลี่ยที่ง่ายที่สุด โดยใช้วิธีคอนโวลูชันซึ่งมีค่าถ่วงน้ำหนักของเทมเพลตตัวกรองเป็น 1 ในทุกตำแหน่ง และค่าถ่วงน้ำหนักนี้ถูกนอร์มอลไลซ์ด้วยจำนวนพิกเซลทั้งหมดในเทมเพลตตัวกรองเท่ากับ $M \times N$ ค่าเฉลี่ยเลขคณิตสามารถเขียนเป็นสมการได้ดังนี้

$$g(x, y) = \frac{1}{mn} \sum_{(s,t) \in Sxy} f(s, t) \quad (2.13)$$

การใช้เทคนิคนี้ไม่สามารถลดผลกระทบจากสัญญาณรบกวน ผลลัพธ์ที่ได้จะทำให้ภาพเบลอ (Blur) ส่งผลให้ความคมชัดของภาพลดลง

2.4.1.2.2 ตัวกรองเฉลี่ยเชิงเรขาคณิต (Geometric Mean Filter)

ตัวกรองเฉลี่ยเชิงเรขาคณิต ทำการคำนวณผลคูณของค่าในตำแหน่งพิกเซลทั้งหมดในเทมเพลตตัวกรอง วิธีนี้ไม่ได้ใช้เทคนิคคอนโวลูชันเหมือนกับตัวกรองค่าเฉลี่ยเชิงเลขคณิต สามารถเขียนเป็นสมการในการคำนวณตัวกรองเฉลี่ยเชิงเรขาคณิต ได้ดังนี้

$$g(x, y) = \left[\prod_{(s,t) \in Sxy} f(s, t) \right]^{\frac{1}{mn}} \quad (2.14)$$

เทคนิคการลดสิ่งรบกวนวิธีนี้จะทำการปรับภาพให้เรียบ (Smoothing) ได้เช่นเดียวกับการใช้วิธีตัวกรองค่าเฉลี่ยเชิงเลขคณิตแต่จะให้ภาพที่มีรายละเอียดและคมชัดมากกว่า แต่การคำนวณจะทำได้ยากและใช้เวลามากขึ้น

2.4.1.2.3 ตัวกรองเฉลี่ยฮาร์โมนิก (Harmonic Mean Filter)

ตัวกรองเฉลี่ยฮาร์โมนิก สามารถเขียนเป็นสมการได้ดังนี้

$$g(x, y) = \frac{mn}{\sum_{(s,t) \in Sxy} \frac{1}{f(s, t)}} \quad (2.15)$$

เทคนิคนี้ใช้ในการลดผลกระทบของสัญญาณรบกวนแบบซอลท์ที่ได้ดี แต่ใช้ไม่ได้กับสัญญาณรบกวนแบบเปปเปอร์

2.4.1.2.4 ตัวกรองเฉลี่ยคอนทราฮาร์โมนิก (Contra Harmonic Mean Filter)

ตัวกรองเฉลี่ยคอนทราฮาร์โมนิก สามารถเขียนเป็นสมการได้ดังนี้

$$g(x, y) = \frac{\sum_{(s,t) \in Sxy} f(s, t)^{Q+1}}{\sum_{(s,t) \in Sxy} f(s, t)^Q} \quad (2.16)$$

โดยค่า Q เป็นค่าลำดับของตัวกรองลำดับ (Order Filter) เทคนิคนี้เหมาะสำหรับใช้ลดผลกระทบของสัญญาณรบกวนแบบซอลท์แอนด์เปปเปอร์ โดย

ค่า Q เป็นบวกจะใช้สำหรับลดผลกระทบของสัญญาณรบกวนแบบเปปเปอร์

ค่า Q เป็นลบจะใช้สำหรับลดผลกระทบของสัญญาณรบกวนแบบซอลท์

เทคนิคนี้ไม่สามารถใช้ลดผลกระทบของสัญญาณรบกวนทั้งสองแบบพร้อมกันได้ จะเห็นได้ว่าตัวกรองเฉลี่ยคอนทราฮาร์โมนิกจะลดรูปเป็นตัวกรองเฉลี่ยเชิงเลขคณิตได้ถ้าค่า $Q = 0$ และจะลดรูปเป็นตัวกรองฮาร์โมนิกถ้า $Q = -1$

แม้ว่าตัวกรองเฉลี่ยคอนทราฮาร์โมนิกจะให้ผลดีแต่มีความยุ่งยากในการเลือกค่าของ Q ให้เหมาะสมกับชนิดของสัญญาณรบกวน ถ้าเลือกค่าของ Q ผิดจะมีผลกระทบต่อการกรอง

2.4.1.3 ตัวกรองแบบปรับเปลี่ยน

เทคนิคการกรองสัญญาณรบกวนแบบปรับเปลี่ยนนั้น จะมีการปรับเปลี่ยนค่าไปตามคุณลักษณะทางสถิติของภาพ ภายในพื้นที่ที่เทมเพลตตัวกรอง ซึ่งมีขนาด $m \times n$ เทคนิคที่นิยมใช้สำหรับตัวกรองแบบปรับเปลี่ยนมีสองแบบคือ ตัวกรองเฉลี่ยแบบปรับเปลี่ยนในพื้นที่ (Adaptive Local Noise Filter) และตัวกรองมัชฐานแบบปรับเปลี่ยนในพื้นที่ (Adaptive Median Filter)

2.4.1.3.1 ตัวกรองเฉลี่ยแบบปรับเปลี่ยนในพื้นที่ (Adaptive Local Noise Filter หรือ Minimum Mean Square Error Filter : MMSE Filter)

ค่าทางสถิติที่วัดได้ดีที่สุดสำหรับตัวแปรสุ่ม คือ ค่าเฉลี่ย (Mean) และค่าความแปรปรวน (Variance) โดยที่ค่าเฉลี่ยจะให้ค่าเฉลี่ยของค่าความเข้มแสงในพื้นที่ และค่าความแปรปรวนจะให้ค่าเฉลี่ยคอนทราสต์ (ค่าความแตกต่างของความเข้มแสง (Contrast)) ในพื้นที่นั้น

ผลตอบสนองจากการกรองสัญญาณเทคนิคนี้ขึ้นอยู่กับค่าสำคัญ 4 ค่าคือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ค่าขนาดของสัญญาณรบกวนในภาพ
2. ค่า σ_n^2 (Variance : ค่าเบี่ยงเบนของสัญญาณรบกวน)
3. m_L ค่า Local Mean ของพิกเซลภายในพื้นที่เทมเพลตตัวกรอง
4. σ_L^2 ของจุดภาพในพื้นที่เทมเพลตตัวกรอง
ซึ่งหาค่าได้ดังนี้

$$g(x, y) = f(x, y) - \frac{\sigma_n^2}{\sigma_L^2} [f(x, y) - m_L] \quad (2.17)$$

ค่าเพียงค่าเดียวที่ต้องรู้หรือประมาณค่า คือ ค่าการเบี่ยงเบนรวมของสัญญาณรบกวน (σ_n^2) ค่านี้จะเป็นตัวแปรที่ใช้ในการปรับปรุงคุณสมบัติของตัวกรองสัญญาณส่วนค่าอื่น ๆ สามารถคำนวณได้จากค่าความเข้มแสงภายในพื้นที่ของเทมเพลตตัวกรอง ซึ่งมีจุดกึ่งกลางอยู่ที่พิกเซลตำแหน่ง (x, y) การลดผลกระทบของสัญญาณรบกวน โดยใช้เทคนิคตัวกรองเคลื่อนแบบปรับเปลี่ยนในพื้นที่จะให้ผลดีกว่าตัวกรองค่าเฉลี่ยเชิงเลขคณิต ซึ่งภาพที่ได้จะมีความคมชัดกว่า

2.4.1.3.2 ตัวกรองมัชฌิมาแบบปรับเปลี่ยนในพื้นที่ (Adaptive Median Filter)

เทคนิคตัวกรองค่ามัชฌิมาทำงานได้ดีในกรณีที่การกระจายของสัญญาณรบกวนมีปริมาณน้อย (โดยปกติคือ P_a และ P_b มีค่าน้อยกว่า 0.2) แต่เทคนิคตัวกรองมัชฌิมาแบบปรับเปลี่ยนในพื้นที่ สามารถใช้จัดการกับสัญญาณรบกวนแบบอิมพัลส์ ซึ่งมีการกระจายมากกว่านี้ได้ ข้อดีของเทคนิคตัวกรองแบบมัชฌิมาแบบปรับเปลี่ยนในพื้นที่อีกอย่างหนึ่งคือ ไม่ทำลายรายละเอียดของภาพในขณะที่ทำการกำจัดสัญญาณรบกวน ซึ่งเทคนิคตัวกรองค่ามัชฌิมาแบบธรรมดาทำไม่ได้

สำหรับตัวกรองแบบมัชฌิมาแบบปรับเปลี่ยนในพื้นที่นี้แตกต่างจากเทคนิคการกรองที่ผ่านมา คือ ขนาดของเทมเพลตตัวกรองจะมีการเปลี่ยนแปลงขณะทำการกรอง ซึ่งขนาดจะขึ้นอยู่กับเงื่อนไขที่ตั้งขึ้น ผลลัพธ์ของการกรองสำหรับจุดภาพแต่ละตำแหน่ง จะถูกนำไปแทนที่ค่าของจุดภาพ ณ ตำแหน่งนั้น ๆ ซึ่งเป็นจุดกึ่งกลางของหน้าต่างตัวกรอง

โครงสร้างของตัวกรองแบบมัชฌิมาแบบปรับเปลี่ยนในพื้นที่แบ่งเป็น 2 ระดับ โดยเมื่อทำการตรวจดูว่าค่าความเข้มแสงมัชฌิมาเป็นค่าของสัญญาณรบกวนแบบอิมพัลส์หรือไม่ ถ้าค่ามัชฌิมาไม่ใช่สัญญาณรบกวน จะทำการตรวจสอบค่าในตำแหน่งกึ่งกลางของเทมเพลตตัวกรอง ถ้าค่าทั้งสองไม่ใช่สัญญาณรบกวน จะทำให้ระดับแรก คือ แทนค่าผลลัพธ์การกรองด้วยค่าตำแหน่งกึ่งกลางของเทมเพลตตัวกรอง การทำเช่นนี้เป็นกรลึกละเอียดการเปลี่ยนแปลงค่าความเข้มแสงของภาพโดยไม่จำเป็น ทำให้คงคุณภาพของค่าพิกเซลตำแหน่งนั้นไว้ ส่วนระดับที่สอง เมื่อพบว่าค่า

ตำแหน่งกึ่งกลางของเทมเพลตตัวกรองเป็นค่าสัญญาณรบกวนแบบอิมพัลส์ จะทำการแทนค่าผลลัพธ์การกรอง ด้วยค่ามัธยฐาน

ในการตรวจสอบค่ามัธยฐานว่าเป็นสัญญาณรบกวนหรือไม่นั้น จากคำจำกัดความของสัญญาณรบกวนแบบอิมพัลส์ จะพบว่าเป็นค่าสูงสุด (จุดขาว) หรือต่ำสุด (จุดดำ) ดังนั้นถ้าค่ามัธยฐานไม่เป็นค่าสูงสุด หรือค่าต่ำสุดหรือเงื่อนไข $z_{\min} < z_{\text{med}} < z_{\max}$ เป็นจริง โดยที่

1. z_{\min} เป็นค่าความเข้มแสงต่ำสุดในหน้าต่างตัวกรอง
2. z_{med} เป็นค่าความเข้มแสงกลาง (Median) ในหน้าต่างตัวกรอง
3. z_{\max} เป็นค่าความเข้มแสงสูงสุดในหน้าต่างตัวกรอง

ถ้าเงื่อนไขนี้เป็นจริงแสดงว่าค่าความเข้มแสงมัธยฐานนั้นไม่ใช่ค่าที่เป็นสัญญาณรบกวนแบบอิมพัลส์ จึงข้ามไปทำระดับที่สอง ซึ่งเป็นการตรวจสอบว่าค่าความเข้มแสง ณ ตำแหน่ง (x, y) กึ่งกลางของเทมเพลตตัวกรองนั้นเป็นค่าของสัญญาณรบกวนแบบอิมพัลส์หรือไม่

- ถ้าเงื่อนไข $z_{\min} < f(x, y) < z_{\max}$ เป็นจริง

แสดงว่าค่าความเข้มแสง ณ ตำแหน่งนั้น ไม่ได้เป็นค่าของสัญญาณรบกวนแบบอิมพัลส์ เราจึงใช้ค่าความเข้มแสงของตำแหน่งนั้น โดยไม่มีการเปลี่ยนแปลงใด ๆ

- ถ้าเงื่อนไข $z_{\min} < f(x, y) < z_{\max}$ เป็นเท็จ

แสดงว่าค่าความเข้มแสงในตำแหน่งกึ่งกลางของเทมเพลตตัวกรองมีโอกาสเป็นค่าของสัญญาณรบกวนแบบอิมพัลส์ ดังนั้นเราจะแทนค่าความเข้มแสง ณ ตำแหน่ง (x, y) นั้นด้วยค่าความเข้มแสงมัธยฐาน z_{med}

- ถ้าเงื่อนไข $z_{\min} < z_{\text{med}} < z_{\max}$ เป็นเท็จ

แสดงว่าค่าความเข้มแสงมัธยฐานนั้น เป็นค่าที่เป็นสัญญาณรบกวนแบบอิมพัลส์ ดังนั้น เราจึงทำการขยายขนาดของเทมเพลตตัวกรอง เพื่อทำการหาค่าความเข้มแสงมัธยฐานที่ไม่มีผลกระทบจากสัญญาณรบกวนแบบอิมพัลส์ต่อไป หรือหยุดทำถ้าขนาดของเทมเพลตตัวกรองเพิ่มขึ้นจนถึงขนาดสูงสุดที่กำหนดไว้

*หมายเหตุ เมื่อเทมเพลตตัวกรองขยายไปยังตำแหน่ง (x, y) ตำแหน่งใหม่ ขนาดของเทมเพลตตัวกรองจะกลับไปเท่ากับขนาดเริ่มต้นทุกครั้ง

2.4.2 การปรับปรุงคุณภาพของภาพในโดเมนความถี่

การปรับปรุงคุณภาพของภาพในโดเมนความถี่นั้น มีหลักการเช่นเดียวกันกับการปรับปรุงคุณภาพของภาพในโดเมนตำแหน่งหรือเวลา แต่เปลี่ยนรูปแบบการประมวลผลไปทำในแกนความถี่ โดยทำการรับค่าสัมประสิทธิ์ความถี่แทนค่าความเข้มแสงหรือสีของพิกเซลในแต่ละตำแหน่ง โดย

เอกสารจะแบ่งเทคนิคการปรับคุณภาพของภาพเป็น 4 กลุ่ม ดังนั้นท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การปรับคุณภาพของภาพเพื่อลดสัญญาณรบกวนในช่วงความถี่ต่ำ หรือปรับความเรียบให้ภาพ (Low Pass Filter)

การปรับคุณภาพของภาพเพื่อลดสัญญาณรบกวนในช่วงความถี่ที่กำหนด (Band Pass Filter)

การปรับคุณภาพของภาพเพื่อเพิ่มความคมชัดของภาพ ด้วยเทคนิคการกรองสัญญาณความถี่สูง (High Pass Filter)

การปรับคุณภาพของภาพเพื่อเพิ่มความคมชัดของภาพ ด้วยเทคนิคการกรองสัญญาณแบบย้อนกลับ (Inverse Filter)

2.5 บทความงานวิจัยที่เกี่ยวข้อง

2.5.1 Learning Classification Rules With Genetic Algorithm

บทความวิจัยนี้เป็นของ Maria Muntean, Corina Rotari และ Ioan Ileana ในปี 2010 โดยบทความวิจัยนี้มีวัตถุประสงค์ในเรื่องของการค้นหาความแม่นยำของการจำแนกประเภท โดยมีการนำเจเนติกอัลกอริทึมมาช่วยในการปรับปรุงคุณภาพของการจำแนกประเภท

การจำแนกประเภทเป็นงานหนึ่งในงานคาด้าไมนิ่ง โดยอัลกอริทึมที่นิยมนำมาใช้ในการแยกประเภทนั้นมีหลายอัลกอริทึม อาทิเช่น decision-tree, rule-induction, association-based หรือแบบ hybrid เป็นต้น โดย IF-Then ก็เป็นกฎหนึ่งที่น่านำมาใช้กับข้อมูลที่มีขนาดตายตัว และนำเจเนติกอัลกอริทึมมาประยุกต์ใช้กับ rule-induction เพื่อเพิ่มความสามารถในการจำแนกประเภท

2.5.1.1 Rule induction using genetic algorithm

เจเนติกอัลกอริทึมถูกนำมาใช้เพิ่มเพื่อประสิทธิภาพสำหรับการจำแนกประเภท ในงานคาด้าไมนิ่ง ประชากรเป็นตัวแทนของ solution search space ซึ่งก็คือ IF-Then rule โครโมโซมคือ IF-Then rule และยีนคือ ส่วนย่อยของ IF-Then rule ซึ่งขึ้นอยู่กับค่าของ fittest ซึ่งประชากรใหม่จะเกิดขึ้นจากกฎดังกล่าว ซึ่งมีการตรวจสอบความถูกต้องจากชุดข้อมูลทดสอบที่เรียนรู้ ขั้นตอนการจำแนกข้อมูลโดยใช้เจเนติกอัลกอริทึมมีดังนี้

1. โหลดชุดข้อมูลที่ต้องการประมวลผล
2. ไล่ IF-Then rule
3. คำนวณจำนวนประเภทของกลุ่มข้อมูล ที่มีการกระจายอยู่ในกฎ
4. เปรียบเทียบการกระจายตัวของประเภทกับกฎและตัวอย่าง
5. กำหนดให้ตัวอย่างที่เป็นจริงแทนด้วยค่าบวก และแทนค่าส่วนที่เป็นเท็จด้วยค่าลบ
6. กำหนดค่าตัวชี้วัด

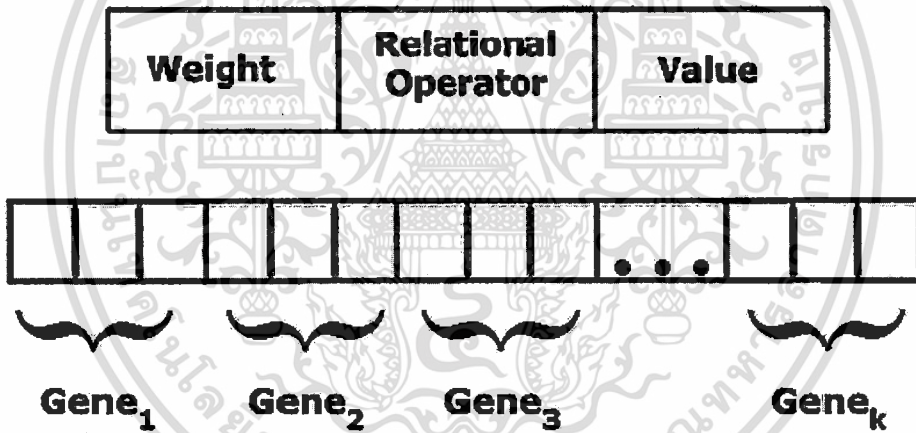
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. ตรวจสอบผลลัพธ์ของฟังก์ชันความเหมาะสมที่ได้กับค่าตัวชี้วัด โดยขั้นตอนที่กล่าวมาแล้วนั้น เป็นขั้นตอนก่อนที่จะเรียนรู้กฎและการคำนวณหาค่าความเหมาะสม

2.5.1.2 การนำเจเนติกอัลกอริทึมมาใช้

หน้าที่หลักของเจเนติกอัลกอริทึมคือการสร้างประชากร และการจำแนกประชากร โดยแทนค่าโครโมโซม ในแต่ละโครโมโซมจะมีการจัดลำดับ การจำแนก และการประเมินค่า

การใช้เจเนติกอัลกอริทึมจะช่วยให้การเข้ารหัสโครโมโซมมีความยืดหยุ่น ซึ่งแต่ละโครโมโซมมีรูปแบบในการแทนค่าเพื่อให้สอดคล้องกับกฎการจำแนก (classification rule) โดยในแต่ละโครโมโซมจะถูกแบ่งออกเป็น 3 ส่วนคือน้ำหนัก(weight), relational operation และค่า (value) โดยน้ำหนักไว้สำหรับเป็นลำดับการตรวจสอบค่าความตรงกันของยีน relational operation ไว้สำหรับ ตรวจสอบค่าของน้ำหนักเปรียบเทียบกับค่า threshold และค่าของยีน



รูปที่ 2.9 ลักษณะของยีน และลักษณะของโครโมโซม

แม้ว่าจำนวนของยีนจะมีค่าคงที่ตามจำนวนของกฎที่เป็นเงื่อนไขต่าง ๆ แต่ก็มีวิธีการคัดเลือกพ่อแม่ ในแต่ละครั้งนั้นพ่อแม่จะถูกคัดเลือกขึ้นมา 3 รอบและในแต่ละรอบโครโมโซมที่มีค่าความเหมาะสมสูงสุดจะถูกเลือกออกมาเป็นพ่อแม่ นอกจากนี้การคัดเลือกในรอบแรกเป็นการสุ่มโครโมโซมขึ้นมา 6 โครโมโซมเพื่อทำการคัดเลือก โดยเลือกโครโมโซมที่มีค่าความเหมาะสมดีที่สุดในอันดับแรกขึ้นมาเป็นโครโมโซมพ่อแม่ เพื่อต้องการให้เกิดโครโมโซมรุ่นลูกที่ดี ซึ่งใช้วิธีการ 2-point crossover ในการดำเนินการเพื่อให้เกิดโครโมโซมรุ่นลูก ซึ่งประชากรรุ่นลูกจะมีอัตราการกลายพันธุ์ 30%

การทำงานของเจเนติกอัลกอริทึมแต่ละครั้งจะสามารถจำแนกประเภทออกได้ 2 ประเภทซึ่งเป้าหมายที่ต้องการคือต้องการทำนายประเภทของกลุ่มตัวอย่างดังนั้นเจเนติกอัลกอริทึม จึงถูกเรียกใช้อย่างน้อย 1 ครั้งในแต่ละประเภท

ตัวอย่างเช่น

กลุ่มตัวอย่างมีทั้งสิ้น 6 ประเภท จะต้องเรียกใช้เจเนติกอัลกอริทึมอย่างน้อย 6 ครั้งเนื่องจากการเรียกใช้เจเนติกอัลกอริทึมครั้งที่ 1 จะใช้สำหรับการทำนายลักษณะของประเภทที่ 1 เมื่อมีการเรียกใช้เจเนติกอัลกอริทึมในครั้งที่ 2 จะเป็นการทำนายลักษณะของประเภทที่ 2 ดังนั้นเจเนติกอัลกอริทึมจึงสามารถเข้ามาใช้ในการดำเนินการกับชุดข้อมูลที่มีจำนวนประเภทไม่คงที่ได้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม

3.1 การนำเจเนติกอัลกอริทึมใช้ในการจำแนกประเภทของข้อมูล

การจำแนกประเภทของข้อมูล เป็นกระบวนการสร้าง โมเดลจัดการชุดข้อมูล โดยจำแนกให้ชุดข้อมูลที่คล้ายกัน อยู่ในกลุ่มข้อมูลประเภทเดียวกัน ซึ่งการสร้างจำแนกประเภทของข้อมูล ออกเป็นจำนวนกี่กลุ่มนั้น ขึ้นอยู่กับการวิเคราะห์กลุ่มของข้อมูลจากชุดข้อมูลเรียนรู้ โดยนำชุดข้อมูลเรียนรู้มาสอนให้ระบบเรียนรู้ลักษณะของกลุ่มของข้อมูลในแต่ละประเภทของชุดข้อมูล ผลลัพธ์ที่ได้จากการเรียนรู้ คือ โมเดลต้นแบบในการจัดประเภทของข้อมูล และได้มีการนำเจเนติกอัลกอริทึมมาใช้ในการสร้าง โมเดลต้นแบบในการจัดประเภทของข้อมูล เพื่อให้ได้โมเดลที่มีประสิทธิภาพในการจำแนกประเภทของข้อมูล

3.1.1 วิธีการสร้างโปรแกรมในการจำแนกประเภทของข้อมูล

โปรแกรมในการจำแนกประเภทของข้อมูลนี้ ออกแบบให้รองรับกับการจำแนกประเภทของข้อมูล ที่มีชุดข้อมูลเป็นประเภทตัวเลขเท่านั้น สามารถกำหนดจำนวนประชากรได้ สามารถปรับเปลี่ยนให้รองรับกับขนาดของโครโมโซม และประเภทที่มีอยู่ในชุดข้อมูลได้เอง ซึ่งมีองค์ประกอบในการพัฒนาโปรแกรม ดังนี้

3.1.1.1 การกำหนดประเภทของชุดข้อมูล

การกำหนดประเภทของชุดข้อมูล โดยทำการวิเคราะห์ชุดข้อมูลทดลองที่นำมาดำเนินการ นั้น ว่าสามารถแบ่งออกประเภทของข้อมูลได้เป็นกี่ประเภท เพื่อนำมาใช้สร้าง โมเดลการจำแนกประเภทของข้อมูลให้รองรับและเหมาะสมกับชุดข้อมูลนั้น ๆ

3.1.1.2 การสร้างประชากร

ประชากร คือ ชุดของโครโมโซมที่สร้างขึ้น เพื่อที่จะนำไปผ่านกระบวนการคัดเลือก เพื่อเป็นโครโมโซมพ่อ – แม่ สำหรับการสร้างประชากรในรุ่นถัดไป

3.1.1.3 การสร้างโครโมโซมสำหรับเป็นต้นแบบในการจำแนกประเภทของข้อมูล

โครโมโซมที่ใช้สำหรับการเป็นโมเดลในการจำแนกประเภทของข้อมูลนั้น ใน 1 โครโมโซมจะต้องมีประเภทหรือกลุ่มของข้อมูลครบตามที่ได้ทำการวิเคราะห์ไว้ในขั้นตอนแรก และมีการสร้างโครโมโซมออกเป็น 2 แบบ แบบแรกคือเป็นตัวเลขฐาน 10 ตามชุดข้อมูล เพื่อใช้

เอกสารนี้เป็นเอกสารลิขสิทธิ์สงวนลิขสิทธิ์ของงานเพื่อการศึกษาเท่านั้น เมื่อผู้จัดทำเผยแพร่เอกสารนี้เป็นการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

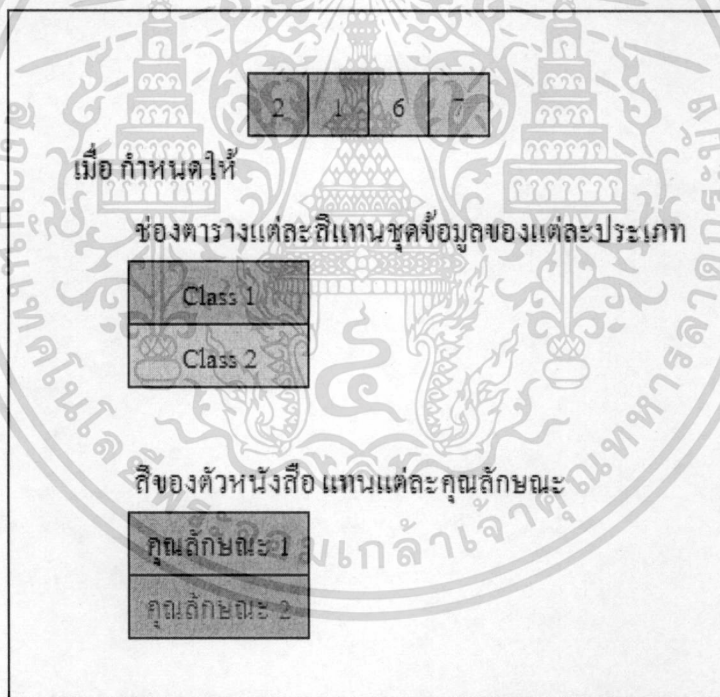
สำหรับการคำนวณหาค่าความเหมาะสม และการจำแนกประเภทของข้อมูล และอีกชุดคือแปลงเป็นตัวเลขฐาน 2 เพื่อใช้สำหรับการทำการไขว้เปลี่ยน และการกลายพันธุ์

ตารางที่ 3.1 ตัวอย่างชุดข้อมูลทดสอบ สำหรับสร้างโครโมโซม

ประเภท	คุณลักษณะ1	คุณลักษณะ2
A	2	1
B	6	7
A	1	1

จากตัวอย่างชุดข้อมูลทดสอบ สามารถนำมาสร้างเป็นโครโมโซม ได้ดังนี้

1. สุ่มเลือกข้อมูลประเภท A และ B มาประเภทละ 1 ชุดเพื่อมาสร้างเป็นโครโมโซม



รูปที่ 3.1 ตัวอย่างจำลองแบบโครโมโซมแบบเลขฐาน 10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. แปลงชุดข้อมูลโครโมโซม จากเลขฐาน 10 เป็นเลขฐาน 2

0	1	0	0	0	1	1	1	0	1	1	1
---	---	---	---	---	---	---	---	---	---	---	---

รูปที่ 3.2 ตัวอย่างแบบจำลองโครโมโซมแบบเลขฐาน 2

3.1.1.4 การคำนวณหาค่าความเหมาะสม

ค่าความเหมาะสม คำนวณได้จาก ค่าความถูกต้องที่โครโมโซม ทำการแบ่งกลุ่มของชุดข้อมูลทดสอบ

$$\text{ค่าความเหมาะสม} = \frac{\text{จำนวนชุดข้อมูลที่โครโมโซมแบ่งประเภทได้ถูกต้อง}}{\text{จำนวนชุดข้อมูลทั้งหมดที่นำมาเรียนรู้}} \quad (3.1)$$

วิธีการคำนวณเพื่อทำนายประเภทของชุดข้อมูล โดยเลือกกลุ่มที่มีผลลัพธ์ของแต่ละกลุ่มน้อยที่สุด เมื่อคำนวณเปรียบเทียบระหว่าง โครโมโซมกับชุดข้อมูลทดสอบ

$$\text{กลุ่ม} = \min(\text{ผลลัพธ์ของกลุ่ม}A, \text{ผลลัพธ์ของกลุ่ม}B) \quad (3.2)$$

สมการที่นำมาใช้คำนวณหาผลลัพธ์ เพื่อทำนายกลุ่ม ใช้ 4 สมการเพื่อนำมาหาคำตอบ ดังนี้

1. ระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

$$dis = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + \dots}$$

2. Mean Absolute Deviation (MAD)

$$MAD = \frac{1}{n} \sum_{i=1}^n |g_i - f_i|$$

3. Mean Square Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^n (g_i - f_i)^2$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. Mean Absolute Percent Error (MAPE)

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{g_i - f_i}{g_i} \right| \times 100}{n}$$

โดยในแต่ละรอบของการทดลองเพื่อนำมาคำนวณหาค่าความเหมาะสมนั้น จะเลือกใช้สมการในการคำนวณ รอบละ 1 สมการเท่านั้น

3.1.1.5 การเลือกโครโมโซมพ่อแม่

ใช้วิธีการคัดเลือกแบบ Roulette Wheel Weighting โดยใช้เทคนิค Rank Weighting ซึ่งมีขั้นตอนในการทำ ดังนี้

1. เรียงลำดับโครโมโซม โดยเรียงจากโครโมโซมที่มีค่าความเหมาะสมจากมากไปน้อย
2. คำนวณหาค่าความน่าจะเป็นของแต่ละโครโมโซม

$$P_n = \frac{N - n + 1}{\sum_{n=1}^N n} \quad (3.3)$$

$$\sum_{n=1}^N n = \frac{N}{2} [2 + (n-1)]$$

3. คำนวณหาพื้นที่ให้กับแต่ละโครโมโซม โดยพื้นที่รวมทั้งหมด คือ 1
4. เลือกโครโมโซมมาเป็นพ่อ - แม่ โดยการสุ่มตัวเลขระหว่าง 0 - 1 เพื่อดูว่าค่าอยู่ในช่วงพื้นที่ของโครโมโซมใด

3.1.1.6 การไขว้เปลี่ยน

นำโครโมโซมพ่อแม่ที่คัดเลือกมาทำการไขว้เปลี่ยน เมื่อให้เกิดโครโมโซมรุ่นใหม่ โดยใช้วิธี Uniform Crossover

1. สุ่มค่าระหว่าง 0 - 1 เพื่อหาค่าความน่าจะเป็นในการไขว้เปลี่ยน หากมีค่าน้อยกว่าค่าความน่าจะเป็นในการไขว้เปลี่ยน ที่กำหนดไว้ ให้ดำเนินการไขว้เปลี่ยน
2. ดำเนินการไขว้เปลี่ยน โดยทำการสลับตำแหน่งระหว่างค่าโครโมโซมพ่อ และแม่ตามวิธี Uniform Crossover

หากยีนใน Crossover mask = 1

โครโมโซมลูกตัวที่ 1 คัดลอกยีนจากพ่อ แม่ตัวที่ 1

โครโมโซมลูกตัวที่ 2 คัดลอกยีนจากพ่อ แม่ตัวที่ 2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หากยีนใน Crossover mask = 0

โครโมโซมลูกตัวที่ 1 คัดลอกยีนจากพ่อ แม่ตัวที่ 2

โครโมโซมลูกตัวที่ 2 คัดลอกยีนจากพ่อ แม่ตัวที่ 1

3.1.1.7 การกลายพันธุ์

นำโครโมโซมลูกมาทำการกลายพันธุ์

1. สุ่มค่าระหว่าง 0 – 1 เพื่อหาค่าความน่าจะเป็นในการกลายพันธุ์ หากมีค่าน้อยกว่าค่าความน่าจะเป็นในการกลายพันธุ์ ที่กำหนดไว้ ให้ดำเนินการกลายพันธุ์
2. สุ่มหาตำแหน่งของยีนที่ถูกกลายพันธุ์
3. ทำการกลายพันธุ์ โดยการสลับค่า จาก 0 เป็น 1 และจาก 1 เป็น 0

3.1.1.8 การสิ้นสุดกระบวนการ

กระบวนการในการสร้างโครโมโซมที่ดีที่สุด จะสิ้นสุดเมื่อ โครโมโซมที่ดีที่สุดในรุ่นสุดท้ายให้ค่าความเหมาะสมที่ไม่เปลี่ยนแปลง หรือวนครบรอบตามที่ได้กำหนดไว้

3.2 การนำวิธีการปรับปรุงคุณภาพของภาพมาประยุกต์ใช้กับการจำแนกประเภทของข้อมูล

ในการปรับปรุงคุณภาพของภาพ มีวิธีการปรับปรุงคุณภาพหลายแบบ หนึ่งในนั้นคือ การกำจัดสิ่งรบกวน (Noise) ในภาพ ที่เรียกว่าเป็นการกรองสัญญาณรบกวน โดยเทคนิคที่นำมาประยุกต์ใช้ คือ การกำจัดสิ่งรบกวนในภาพในโดเมนตำแหน่ง โดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

3.2.1 การนำตัวกรองค่ามัธยฐานมาประยุกต์ใช้กับชุดข้อมูลเรียนรู้

จะใช้วิธีการแทนที่ค่าของพิกเซลด้วยค่ามัธยฐาน (Median) ของค่าความเข้มแสงในพิกเซลที่อยู่ในพื้นที่ภายในเทมเพลตตัวกรองนั้น

$$g(x, y) = \underset{(s,t) \in Sxy}{\text{median}} \{f(s, t)\}$$

3.2.1.1 การนำมาประยุกต์ใช้กับชุดข้อมูลเรียนรู้

1. นำชุดข้อมูลสำหรับการทดสอบมาแบ่งออกเป็น 2 กลุ่ม คือ ชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. จัดเตรียมชุดข้อมูลเรียนรู้ โดยเตรียมข้อมูลเรียนรู้สำหรับไว้ปรับปรุง ทีละกลุ่มข้อมูล
3. ใช้โปรแกรม Matlab มาช่วยในการปรับปรุงชุดข้อมูลเรียนรู้ โดยวิธีการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน
4. นำข้อมูลที่ผ่านการปรับปรุงแล้ว เข้าสู่ฐานข้อมูล เพื่อเตรียมไว้ให้โปรแกรมจำแนกประเภทของข้อมูล มาใช้ชุดข้อมูลที่ปรับปรุงแล้วเป็นชุดข้อมูลเรียนรู้ สำหรับสร้างเป็นโครโมโซมต้นแบบสำหรับการใช้ในการจำแนกประเภทของข้อมูล
5. โปรแกรมจำแนกประเภทของข้อมูลมาทำการสร้างโครโมโซมต้นแบบ

3.2.1.2 ขั้นตอนในการนำชุดข้อมูลมาผ่านกระบวนการปรับปรุง

ตารางที่ 3.2 ตัวอย่างชุดข้อมูลเรียนรู้ ที่เตรียมไปปรับปรุงโดยวิธีการกรองสัญญาณเชิงซ้อนหรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

คุณลักษณะ 1	คุณลักษณะ 2	คุณลักษณะ 3	คุณลักษณะ 4	ประเภท
5.1	3.5	1.4	0.2	A
4.9	3	1.4	0.2	A
4.7	3.2	1.3	0.2	A
4.6	3.1	1.5	0.2	A
5	3.6	1.4	0.2	A
5.4	3.9	1.7	0.4	A
7	3.2	4.7	1.4	B
6.4	3.2	4.5	1.5	B
6.9	3.1	4.9	1.5	B
5.5	2.3	4	1.3	B
6.5	2.8	4.6	1.5	B
5.7	2.8	4.5	1.3	B

1. เตรียมข้อมูลสำหรับนำเข้าโปรแกรม Matlab ทีละประเภท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ตัวอย่างชุดข้อมูลประเภท A ที่เตรียมเข้าสู่การปรับปรุงฯ

5.1	3.5	1.4	0.2
4.9	3	1.4	0.2
4.7	3.2	1.3	0.2
4.6	3.1	1.5	0.2
5	3.6	1.4	0.2
5.4	3.9	1.7	0.4

ตารางที่ 3.4 ตัวอย่างชุดข้อมูลประเภท B ที่เตรียมเข้าสู่การปรับปรุงฯ

7	3.2	4.7	1.4
6.4	3.2	4.5	1.5
6.9	3.1	4.9	1.5
5.5	2.3	4	1.3
6.5	2.8	4.6	1.5
5.7	2.8	4.5	1.3

- ใช้โปรแกรม Matlab ในการปรับปรุงชุดข้อมูลที่ละประเภทโดยใช้คำสั่ง `medfilt2()`

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

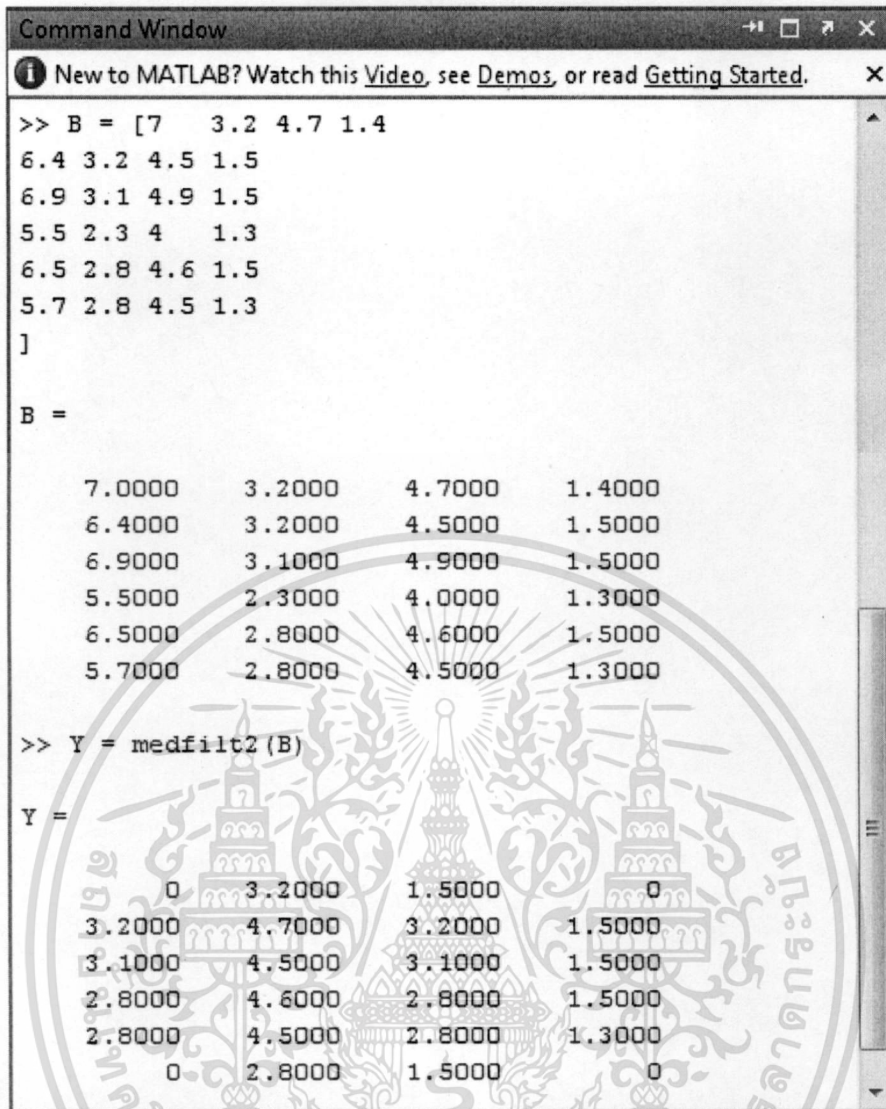
Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
A = [5.1    3.5 1.4 0.2
4.9 3    1.4 0.2
4.7 3.2 1.3 0.2
4.6 3.1 1.5 0.2
5    3.6 1.4 0.2
5.4 3.9 1.7 0.4
]
A =
5.1000    3.5000    1.4000    0.2000
4.9000    3.0000    1.4000    0.2000
4.7000    3.2000    1.3000    0.2000
4.6000    3.1000    1.5000    0.2000
5.0000    3.6000    1.4000    0.2000
5.4000    3.9000    1.7000    0.4000

>> Y = medfilt2(A)
Y =
0.0000    1.4000    0.2000    0.0000
3.2000    3.2000    1.4000    0.2000
3.1000    3.1000    1.4000    0.2000
3.2000    3.2000    1.4000    0.2000
3.6000    3.6000    1.5000    0.2000
0.0000    1.7000    0.4000    0.0000

```

รูปที่ 3.3 โปรแกรม Matlab ที่ทำการปรับปรุงชุดข้อมูลประเภท A

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



```

Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
>> B = [7 3.2 4.7 1.4
6.4 3.2 4.5 1.5
6.9 3.1 4.9 1.5
5.5 2.3 4 1.3
6.5 2.8 4.6 1.5
5.7 2.8 4.5 1.3
]
B =
7.0000 3.2000 4.7000 1.4000
6.4000 3.2000 4.5000 1.5000
6.9000 3.1000 4.9000 1.5000
5.5000 2.3000 4.0000 1.3000
6.5000 2.8000 4.6000 1.5000
5.7000 2.8000 4.5000 1.3000

>> Y = medfilt2(B)
Y =
0 3.2000 1.5000 0
3.2000 4.7000 3.2000 1.5000
3.1000 4.5000 3.1000 1.5000
2.8000 4.6000 2.8000 1.5000
2.8000 4.5000 2.8000 1.3000
0 2.8000 1.5000 0

```

รูปที่ 3.4 โปรแกรม Matlab ที่การปรับปรุงชุดข้อมูลประเภท B

- นำข้อมูลที่ละชุดที่ผ่านการปรับปรุงแล้ว เข้าสู่ตารางฐานข้อมูล สำหรับให้โปรแกรมจำแนกประเภทของข้อมูลมาทำการเรียนรู้ เพื่อสร้างเป็นโครโมโซมต้นแบบ โดยตัดบรรทัดบนสุด และล่างสุดของชุดข้อมูลที่ผ่านการปรับปรุงออก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.5 ตัวอย่างชุดข้อมูลที่ผ่านการปรับปรุงฯ และพร้อมเข้าสู่ตารางฐานข้อมูล

คุณลักษณะ 1	คุณลักษณะ 2	คุณลักษณะ 3	คุณลักษณะ 4	ประเภท
3.2	3.2	1.4	0.2	A
3.1	3.1	1.4	0.2	A
3.2	3.2	1.4	0.2	A
3.6	3.6	1.5	0.2	A
3.2	4.7	3.2	1.5	B
3.1	4.5	3.1	1.5	B
2.8	4.6	2.8	1.5	B
2.8	4.5	2.8	1.3	B

3.2.1.3 การใช้งานร่วมกับเจเนติกอัลกอริทึม

ใช้โปรแกรมจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม มาทำการสร้างโครโมโซมต้นแบบ โดยใช้ชุดข้อมูลเรียนรู้ ที่ผ่านกระบวนการปรับปรุงโดยวิธีการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

3.3 การทดสอบประสิทธิภาพของโครโมโซมต้นแบบในการจำแนกประเภทของข้อมูล

1. สร้างโครโมโซมต้นแบบ ที่ใช้สมการในการทำนายประเภทของชุดข้อมูล ทั้ง 4 แบบ ดำเนินการทีละสมการ สำหรับการคำนวณหาค่าความเหมาะสม ซึ่งโครโมโซมต้นแบบ จะได้จากการเรียนรู้ในชุดข้อมูลเรียนรู้ปกติ และชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยวิธีการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามาตรฐาน
2. ภายหลังจากได้โครโมโซมต้นแบบมาแล้ว ให้นำโครโมโซมต้นแบบ มาทำการทำนายประเภทของชุดข้อมูล โดยใช้สมการทั้ง 4 สมการมาทำนายประเภทของข้อมูล
3. คำนวณหาเปอร์เซ็นต์ความถูกต้องในการจำแนกประเภทของข้อมูล ของแต่ละสมการ เพื่อให้สามารถนำผลการทำนายจากแต่ละสมการมาทำการเปรียบเทียบกันได้

บทที่ 4

ผลการทดลอง

4.1 ตัวแปรที่เกี่ยวข้องในการทดลองการจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม

4.1.1 ตัวแปรต้น

ชุดข้อมูลเรียนรู้ ที่แยกออกจากชุดข้อมูลทดลอง ซึ่งประกอบไปด้วยชุดข้อมูลทดลองทั้ง 10 ชุด ดังที่ดังนี้

ชุดข้อมูลที่ใช้ในการทดลองนี้ ได้ใช้ชุดข้อมูลมาตรฐานจาก UCI Machine Learning Repository จำนวน 10 ชุด ดังนี้

4.1.1.1 ชุดข้อมูลผลการทดลองทางจิตวิทยา (Balance-Scale Data Set)

เป็นชุดข้อมูลผลการทดลองของแบบจำลองทางจิตวิทยาในด้านความสมดุลของสมอง โดยคำนวณจาก ค่าระยะทางซีกซ้าย x ค่าน้ำหนักซีกซ้าย และ ค่าระยะทางซีกขวา x ค่าน้ำหนักซีกขวา ถ้าเปรียบเทียบแล้วเท่ากัน แสดงว่ามีความสมดุล โดยผลการทดลองทั้งหมด 625 พบว่ามีทั้งสิ้น 49 ที่มีความสมดุล และเป็นด้านซ้ายและขวาอย่างละ 288 ข้อมูล

4.1.1.2 ชุดข้อมูลการบริจาคโลหิต (Blood-Transfusion Data Set)

ข้อมูลการบริจาคโลหิตจาก ที่ศูนย์บริการเมือง Hsin-Chu ในไต้หวัน โดยสุ่มข้อมูลจากฐานข้อมูลผู้บริจาคโลหิตจำนวน 748 ข้อมูล โดยข้อมูลที่ทำการเก็บได้แก่ ระยะเวลา(เดือน)นับจากบริจาคครั้งสุดท้าย จำนวนครั้งที่บริจาค ปริมาณโลหิตที่บริจาค และระยะเวลา(เดือน)นับจากบริจาคเลือดครั้งแรก

4.1.1.3 ชุดข้อมูลแก้ว (Glass Data Set)

ชุดข้อมูลคุณลักษณะ และประเภทของแก้ว เป็นข้อมูลจากสถาบันวิจัยกลาง สำนักงานบริการนิติวิทยาศาสตร์ ที่ได้เก็บคุณสมบัติ องค์ประกอบ และการแยกประเภทของแก้ว ข้อมูลองค์ประกอบของแก้ว ได้ทำการเก็บข้อมูลเกี่ยวกับ ดัชนีหักเห โซเดียม แมกนีเซียม อลูมิเนียม ซิลิกอน โพแทสเซียม แคลเซียม แบเรียม เหล็ก โดยแบ่งออกเป็นประเภทของแก้ว คือ

- 1 building_windows_float_processed
- 2 building_windows_non_float_processed

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 3 vehicle_windows_float_processed
- 4 vehicle_windows_non_float_processed (ไม่มีในชุดข้อมูลนี้)
- 5 containers
- 6 tableware
- 7 headlamps

4.1.1.4 ชุดข้อมูลการอยู่รอดของผู้รักษามะเร็งเต้านม (Haberman Data Set)

ชุดข้อมูลนี้ที่มีจากกรณีศึกษาที่ได้ดำเนินการระหว่างปี 1958 ถึง 1970 ที่มหาวิทยาลัยโรงพยาบาลบิลลิงส์ซิกาโก สำหรับการอยู่รอดของผู้ป่วยที่ได้รับการผ่าตัดมะเร็งเต้านมจำนวน 306 ข้อมูล โดยเก็บข้อมูลอายุของผู้ป่วยในช่วงเวลาของการดำเนินงาน ปีที่ผ่าตัด จำนวนของ axillary ที่ตรวจพบ สถานะการอยู่รอด

- 1 = ผู้ป่วยที่รอดชีวิต 5 ปีหรือนานกว่า
- 2 = ผู้ป่วยเสียชีวิตภายใน 5 ปี

4.1.1.5 ชุดข้อมูลดอกไอริส (Iris Data Set)

ชุดข้อมูลดอกไอริส เป็นชุดข้อมูลที่เกี่ยวข้องกับดอกไอริส จำนวน 150 ข้อมูล แต่ละข้อมูลประกอบด้วย 4 คุณลักษณะ ได้แก่ ความกว้างของกลีบเลี้ยง ความยาวของกลีบเลี้ยง ความกว้างของกลีบดอก และความยาวของกลีบดอก โดยดอกไอริสที่นำมาทดสอบ สุ่มมาจากดอกไอริส 3 สปีชีส์ ได้แก่ setosa, versicolor และ virginica สปีชีส์ละ 50 ข้อมูล ดังนั้นกลุ่มข้อมูลของชุดนี้คือ 3

4.1.1.6 ชุดข้อมูลความผิดปกติของตับ (Liver-disorders Data Set)

ชุดข้อมูลความผิดปกติของตับ จาก BUPA Medical Research Ltd. โดยมีจำนวนข้อมูลทั้งหมด 345 โดยเก็บข้อมูล mean corpuscular volume, alkphos alkaline phosphotase, sgpt alamine aminotransferase, sgot aspartate aminotransferase, gammagt gamma-glutamyl transpeptidase, drinks number of half-pint equivalents of alcoholic beverages drunk per day และสามารถแยกประเภทของตับออกเป็น 2 ประเภท

4.1.1.7 ชุดข้อมูลคนไข้โรคเบาหวานของชนเผ่าอินเดียนแดง (Pima Indians diabetes Data Set)

ชุดข้อมูลคนไข้โรคเบาหวานของชนเผ่าอินเดียนแดง เป็นชุดข้อมูลสำหรับทำนายผลตรวจสอบคนไข้ว่าเป็นโรคเบาหวานหรือไม่ โดยการอ้างอิงข้อมูลจากองค์การอนามัยโลก ข้อมูลประกอบด้วย ข้อมูลของคนไข้เพศหญิง อายุไม่ต่ำกว่า 21 ปี จากชนเผ่าอินเดียนแดง ที่อาศัยอยู่ใกล้เมืองฟีนิก มลรัฐแอริโซนา ประเทศสหรัฐอเมริกา ข้อมูลทั้งหมดจำนวน 768 ข้อมูล แต่ละข้อมูล

ประกอบด้วยคุณลักษณะจำนวน 8 คุณลักษณะ ข้อมูลมีทั้งหมด 2 กลุ่ม คือคนไข้ที่เป็นโรคเบาหวาน และคนไข้ที่ไม่เป็นโรคเบาหวาน โดยกลุ่มคนไข้ที่เป็นโรคเบาหวานมีทั้งหมด 268 คน และไม่เป็นโรคเบาหวานจำนวน 500 คน

4.1.1.8 ชุดข้อมูลระบบกระบวนการหาตำแหน่งของวัตถุใต้น้ำ โดยการส่งคลื่นโซนาร์ (Sonar Data Set)

ชุดข้อมูลการหาตำแหน่งของวัตถุใต้น้ำด้วยวิธีโซนาร์นั้น มีจำนวนทั้งสิ้น 208 ข้อมูล ประกอบด้วยคุณลักษณะจำนวน 60 คุณลักษณะ และแบ่งกลุ่มของชุดข้อมูลออกเป็น แร่ และก้อนหิน โดยเป็นก้อนหินจำนวน 97 ข้อมูล และแร่จำนวน 111 ข้อมูล กลุ่มข้อมูลของชุดนี้คือ 2

4.1.1.9 ชุดข้อมูลยานพาหนะ (Vehicle Data Set)

ชุดข้อมูลยานพาหนะ จาก Turing Institute, Glasgow, Scotland. โดยเก็บข้อมูลคุณลักษณะของยานพาหนะจำนวน 18 คุณสมบัติทั้งหมด 946 ชุดข้อมูล โดยแบ่งออกเป็น 4 ประเภทคือ opel จำนวน 240 ข้อมูล saab จำนวน 240 ข้อมูล bus จำนวน 240 ข้อมูล และ van จำนวน 226 ข้อมูล

4.1.1.10 ชุดข้อมูลไวน์ (Wine Data Set)

ชุดข้อมูลไวน์ โดยเป็นข้อมูลผลของการวิเคราะห์ทางเคมีของไวน์ 3 ชนิด จากภูมิภาคเดียวกันในอิตาลี ซึ่งเก็บข้อมูลทั้งหมด 178 ชุดข้อมูล และเก็บคุณลักษณะ 13 คุณลักษณะคือ Alcohol, Malic acid, Ash, Alkalinity of ash, Magnesium, Total phenols, Flavanoids, Nonflavanoid phenols, Proanthocyanins, Color intensity, Hue, OD280/OD315 of diluted wines, Proline

ตารางที่ 4.1 สรุปคุณลักษณะที่สำคัญของชุดข้อมูลทดสอบ

ชื่อชุดข้อมูล	จำนวนคุณลักษณะ	จำนวนประเภทของข้อมูล	ประเภทของข้อมูล	จำนวนข้อมูล
Balance-Scale	4	3	B, L, R	625
Blood-Transfusion	4	2	0, 1	748
Glass	9	6	1, 2, 3, 5, 6, 7	214
Haberman	3	2	1, 2	306
Iris	4	3	Iris-setosa, Iris-versicolor, Iris-virginica	150
Liver-disorders	6	2	1, 2	345
Pima Indians diabetes	8	2	0, 1	768
Sonar	60	2	M, R	208
Vehicle	18	4	bus, opel, saab, van	846
Wine	13	3	1, 2, 3	178

4.1.2 ตัวแปรตาม

ผลการจำแนกประเภทข้อมูล และประสิทธิภาพในการจำแนกประเภทของข้อมูล การวัดประสิทธิภาพ ในการจำแนกประเภทของข้อมูล

$$\text{ประสิทธิภาพ} = \frac{\text{จำนวนชุดข้อมูลที่โคร โม โชมแบ่งประเภทได้ถูกต้อง}}{\text{จำนวนชุดข้อมูลทั้งหมดที่นำมาเรียนรู้}} \times 100$$

4.1.3 ตัวแปรควบคุม

ในการทดลองการจำแนกข้อมูลด้วยเจเน็ติกอัลกอริทึมนี้ มีการกำหนดค่าตัวแปรที่ทำการควบคุมไว้ ดังนี้

ขนาดของจำนวนประชากร = 6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความน่าจะเป็นในการไขว้เปลี่ยน = 0.9

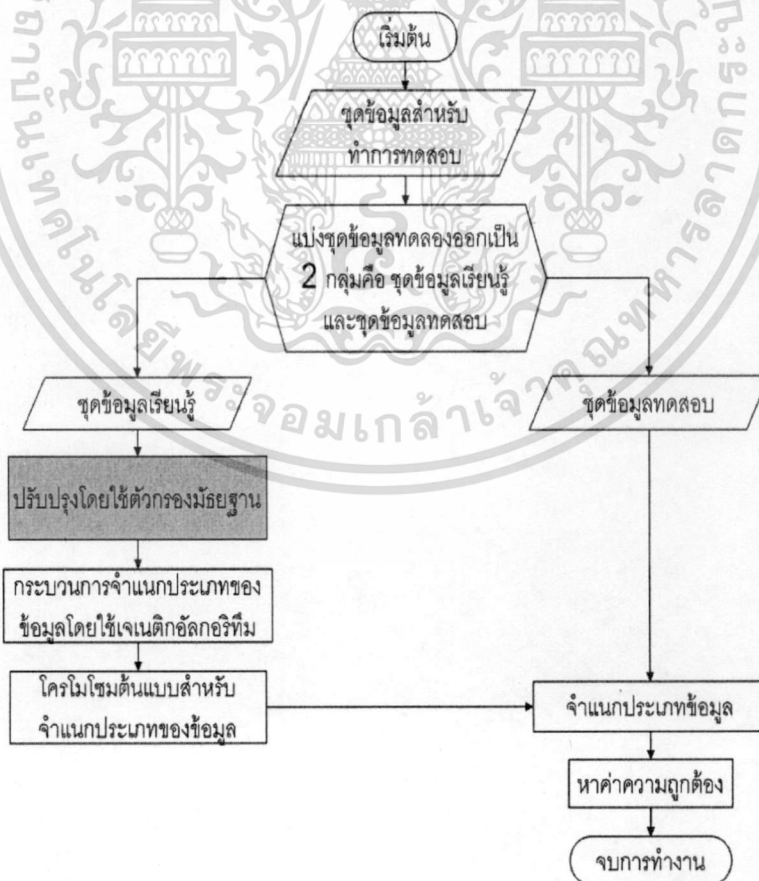
ความน่าจะเป็นในการกลายพันธุ์ = 0.01

ชุดข้อมูลที่นำมาทำการหาโครโมโซมต้นแบบ ในแต่ละรอบของการจำแนกประเภทของข้อมูล เพื่อให้สามารถนำมาเปรียบเทียบประสิทธิภาพในการทำนายประเภทของข้อมูลของสมการทั้ง 4 สมการ

4.2 การทดลองเพื่อทดสอบจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม

การทดลองนี้ได้พัฒนาโปรแกรมสำหรับการจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม ด้วยภาษาซีชาร์ป (C#) ที่มีการเก็บชุดข้อมูลทดลองลงบนฐานข้อมูล SQL Server 2008 และใช้โปรแกรม Matlab Version R2008a ในการปรับปรุงชุดข้อมูลเรียนรู้

การทดลองนี้ ได้ทำการทดสอบประสิทธิภาพในการจำแนกประเภทของข้อมูล โดยได้ทำการทดลองเพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทข้อมูลระหว่างการเรียนรู้จากชุดข้อมูลเรียนรู้แบบปกติ และการเรียนรู้จากชุดข้อมูลเรียนรู้ที่ผ่านกระบวนการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน



รูปที่ 4.1 ขั้นตอนในการดำเนินการทดลอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เฉพาะในเพื่อการศึกษาเท่านั้น เมื่อผู้ดูแลเห็นว่าไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนในการจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึมนั้น เริ่มต้นจากการแบ่งชุดข้อมูลสำหรับการทดสอบออกเป็น 2 กลุ่ม คือ ชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ จากนั้นนำชุดข้อมูลทดสอบมาเข้าสู่กระบวนการจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม เพื่อให้ได้โครโมโซมต้นแบบ สำหรับการจำแนกประเภทของข้อมูล จากนั้นให้นำโครโมโซมต้นแบบมาทำการจำแนกประเภทของข้อมูล และหาค่าความถูกต้องในการจำแนกประเภทของข้อมูล

ส่วนการนำวิธีการปรับปรุงคุณภาพของภาพมาประยุกต์ใช้กับการจำแนกประเภทของข้อมูลนั้น มีการเพิ่มขั้นตอนการปรับปรุงชุดข้อมูลเรียนรู้ โดยการนำชุดข้อมูลเรียนรู้มาผ่านกระบวนการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน ก่อนนำเข้าสู่กระบวนการจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึมตามปกติ

ทั้งนี้นอกการทดลองเพื่อทดสอบประสิทธิภาพโดยเปลี่ยนชุดข้อมูลในการเรียนรู้ ดังได้กล่าวมาแล้วนั้น ยังมีการเปลี่ยนสมการที่ใช้ในการจำแนกประเภทของข้อมูลอีก 4 แบบ คือ ระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง , Mean Absolute Deviation (MAD), Mean Absolute Percent Error (MAPE) และ Mean Square Error (MSE)



4.3 ผลการทดลองการทดสอบการจำแนกประเภทของข้อมูลโดยใช้เจเนติกอัลกอริทึม

ผลการทดลองการจำแนกประเภทข้อมูลโดยใช้เจเนติกอัลกอริทึม ดังนี้

4.3.1 ผลการทดลองโดยใช้ชุดข้อมูล Balance-Scale

4.3.1.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.2 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	78.5600	77.2800	75.5200	78.5600
2	74.7200	65.7600	57.7600	74.7200
3	66.7200	56.6400	56.4800	66.7200
4	63.2000	64.9600	46.0800	63.2000
5	61.6000	55.0400	35.5200	61.6000

ตารางที่ 4.3 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	83.1987	68.1745	84.4911	83.1987
2	82.7141	71.0824	86.2682	82.7141
3	82.3910	76.2520	81.7448	82.3910
4	80.1292	79.3215	66.8821	80.1292
5	78.0291	69.4669	80.7754	78.0291

4.3.1.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.4 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.1600	70.8800	72.8000	76.1600
2	71.0400	65.1200	60.9600	71.0400
3	67.8400	63.6800	30.2400	67.8400
4	66.8800	68.9600	46.0800	66.8800
5	42.0800	45.6000	39.2000	42.0800

ตารางที่ 4.5 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	85.9451	70.5977	86.9144	85.9451
2	81.5832	71.2439	83.0372	81.5832
3	80.1292	82.7141	53.4733	80.1292
4	77.3829	69.6284	46.8498	77.3829
5	71.7286	63.3280	44.2649	71.7286

4.3.1.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.6 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	72.6400	64.8000	78.4000	72.6400
2	69.9200	63.3600	41.2800	69.9200
3	67.0400	60.6400	70.0800	67.0400
4	58.8800	51.3600	56.6400	58.8800
5	49.9200	50.0800	45.7600	49.9200

ตารางที่ 4.7 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	84.3296	68.9822	66.0743	84.3296
2	71.4055	63.9742	80.1292	71.4055
3	69.4669	68.9822	51.8578	69.4669
4	68.0129	61.3893	71.8901	68.0129
5	67.8514	61.3893	84.1680	67.8514

4.3.1.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.8 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.0000	69.4400	71.0400	76.0000
2	73.2800	71.2000	62.2400	73.2800
3	71.0400	60.0000	57.1200	71.0400
4	69.1200	66.7200	73.4400	69.1200
5	63.6800	58.5600	63.5200	63.6800

ตารางที่ 4.9 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Balance-Scale ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	84.1680	68.8207	84.1680	84.1680
2	83.1987	68.1745	84.4911	83.1987
3	74.1519	64.6204	49.2730	74.1519
4	68.8207	69.3053	30.2100	68.8207
5	63.6511	61.2278	73.6672	63.6511

4.3.2 ผลการทดลองโดยใช้ชุดข้อมูล Blood-Transfusion

4.3.2.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.10 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	77.4064	76.8717	73.1283	77.4064
2	76.4706	74.8663	76.2032	76.4706
3	76.0695	76.7380	76.2032	76.0695
4	70.0535	68.0481	43.9840	67.1123
5	64.4385	62.1658	58.8235	62.9679

ตารางที่ 4.11 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.3441	76.3441	76.3441	76.3441
2	76.3441	69.0860	65.3226	76.3441
3	74.5968	73.5215	53.8979	74.5968
4	73.9247	72.9839	56.9893	73.9247
5	71.9086	71.9086	42.6075	71.9086

4.3.2.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.12 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	77.4064	76.7380	39.5722	77.2727
2	77.4064	76.7380	26.8717	77.2727
3	77.1390	76.8717	56.6845	77.5401
4	76.6043	76.8717	54.9465	77.0054
5	76.6043	76.7380	70.3209	76.8717

ตารางที่ 4.13 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	66.1290	66.1290	40.3226	66.1290
2	60.4839	58.6022	69.4893	60.4839
3	59.2742	59.6774	63.8441	59.2742
4	58.1989	59.5430	50.2688	58.1989
5	56.9893	46.2366	35.0807	56.9893

4.3.2.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.14 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.6043	76.7380	27.0054	76.6043
2	76.4706	75.5348	68.5829	75.1337
3	76.3369	75.2674	75.4011	75.0000
4	76.2032	72.7273	51.7380	74.7326
5	74.1979	73.1283	77.4064	72.8610

ตารางที่ 4.15 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.3441	76.3441	76.3441	76.3441
2	76.3441	76.3441	74.5968	76.3441
3	76.3441	76.3441	74.3280	76.3441
4	54.8387	58.6022	40.9946	54.8387
5	53.3602	53.3602	59.0054	53.3602

4.3.2.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.16 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.8717	76.7380	53.0749	76.2032
2	76.2032	75.9358	76.2032	76.6043
3	70.1872	70.0535	76.2032	69.1177
4	64.0374	62.9679	39.7059	63.5027
5	63.6364	67.2460	28.2086	67.1123

ตารางที่ 4.17 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Blood-Transfusion ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.8817	52.9570	54.5699	76.8817
2	76.3441	76.3441	69.7581	76.3441
3	76.3441	76.3441	30.1075	76.3441
4	76.2097	76.2097	51.7473	76.2097
5	71.2366	71.2366	42.7419	71.2366

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.3 ผลการทดลองโดยใช้ชุดข้อมูล Glass

4.3.3.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.18 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	52.8037	53.2710	0.0000	52.8037
2	50.9346	46.7290	6.0748	50.9346
3	50.4673	49.0654	0.0000	50.4673
4	48.5981	48.5981	0.0000	48.5981
5	45.7944	51.4019	0.0000	45.7944

ตารางที่ 4.19 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	79.2079	79.2079	0.0000	79.2079
2	77.2277	75.2475	0.0000	77.2277
3	77.2277	73.7624	13.3663	77.2277
4	76.2376	77.7228	13.3663	76.2376
5	72.2772	72.2772	0.0000	72.2772

4.3.3.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.20 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	53.7383	55.1402	0.0000	53.7383
2	53.7383	52.8037	0.0000	53.7383
3	53.7383	52.3365	0.0000	53.7383
4	50.4673	53.7383	0.0000	50.4673
5	45.3271	45.7944	0.0000	45.3271

ตารางที่ 4.21 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.7327	77.2277	0.0000	76.7327
2	76.7327	76.2376	0.0000	76.7327
3	76.2376	73.7624	0.0000	76.2376
4	75.7426	82.1782	0.0000	75.7426
5	75.7426	76.2376	0.0000	75.7426

4.3.3.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.22 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	55.1402	56.5421	0.0000	55.1402
2	50.4673	52.3365	0.0000	50.4673
3	50.0000	47.6636	32.7103	50.0000
4	50.0000	47.1963	0.0000	50.0000
5	48.1308	47.6636	0.0000	48.1308

ตารางที่ 4.23 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	78.2178	80.6931	13.3663	78.2178
2	75.7426	75.7426	13.3663	75.7426
3	72.7723	71.7822	0.0000	72.7723
4	65.8416	70.2970	0.0000	65.8416
5	61.8812	61.8812	0.0000	61.8812

4.3.3.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.24 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	58.4112	59.8131	0.0000	58.4112
2	51.8692	51.4019	0.0000	51.8692
3	50.0000	49.0654	0.0000	50.0000
4	47.1963	48.1308	0.0000	47.1963
5	44.8598	44.3925	0.0000	44.8598

ตารางที่ 4.25 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Glass ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อนหรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	78.7129	67.3267	13.3663	78.7129
2	77.2277	77.2277	0.0000	77.2277
3	76.7327	75.2475	13.3663	76.7327
4	72.7723	75.2475	13.3663	72.7723
5	62.3762	61.3861	0.0000	62.3762

4.3.4 ผลการทดลองโดยใช้ชุดข้อมูล Haberman

4.3.4.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.26 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.7974	76.4706	44.7712	76.7974
2	76.1438	75.8170	26.4706	76.1438
3	74.8366	74.1830	26.4706	74.8366
4	73.5294	73.5294	73.5294	73.5294
5	71.2418	67.6471	0.0000	71.2418

ตารางที่ 4.27 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	82.1192	82.1192	82.1192	82.1192
2	64.2384	61.5894	67.8808	64.2384
3	55.9603	55.2980	64.2384	55.9603
4	50.3311	50.6623	49.3378	50.3311
5	47.6821	47.6821	26.1589	47.6821

4.3.4.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.28 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	75.4902	75.1634	47.0588	75.4902
2	75.1634	76.7974	37.5817	75.1634
3	73.5294	73.5294	26.4706	73.5294
4	71.5686	77.1242	48.0392	71.5686
5	70.5882	76.4706	26.4706	70.5882

ตารางที่ 4.29 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	81.1258	81.7881	26.1589	81.1258
2	74.8344	64.5695	65.2318	74.8344
3	61.9205	61.9205	61.9205	61.9205
4	58.9404	82.4503	26.1589	58.9404
5	56.6225	58.9404	54.6358	56.6225

4.3.4.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.30 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	76.1438	75.1634	26.4706	76.1438
2	62.4183	61.4379	73.5294	62.4183
3	57.5163	47.7124	43.1373	57.5163
4	51.3072	55.8824	73.5294	51.3072
5	37.5817	34.9673	50.0000	37.5817

ตารางที่ 4.31 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	81.1258	81.7881	26.1589	81.1258
2	80.4636	81.1258	53.9735	80.4636
3	80.4636	80.4636	26.1589	80.4636
4	73.8411	73.8411	73.8411	73.8411
5	53.9735	53.9735	53.6424	53.9735

4.3.4.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.32 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	77.1242	73.8562	26.4706	77.1242
2	76.4706	75.4902	26.4706	76.4706
3	75.1634	75.4902	46.0784	75.1634
4	73.5294	73.5294	74.1830	73.5294
5	72.5490	73.8562	26.4706	72.5490

ตารางที่ 4.33 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Haberman ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	83.1126	82.1192	26.1589	83.1126
2	81.7881	81.7881	26.1589	81.7881
3	80.7947	80.7947	74.8344	80.7947
4	80.4636	80.4636	26.1589	80.4636
5	60.5960	59.2715	26.1589	60.5960

4.3.5 ผลการทดลองโดยใช้ชุดข้อมูล Iris

4.3.5.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.34 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	97.3333	97.3333	86.0000	97.3333
2	96.6667	97.3333	95.3333	96.6667
3	96.0000	92.0000	92.6667	96.0000
4	92.0000	90.6667	91.3333	92.0000
5	88.6667	88.0000	80.6667	88.6667

ตารางที่ 4.35 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	100.0000	98.6111	92.3611	100.0000
2	99.3056	97.2222	97.9167	99.3056
3	98.6111	96.5278	99.3056	98.6111
4	95.1389	96.5278	98.6111	95.1389
5	95.1389	93.7500	84.0278	95.1389

4.3.5.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.36 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	95.3333	96.0000	94.0000	95.3333
2	95.3333	95.3333	90.0000	95.3333
3	94.0000	95.3333	97.3333	94.0000
4	92.0000	95.3333	96.6667	92.0000
5	91.3333	95.3333	94.6667	91.3333

ตารางที่ 4.37 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	100.0000	100.0000	100.0000	100.0000
2	98.6111	96.5278	98.6111	98.6111
3	98.6111	96.5278	97.9167	98.6111
4	97.9167	98.6111	95.8333	97.9167
5	97.9167	97.2222	88.8889	97.9167

4.3.5.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.38 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	97.3333	85.3333	84.0000	97.3333
2	96.0000	94.0000	86.6667	96.0000
3	88.0000	88.0000	90.6667	88.0000
4	86.0000	85.3333	83.3333	86.0000
5	86.0000	85.3333	82.0000	86.0000

ตารางที่ 4.39 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	100.0000	100.0000	100.0000	100.0000
2	99.3056	98.6111	95.1389	99.3056
3	98.6111	99.3056	100.0000	98.6111
4	98.6111	97.9167	96.5278	98.6111
5	97.9167	98.6111	97.2222	97.9167

4.3.5.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.40 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	96.0000	94.0000	86.6667	96.0000
2	90.6667	92.0000	94.0000	90.6667
3	90.6667	89.3333	94.6667	90.6667
4	90.0000	88.6667	94.0000	90.0000
5	89.3333	91.3333	94.6667	89.3333

ตารางที่ 4.41 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Iris ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อนหรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	100.0000	100.0000	100.0000	100.0000
2	100.0000	100.0000	100.0000	100.0000
3	100.0000	100.0000	100.0000	100.0000
4	100.0000	98.6111	98.6111	100.0000
5	99.3056	97.2222	97.9167	99.3056

4.3.6 ผลการทดลองโดยใช้ชุดข้อมูล Liver-disorders

4.3.6.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.42 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	62.0290	63.1884	58.8406	62.0290
2	55.6522	55.6522	61.4493	55.6522
3	54.2029	54.7826	47.8261	54.2029
4	53.0435	54.7826	47.2464	53.0435
5	52.7536	52.1739	60.5797	52.7536

ตารางที่ 4.43 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	63.6364	63.9296	62.4633	63.6364
2	63.6364	62.1701	43.1085	63.6364
3	56.5982	56.5982	43.1085	56.5982
4	56.3050	53.6657	58.0645	56.3050
5	54.2522	53.6657	56.5982	54.2522

4.3.6.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.44 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	65.5073	67.2464	68.9855	65.5073
2	60.2899	67.5362	49.8551	60.2899
3	59.7101	61.4493	66.6667	59.7101
4	58.2609	58.5507	60.2899	58.2609
5	57.9710	55.9420	44.9275	57.9710

ตารางที่ 4.45 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	63.9296	63.9296	61.5836	63.9296
2	59.2375	64.2229	63.3431	59.2375
3	57.4780	48.6804	56.0117	57.4780
4	56.0117	57.7713	56.8915	56.0117
5	55.4252	55.7185	43.1085	55.4252

4.3.6.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.46 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	63.4783	59.7101	68.9855	63.4783
2	57.3913	55.0725	44.9275	57.3913
3	53.9130	54.4928	60.0000	53.9130
4	53.6232	58.2609	43.1884	53.6232
5	53.0435	49.5652	59.1304	53.0435

ตารางที่ 4.47 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	65.6892	64.5161	65.1026	65.6892
2	63.6364	60.1173	62.7566	63.6364
3	63.0499	62.4633	67.4487	63.0499
4	60.9971	61.8768	47.5073	60.9971
5	59.2375	58.6510	65.3959	59.2375

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.6.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.48 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	66.3768	58.8406	57.9710	66.3768
2	62.6087	56.2319	61.7391	62.6087
3	61.4493	59.7101	45.2174	61.4493
4	55.6522	56.2319	53.3333	55.6522
5	54.7826	55.3623	42.8986	54.7826

ตารางที่ 4.49 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Liver-disorders ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	65.1026	65.3959	62.7566	65.1026
2	63.6364	62.1701	43.1085	63.6364
3	63.0499	63.3431	66.8622	63.0499
4	61.8768	60.7038	64.5161	61.8768
5	58.3578	58.6510	44.8680	58.3578

4.3.7 ผลการทดลองโดยใช้ชุดข้อมูล Pima Indians diabetes

4.3.7.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.50 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	75.6510	74.3490	0.0000	75.6510
2	75.6510	74.2188	0.0000	75.6510
3	75.3906	76.3021	0.0000	75.3906
4	72.7865	72.5260	0.0000	72.7865
5	64.9740	64.8438	0.0000	64.9740

ตารางที่ 4.51 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	73.6911	75.3927	63.0890	73.6911
2	70.5497	68.3246	45.5497	70.5497
3	65.0524	62.1728	65.9686	65.0524
4	54.9738	62.3037	59.5550	54.9738
5	50.1309	51.8325	61.6492	50.1309

4.3.7.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.52 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	74.2188	73.4375	0.0000	74.2188
2	74.0885	74.0885	0.0000	74.0885
3	66.0156	65.8854	34.8958	66.0156
4	65.6250	66.5365	34.8958	65.6250
5	52.2135	52.6042	65.1042	52.2135

ตารางที่ 4.53 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	63.0890	63.4817	58.5079	63.0890
2	59.4241	63.7435	50.2618	59.4241
3	55.6283	56.2827	65.1832	55.6283
4	53.4031	54.0576	65.1832	53.4031
5	51.8325	51.8325	65.1832	51.8325

4.3.7.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.54 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	74.6094	74.8698	0.0000	74.6094
2	74.2188	73.0469	0.0000	74.2188
3	66.7969	69.0104	61.4583	66.7969
4	65.6250	65.6250	54.2969	65.6250
5	61.1979	63.0208	39.7135	61.1979

ตารางที่ 4.55 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	70.2880	72.1204	34.8168	70.2880
2	68.0628	67.4084	43.7173	68.0628
3	67.9319	69.1100	61.6492	67.9319
4	60.9948	59.8168	65.1832	60.9948
5	57.8534	60.6021	47.7749	57.8534

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.7.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.56 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	75.5208	75.2604	0.0000	75.5208
2	75.1302	72.3958	0.0000	75.1302
3	74.8698	75.0000	0.0000	74.8698
4	69.1406	67.4479	57.1615	69.1406
5	65.7552	65.2344	39.8438	65.7552

ตารางที่ 4.57 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Pima Indians diabetes ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	64.5288	65.5759	34.8168	64.5288
2	59.5550	64.0052	48.0367	59.5550
3	58.3770	63.0890	49.3456	58.3770
4	55.2356	57.9843	48.6911	55.2356
5	54.8429	61.6492	65.1832	54.8429

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.8 ผลการทดลองโดยใช้ชุดข้อมูล Sonar

4.3.8.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.58 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	68.7500	68.7500	44.7115	68.7500
2	61.0577	58.6539	45.6731	61.0577
3	59.6154	57.6923	53.3654	59.6154
4	57.6923	63.9423	47.5962	57.6923
5	53.3654	58.1731	55.7692	53.3654

ตารางที่ 4.59 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	71.0784	73.0392	75.9804	71.0784
2	70.5882	63.7255	72.0588	70.5882
3	66.1765	67.6471	68.1373	66.1765
4	63.2353	60.7843	66.6667	63.2353
5	53.9216	55.3922	71.5686	53.9216

4.3.8.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.60 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	70.1923	71.1539	47.1154	70.1923
2	69.2308	71.6346	69.2308	69.2308
3	61.5385	69.2308	70.6731	61.5385
4	59.1346	58.6539	58.6539	59.1346
5	56.7308	55.2885	51.9231	56.7308

ตารางที่ 4.61 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	69.6078	73.0392	73.5294	69.6078
2	69.1177	66.6667	56.3726	69.1177
3	63.2353	64.7059	69.1177	63.2353
4	59.3137	60.7843	60.2941	59.3137
5	54.4118	54.4118	50.4902	54.4118

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.8.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.62 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	60.0962	62.9808	66.3462	60.0962
2	60.0962	61.5385	72.5962	60.0962
3	59.6154	62.9808	70.6731	59.6154
4	58.6539	57.6923	65.8654	58.6539
5	54.3269	55.2885	59.1346	54.3269

ตารางที่ 4.63 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	64.7059	67.6471	75.9804	64.7059
2	63.7255	68.1373	77.4510	63.7255
3	62.7451	65.6863	75.0000	62.7451
4	61.7647	65.6863	58.8235	61.7647
5	54.9020	52.4510	50.0000	54.9020

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.8.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.64 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	62.0192	66.3462	53.3654	62.0192
2	60.0962	59.1346	59.6154	60.0962
3	57.2115	60.0962	54.3269	57.2115
4	56.7308	56.7308	55.2885	56.7308
5	55.7692	62.0192	53.3654	55.7692

ตารางที่ 4.65 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Sonar ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อนหรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	71.5686	68.6275	58.8235	71.5686
2	71.0784	74.5098	68.1373	71.0784
3	71.0784	73.0392	75.9804	71.0784
4	68.6275	68.6275	62.7451	68.6275
5	51.9608	65.6863	61.7647	51.9608

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.9 ผลการทดลองโดยใช้ชุดข้อมูล Vehicle

4.3.9.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.66 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	45.0355	46.2175	33.3333	45.0355
2	44.6809	44.2080	35.8156	44.6809
3	43.0260	44.4444	46.3357	43.0260
4	41.9622	43.8534	40.6619	41.9622
5	39.2435	37.7069	31.2057	39.2435

ตารางที่ 4.67 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	56.8019	58.4726	51.5513	56.8019
2	54.2959	53.2220	41.0501	54.2959
3	52.5060	55.2506	46.7780	52.5060
4	47.3747	48.4487	50.9547	47.3747
5	46.7780	47.0167	43.6754	46.7780

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.9.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.68 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	44.0898	43.4988	25.7683	44.0898
2	41.7258	39.8345	37.8251	41.7258
3	39.8345	38.1797	33.5697	39.8345
4	39.0071	40.7801	39.9527	39.0071
5	36.8794	38.1797	34.0426	36.8794

ตารางที่ 4.69 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	53.6993	57.6372	56.8019	53.6993
2	52.1480	55.7279	50.9547	52.1480
3	51.6706	57.9952	53.8186	51.6706
4	48.9260	50.2387	47.6134	48.9260
5	47.4940	46.6587	50.5967	47.4940

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.9.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.70 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	46.0993	45.1537	40.6619	46.0993
2	44.2080	43.9716	28.7234	44.2080
3	41.4894	40.4255	39.9527	41.4894
4	40.8984	41.1348	38.4161	40.8984
5	37.5887	28.4870	25.4137	37.5887

ตารางที่ 4.71 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	55.6086	56.5633	50.0000	55.6086
2	54.2959	55.8473	40.9308	54.2959
3	51.1933	51.4320	53.6993	51.1933
4	48.8067	46.4201	42.1241	48.8067
5	47.2554	49.6420	45.4654	47.2554

4.3.9.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.72 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	46.9267	43.2624	26.2411	46.9267
2	46.6903	46.9267	35.6974	46.6903
3	46.6903	44.2080	35.6974	46.6903
4	45.2719	45.3901	33.8061	45.2719
5	34.7518	34.2790	28.6052	34.7518

ตารางที่ 4.73 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Vehicle ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	55.2506	58.9499	48.6874	55.2506
2	55.0119	52.5060	36.8735	55.0119
3	52.1480	53.2220	48.0907	52.1480
4	49.6420	52.8640	46.7780	49.6420
5	47.4940	47.2554	49.7613	47.4940

4.3.10 ผลการทดลองโดยใช้ชุดข้อมูล Wine

4.3.10.1 ผลการทดลองโดยใช้สมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง

ตารางที่ 4.74 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	79.2135	89.3258	84.2697	79.2135
2	74.1573	73.5955	83.7079	74.1573
3	73.5955	73.5955	76.9663	73.5955
4	73.0337	73.0337	75.8427	73.0337
5	73.0337	73.0337	74.7191	73.0337

ตารางที่ 4.75 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการระยะทางระหว่างตำแหน่ง 2 ตำแหน่ง โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	83.7209	87.7907	87.7907	83.7209
2	73.2558	82.5581	82.5581	73.2558
3	72.0930	86.6279	85.4651	72.0930
4	62.7907	72.6744	60.4651	62.7907
5	62.2093	68.0233	80.8140	62.2093

4.3.10.2 ผลการทดลองโดยใช้สมการ Mean Absolute Deviation (MAD)

ตารางที่ 4.76 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	73.5955	74.7191	85.9551	73.5955
2	73.5955	74.1573	70.7865	73.5955
3	72.4719	75.2809	84.8315	72.4719
4	72.4719	71.9101	64.6067	72.4719
5	70.7865	71.3483	81.4607	70.7865

ตารางที่ 4.77 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Deviation (MAD) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	78.4884	85.4651	76.7442	78.4884
2	77.9070	87.7907	81.9768	77.9070
3	76.1628	84.8837	83.1395	76.1628
4	75.5814	79.6512	67.4419	75.5814
5	75.0000	92.4419	76.1628	75.0000

4.3.10.3 ผลการทดลองโดยใช้สมการ Mean Absolute Percent Error (MAPE)

ตารางที่ 4.78 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	73.0337	74.1573	89.3258	73.0337
2	73.0337	73.0337	86.5169	73.0337
3	71.9101	71.9101	93.2584	71.9101
4	71.9101	71.9101	71.3483	71.9101
5	70.2247	72.4719	89.3258	70.2247

ตารางที่ 4.79 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Absolute Percent Error (MAPE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	86.6279	91.2791	87.7907	86.6279
2	80.8140	83.7209	86.0465	80.8140
3	76.7442	71.5116	65.1163	76.7442
4	72.6744	84.8837	89.5349	72.6744
5	68.6047	75.5814	85.4651	68.6047

4.3.10.4 ผลการทดลองโดยใช้สมการ Mean Square Error (MSE)

ตารางที่ 4.80 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ปกติ

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	71.9101	71.9101	84.8315	71.9101
2	71.3483	74.1573	79.7753	71.3483
3	69.6629	70.7865	86.5169	69.6629
4	69.6629	70.2247	61.7978	69.6629
5	64.6067	66.2921	76.4045	64.6067

ตารางที่ 4.81 ผลการจำแนกประเภทข้อมูลชุดข้อมูล Wine ด้วยสมการ Mean Square Error (MSE) โดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อนหรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน

รอบที่	ความถูกต้องในการจำแนกประเภทข้อมูล (%)			
	สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
1	75.5814	86.6279	80.2326	75.5814
2	74.4186	78.4884	88.3721	74.4186
3	72.6744	77.9070	80.8140	72.6744
4	66.8605	75.5814	64.5349	66.8605
5	63.3721	75.0000	75.5814	63.3721

4.4 ค่าเฉลี่ยของผลการทดลองของแต่ละวิธีทดลอง

จากผลทดลองทั้งหมด สามารถนำมาสรุปเป็นค่าเฉลี่ยของผลการทดลอง เพื่อให้สามารถมองเป็นภาพรวม และเปรียบเทียบประสิทธิภาพในการจำแนกประเภทข้อมูล ของแต่ละวิธีได้

ตารางที่ 4.82 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ปกติ เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูล ทั้ง 4 สมการ

ชุดข้อมูล	สมการที่ใช้ในขั้นตอนการหาค่าความเหมาะสม	ค่าเฉลี่ยของความถูกต้องในการจำแนกประเภทข้อมูล (%)			
		สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
Balance-Scale	DIS	68.9600	63.9360	54.2720	68.9600
	MAD	64.8000	62.8480	49.8560	64.8000
	MAPE	63.6800	58.0480	58.4320	63.6800
	MSE	70.6240	65.1840	65.4720	70.6240
Blood-Transfusion	DIS	72.8877	71.7380	65.6685	72.0053
	MAD	77.0321	76.7914	49.6791	77.1925
	MAPE	75.9626	74.6791	60.0267	74.8663
	MSE	70.1872	70.5882	54.6791	70.5080
Glass	DIS	49.7196	49.8131	1.2150	49.7196
	MAD	51.4019	51.9626	0.0000	51.4019
	MAPE	50.7477	50.2804	6.5421	50.7477
	MSE	50.4673	50.5607	0.0000	50.4673
Haberman	DIS	74.5098	73.5294	34.2484	74.5098
	MAD	73.2680	75.8170	37.1242	73.2680
	MAPE	56.9935	55.0327	53.3333	56.9935
	MSE	74.9673	74.4444	39.9346	74.9673
Iris	DIS	94.1333	93.0667	89.2000	94.1333
	MAD	93.6000	95.4667	94.5333	93.6000
	MAPE	90.6667	87.6000	85.3333	90.6667
	MSE	91.3333	91.0667	92.8000	91.3333

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับบุคคลในแผนกหรือองค์กรศึกษาเท่านั้น ไม่อนุญาตให้นำไปเผยแพร่หรือใช้ขึ้นต้นการวิจัย

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.83 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ปกติ เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูลทั้ง 4 สมการ (ต่อ)

ชุดข้อมูล	สมการที่ใช้ในขั้นตอนการหาค่าความเหมาะสม	ค่าเฉลี่ยของความถูกต้องในการจำแนกประเภทข้อมูล (%)			
		สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
Liver-disorders	DIS	55.5362	56.1159	55.1884	55.5362
	MAD	60.3478	62.1449	58.1449	60.3478
	MAPE	56.2899	55.4203	55.2464	56.2899
	MSE	60.1739	57.2754	52.2319	60.1739
Pima Indians diabetes	DIS	72.8906	72.4479	0.0000	72.8906
	MAD	66.4323	66.5104	26.9792	66.4323
	MAPE	68.4896	69.1146	31.0938	68.4896
	MSE	72.0833	71.0677	19.4010	72.0833
Sonar	DIS	60.0962	61.4423	49.4231	60.0962
	MAD	63.3654	65.1923	59.5192	63.3654
	MAPE	58.5577	60.0962	66.9231	58.5577
	MSE	58.3654	60.8654	55.1923	58.3654
Vehicle	DIS	42.7896	43.2861	37.4704	42.7896
	MAD	40.3073	40.0946	34.2317	40.3073
	MAPE	42.0567	39.8345	34.6336	42.0567
	MSE	44.0662	42.8132	32.0095	44.0662
Wine	DIS	74.6067	76.5169	79.1011	74.6067
	MAD	72.5843	73.4831	77.5281	72.5843
	MAPE	72.0225	72.6966	85.9551	72.0225
	MSE	69.4382	70.6742	77.8652	69.4382

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.84 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงฯ เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูลทั้ง 4 สมการ

ชุดข้อมูล	สมการที่ใช้ใน ขั้นตอนการหา ค่าความ เหมาะสม	ค่าเฉลี่ยของความถูกต้องในการจำแนกประเภทข้อมูล (%)			
		สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
Balance-Scale	DIS	81.2924	72.8595	80.0323	81.2924
	MAD	79.3538	71.5024	62.9079	79.3538
	MAPD	72.2132	64.9435	70.8239	72.2132
	MSD	74.7981	66.4297	64.3619	74.7981
Blood- Transfusion	DIS	74.6237	72.7688	59.0323	74.6237
	MAD	60.2151	58.0376	51.8011	60.2151
	MAPD	67.4462	68.1989	65.0538	67.4462
	MSD	75.4032	70.6183	49.7849	75.4032
Glass	DIS	76.4356	75.6436	5.3465	76.4356
	MAD	76.2376	77.1287	0.0000	76.2376
	MAPD	70.8911	72.0792	5.3465	70.8911
	MSD	73.5644	71.2871	8.0198	73.5644
Haberman	DIS	60.0662	59.4702	57.9470	60.0662
	MAD	66.6887	69.9338	46.8212	66.6887
	MAPD	73.9735	74.2384	46.7550	73.9735
	MSD	77.3510	76.8874	35.8940	77.3510
Iris	DIS	97.6389	96.5278	94.4444	97.6389
	MAD	98.6111	97.7778	96.2500	98.6111
	MAPD	98.8889	98.8889	97.7778	98.8889
	MSD	99.8611	99.1667	99.3056	99.8611

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.85 ค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงฯ เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูลทั้ง 4 สมการ (ต่อ)

ชุดข้อมูล	สมการที่ใช้ในขั้นตอนการหาค่าความเหมาะสม	ค่าเฉลี่ยของความถูกต้องในการจำแนกประเภทข้อมูล (%)			
		สมการ DIS	สมการ MAD	สมการ MAPE	สมการ MSE
Liver-disorders	DIS	58.8856	58.0059	52.6686	58.8856
	MAD	58.4164	58.0645	56.1877	58.4164
	MAPD	62.5220	61.5249	61.6422	62.5220
	MSD	62.4047	62.0528	56.4223	62.4047
Pima Indians diabetes	DIS	62.8796	64.0052	59.1623	62.8796
	MAD	56.6754	57.8796	60.8639	56.6754
	MAPD	65.0262	65.8115	50.6283	65.0262
	MSD	58.5079	62.4607	49.2147	58.5079
Sonar	DIS	65.0000	64.1176	70.8824	65.0000
	MAD	63.1373	63.9216	61.9608	63.1373
	MAPD	61.5686	63.9216	67.4510	61.5686
	MSD	66.8627	70.0980	65.4902	66.8627
Vehicle	DIS	51.5513	52.4821	46.8019	51.5513
	MAD	50.7876	53.6516	51.9570	50.7876
	MAPD	51.4320	51.9809	46.4439	51.4320
	MSD	51.9093	52.9594	46.0382	51.9093
Wine	DIS	70.8140	79.5349	79.4186	70.8140
	MAD	76.6279	86.0465	77.0930	76.6279
	MAPD	77.0930	81.3954	82.7907	77.0930
	MSD	70.5814	78.7209	77.9070	70.5814

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.86 ค่าเฉลี่ยผลการทดลอง เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างชุดข้อมูลเรียนรู้ปกติ และชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงฯ

ชุดข้อมูล	ค่าเฉลี่ยของความถูกต้องในการจำแนกประเภทข้อมูล (%)		ประสิทธิภาพในการจำแนกประเภทข้อมูล(%)
	ชุดข้อมูลเรียนรู้ปกติ	ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อนแบบเลือกใช้ตัวกรองคำมัธยฐาน	
Balance-Scale	63.3860	73.0735	9.6875
Blood-Transfusion	69.6557	65.6670	-3.9887
Glass	38.4404	56.8193	18.3789
Haberman	62.6838	64.0066	1.3228
Iris	91.7833	98.1337	6.3503
Liver-disorders	57.2790	59.4391	2.1602
Pima Indians diabetes	57.2754	59.7628	2.4874
Sonar	59.9639	65.0613	5.0973
Vehicle	40.1758	50.8547	10.6789
Wine	74.4452	77.0712	2.6260

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 วิจารณ์ผลการทดลอง

จากตารางค่าเฉลี่ยผลการทดลองเมื่อใช้ชุดข้อมูลเรียนรู้ปกติและชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงโดยเทคนิคการกรองสัญญาณเชิงซ้อนแบบเลือกใช้ตัวกรองค่ามัธยฐาน เพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างสมการที่ใช้ในการจำแนกประเภทของข้อมูลทั้ง 4 สมการจะเห็นได้ว่า การใช้สมการ DIS และ MSE จะให้ค่าสำหรับการจำแนกประเภทของข้อมูลได้เท่ากัน และประสิทธิภาพในการจำแนกประเภทของข้อมูลของทั้ง 4 สมการนั้นจะให้ค่าประสิทธิภาพในการจำแนกประเภทของข้อมูลใกล้เคียงกัน ยกเว้นการใช้สมการ MAPE กับชุดข้อมูล Glass จะให้ค่าประสิทธิภาพในการจำแนกประเภทของข้อมูลได้น้อยกว่าสมการอื่น ๆ

จากตารางค่าเฉลี่ยผลการทดลองของเพื่อเปรียบเทียบประสิทธิภาพในการจำแนกประเภทของข้อมูลระหว่างชุดข้อมูลเรียนรู้ปกติ และชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุง จะเห็นว่าประสิทธิภาพในการจำแนกประเภทของข้อมูลของชุดข้อมูล Balance-Scale, Glass, Sonar และ Vehicle เมื่อใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐานนั้น มีประสิทธิภาพในการจำแนกประเภทข้อมูลค่อนข้างมาก เมื่อพิจารณาไปถึงชุดข้อมูลตั้งต้นจะพบว่า ในแต่ละประเภทของข้อมูลนั้นจะมีค่าความแตกต่างของชุดข้อมูลอยู่ค่อนข้างมาก และมีความแตกต่างในแต่ละคุณสมบัติ เมื่อผ่านตัวกรองแล้วชุดข้อมูลจะมีค่าในแต่ละคุณสมบัติที่มีความต่างกันลดลง มักจะรวมกลุ่มอยู่ในแนวตรงกึ่งกลางของชุดข้อมูลในแต่ละคุณลักษณะ เมื่อเจเนติกอัลกอริทึมมาทำการเรียนรู้จะทำให้ได้โครโมโซมต้นแบบที่สามารถนำไปจำแนกประเภทข้อมูลได้มีประสิทธิภาพมากยิ่งขึ้น

ส่วนในชุดข้อมูล Iris นั้น ในแต่ละประเภทจะมีคุณลักษณะที่ต่างกันมาก ทำให้เมื่อใช้เจเนติกอัลกอริทึมมาทำการจำแนกประเภทข้อมูลจะเห็นได้ว่า มีประสิทธิภาพที่สูง แต่อย่างไรก็ตามเมื่อใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงด้วยเทคนิคการกรองสัญญาณเชิงซ้อนแล้ว ยิ่งทำให้ประสิทธิภาพในการจำแนกประเภทข้อมูลมีความแม่นยำสูงมากยิ่งขึ้น ในบางครั้งของการทดลองที่ได้ค่าความถูกต้องในการจำแนกประเภทข้อมูลเป็น 100%

ส่วนในชุดข้อมูล Blood-Transfusion, Haberman, Liver-disorders และ Wine ผลการทดลองการจำแนกประเภทข้อมูลระหว่างชุดข้อมูลเรียนรู้ปกติ และชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน มีประสิทธิภาพที่ไม่แตกต่างกันมาก

และชุดข้อมูล Pima Indians diabetes การจำแนกประเภทข้อมูลโดยใช้ชุดข้อมูลเรียนรู้ที่ผ่านการปรับปรุงด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน ไม่ทำให้ประสิทธิภาพในการจำแนกประเภทข้อมูลเพิ่มขึ้น ในทางกลับกันยังทำให้การจำแนกประเภทข้อมูลโดยใช้เจเนติกอัลกอริทึมมีประสิทธิภาพลดลง เมื่อเข้าไปพิจารณา จะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พบว่าข้อมูลพบว่า ในแต่ละประเภทมีคุณลักษณะที่ไม่แตกต่างกันมาก และในแต่ละคุณลักษณะจะมีข้อมูลที่มีลักษณะเป็นการกระจายค่อนข้างสูง นอกจากนี้ยังพบว่าข้อมูลระหว่างคุณลักษณะมีความแตกต่างกันมาก ทำให้เมื่อใช้ตัวกรองแบบค่ามัธยฐานแล้ว ชุดข้อมูลที่ได้จะมีความแตกต่างจากชุดข้อมูลตั้งต้นมาก ส่งผลให้โครโมโซมต้นแบบที่ได้ จะมีค่าในแต่ละยีนที่ต่างจากชุดข้อมูลปกติมาก ส่งผลให้ประสิทธิภาพในการจำแนกประเภทข้อมูลลดลง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการทดลอง และข้อเสนอแนะ

5.1 สรุปผลการทดลอง

การนำหลักการในการปรับปรุงคุณภาพของภาพ โดยการกำจัดสิ่งรบกวนในภาพในโดเมนตำแหน่ง ด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน มาปรับปรุงชุดข้อมูลเรียนรู้ในขั้นตอนการเรียนรู้ ของการจำแนกประเภทข้อมูล โดยใช้เจเนติกอัลกอริทึมนี้ สามารถสร้างโครโมโซมต้นแบบที่มีความสามารถในการจำแนกประเภทข้อมูลได้มีประสิทธิภาพมากขึ้นให้กับชุดข้อมูลที่มีข้อมูลระหว่างประเภทที่มีคุณลักษณะแตกต่างกัน และมีข้อมูลระหว่างคุณลักษณะไม่แตกต่างกันมาก ซึ่งสอดคล้องกับเทคนิคการปรับปรุงคุณภาพ ๆ ที่การปรับปรุงคุณภาพของภาพนั้น ไม่ได้ระบุเฉพาะเจาะจงว่าเป็นวิธีใด แต่ขึ้นอยู่กับภาพนั้น ๆว่าจะเหมาะกับวิธีใด เช่นเดียวกับชุดข้อมูล ที่บางประเภทการปรับปรุงโดยใช้วิธีการกำจัดสิ่งรบกวนในภาพในโดเมนตำแหน่ง ด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน สามารถเพิ่มประสิทธิภาพการจำแนกประเภทของข้อมูลได้มาก แต่กับบางชุดข้อมูลที่ไม่เหมาะสมก็ไม่ให้ประสิทธิภาพในการจำแนกประเภทของข้อมูลเพิ่มขึ้น

5.2 ข้อเสนอแนะ

ในการทดลองการจำแนกประเภทข้อมูลในครั้งนี้พบว่าสามารถนำเทคนิคในการปรับปรุงคุณภาพของภาพ มาประยุกต์ใช้เพื่อเพิ่มประสิทธิภาพในการจำแนกประเภทของข้อมูลได้ โดยมุ่งใช้เทคนิคการกำจัดสิ่งรบกวนในภาพในโดเมนตำแหน่ง ด้วยเทคนิคการกรองสัญญาณเชิงซ้อน หรือตัวกรองลำดับสถิติ แบบเลือกใช้ตัวกรองค่ามัธยฐาน ซึ่งได้ผลลัพธ์เป็นที่น่าพอใจกับชุดข้อมูลที่ข้อมูลในระหว่างคุณลักษณะของชุดข้อมูลประเภทเดียวกัน ไม่แตกต่างกันมาก การทดลองในครั้งต่อไป อาจนำเสนอการใช้เทคนิคการปรับปรุงคุณภาพด้วยเทคนิคอื่น ๆ มาทำการปรับปรุงในชุดข้อมูลเรียนรู้ซึ่ง เพื่อให้การจำแนกประเภทของข้อมูลมีประสิทธิภาพมากยิ่งขึ้น หรือสามารถใช้ได้กับชุดข้อมูลคุณลักษณะอื่น ๆ

บรรณานุกรม

- จิตรภรณ์ มวลวงศ์. 2550. "การจัดกลุ่มข้อมูลโดยใช้วิวัฒนาการทางด้านพฤติกรรมทางสังคมของมนุษย์." วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- จิราภรณ์ มั่นศิลป์. 2553. "กระบวนการปรับปรุงภาพและการประยุกต์ใช้กับนิติวิทยาศาสตร์." **Veridian E – Journal, Silpakorn University.** 3(1) : 245-261.
- บุญญาพร เข้มปัญญา. 2555. "การคัดเลือกลักษณะด้วยอัตราส่วนเกนสำหรับการจำแนกประเภทโดยใช้ขั้นตอนวิธีเจเนติก." วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาการคอมพิวเตอร์ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- พรชัย ภาวรงค์ศักดิ์. 2000. การประมวลผลสัญญาณดิจิทัลเบื้องต้น. [Online] Available: www.kmitl.ac.th/~kksasems/dsp/DSP_R10.pdf.
- ศุภวีร์ สายพิมพ์. 2548. "การจัดกลุ่มข้อมูลโดยใช้จีเน็ตออลกอริทึม" วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- อรณัทร จิตต์โสภักตร์. 2552. **Digital Image Processing**. กรุงเทพฯ: สงวนกิจ พรินท์ แอนด์ มีเดีย.
- อุไรวรรณ กะกุลพิมพ์. 2549. "การนำประวัติของโครโมโซมมาใช้ในการปรับปรุงการจัดกลุ่มด้วยจีเน็ตออลกอริทึม." วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- Bandyopadhyay, Sanghamitra. 2007. **Classification and learning using genetic algorithms [electronic resource] : applications in bioinformatics and web intelligence**. Berlin: Springer.
- Berson, Alex. 1997. **Data warehousing, data mining, and OLAP**. New York: McGraw-Hill.
- Dubitzky, Werner, ed. 2007. **Fundamentals of data mining in genomics and proteomics**. New York: Springer.
- Jordi Corbilla. 2010. **Genetic Algorithm (GA)**. [Online] Available: <http://thundaxsoftware.blogspot.com/2010/11/genetic-algorithms-ga.html>.
- "Korkut Koray. 2004. **Mining Classification Rules by Using Genetic Algorithms with Non-random Initial Population and Uniform Operator**. [Online] Available: <http://www.lib.kmitl.ac.th/central/index.php?lang=th>.

Muntean, M. 2010. **Learning classification rules with genetic algorithm.** [Online] Available:

<http://www.lib.kmitl.ac.th/central/index.php?lang=th>.

Ricardo A. Losada. 2008. **Digital Filters with MATLAB.** [Online] Available:

http://www.mathworks.com/tagteam/55876_digfilt.pdf.

Satchidananda Dehuri. 2006. **Genetic Algorithms for Multi-Criterion Classification and Clustering in Data Mining.** [Online] Available:

<http://www.lib.kmitl.ac.th/central/index.php?lang=th>.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้