

วิธีการลดมิติแบบผสมผสานโดยใช้ NIA_s-CCA/DCCA
สำหรับหลายชุดข้อมูลของการวิเคราะห์การนอน

HYBRID DIMENSIONALITY REDUCTION BASED ON NIA_s-CCA/DCCA
FOR MULTISETS OF SLEEP ANALYSIS



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2562

KMITL-2019-EN-D-018-024

วิธีการลดมิติแบบผสมผสานโดยใช้ NIA_s-CCA/DCCA
สำหรับหลายชุดข้อมูลของการวิเคราะห์การนอน

HYBRID DIMENSIONALITY REDUCTION BASED ON NIA_s-CCA/DCCA
FOR MULTISETS OF SLEEP ANALYSIS

พิมพ์พร เหมยน้อย
PIMPORN MOEYNOI

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรดุษฎีบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้า

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2562

KMITL-2019-EN-D-018-024

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

HYBRID DIMENSIONALITY REDUCTION BASED ON NIAS-CCA/DCCA
FOR MULTISSETS OF SLEEP ANALYSIS



PIMPORN MOEYNOI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
DOCTOR OF ENGINEERING IN ELECTRICAL ENGINEERING
FACULTY OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2019

KMITL-2019-EN-D-018-024

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2019

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	วิธีการลดมิติแบบผสมผสานโดยใช้ NIAs-CCA/DCCA สำหรับหลายชุดข้อมูลของการวิเคราะห์การนอน
นักศึกษา	นางสาวพิมพ์พร เหมยน้อย
รหัสประจำตัว	56601036
ปริญญา	วิศวกรรมศาสตรดุษฎีบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้า
พ.ศ.	2562
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร.ยุทธนา คิดใจเดียว

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้นำเสนออัลกอริทึมการลดมิติแบบผสมผสานใหม่สำหรับหลายชุดข้อมูลซึ่งวิเคราะห์พร้อมกับการจำแนกกระยะการนอนและการจำแนกภาวะหยุดหายใจขณะนอนหลับ การวิเคราะห์การนอนเป็นหัวข้องานวิจัยที่น่าสนใจและต้องการพัฒนาการออกแบบอัลกอริทึมเพื่อนำไปประยุกต์ใช้อย่างมีประสิทธิภาพสูงสุด โดยทั่วไปการวิเคราะห์การนอนด้วยหลายสัญญาณเป็นวิธีที่สามารถวิเคราะห์การนอนได้อย่างมีประสิทธิภาพ แต่อย่างไรก็ตามจำนวนของตัวแปรที่มากนั้นเป็นสาเหตุของการเกิดปัญหามิติข้อมูลสูง เพื่อการแก้ไขปัญหาดังกล่าวนั้นวิทยานิพนธ์ฉบับนี้มุ่งเน้นการออกแบบแต่ละขั้นตอนของการวิเคราะห์การนอนแบบอัตโนมัติอย่างเหมาะสม กระบวนการของการวิเคราะห์การนอนแบบอัตโนมัติประกอบด้วย 4 ขั้นตอน ได้แก่ ขั้นตอนการประมวลผลก่อน ขั้นตอนการสกัดคุณลักษณะ ขั้นตอนการลดมิติ และขั้นตอนการจำแนกข้อมูล วิทยานิพนธ์ฉบับนี้ได้นำเสนออัลกอริทึมการเลือกคุณลักษณะใหม่ที่เรียกว่า DisBGWO ซึ่งเป็นวิธีที่ดัดแปลงมาจากวิธีเลือกคุณลักษณะด้วยวิธีค้นหาค่าตอบที่เหมาะสมด้วยแนวคิดของฝูงหมาป่าสีเทาแบบเดิมและวิธีการวิวัฒนาการโดยใช้ความต่าง นอกจากนี้วิทยานิพนธ์ฉบับนี้ได้นำเสนอวิธีการลดมิติแบบผสมผสานแบบใหม่ที่เรียกว่า CCA/DCCA-NIAs และ NIAs-CCA/DCCA ซึ่งทั้งสองอัลกอริทึมเป็นการหลอมรวมระหว่างวิธีการเลือกคุณลักษณะและวิธีการแปลงคุณลักษณะ การทดสอบประสิทธิภาพและเปรียบเทียบวิธีการลดมิติที่วิทยานิพนธ์ฉบับนี้นำเสนอกับวิธีการลดมิติอื่นๆ ด้วยการใช้อัตราการนอนจากฐานข้อมูล SHHS, ฐานข้อมูล Dream database และฐานข้อมูล Physionet จากผลการทดลองแสดงให้เห็นว่าวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA เป็นวิธีที่ดีที่สุดด้วยค่าความถูกต้องแม่นยำของการวิเคราะห์การนอนที่สูง และใช้เวลาในการประมวลผลที่ดีกว่าวิธีการอื่นๆ ยิ่งไปกว่านั้นวิธีที่วิทยานิพนธ์ฉบับนี้เสนอสามารถระบุภาวะหยุดหายใจขณะนอนได้อย่างมีประสิทธิภาพทั้งในกรณีที่ใช้หลายสัญญาณและใช้เพียงสัญญาณเดียว

Thesis	Hybrid Dimensionality Reduction based on NIAs-CCA/DCCA for Multisets of Sleep Analysis
Student	Miss.Pimporn Moeynoi
Student ID.	56601036
Degree	Doctor of Engineering
Program	Electrical Engineering
Year	2019
Thesis Advisor	Asst. Prof. Dr. Yuttana Kidjaidure

ABSTRACT

This thesis presents a new hybrid dimensionality reduction algorithm for multi sets of sleep stage classification and sleep apnea disorder detection. Automatic sleep analysis is popular research and demands to develop an algorithm for maximum efficiency. Normally, the automatic sleep analysis using various signals is an important tool that has ability to analyze sleep effectively. However, the numerous variables from multi signals lead to the high dimension problems. To alleviate problem, this thesis focuses on designing each step of automatic system appropriately. A procedure of automatic sleep analysis system has four steps including the data acquisition, the feature extraction, the dimensionality reduction, and the classification. This research proposes a new method for feature selection called the DisBGWO algorithm modified from conventional binary grey wolf optimization and the differential evolution technique. Moreover, we propose a new hybrid dimensionality reduction called CCA/DCCA-NIAs and NIAs-CCA/DCCA. Both methods are the combination between the feature selection method and the feature transformation. This research evaluates the performance of our proposed method and other methods using sleep data sets from SHHS database, Dream database, and Physionet database. Based on experiment results, the DisBGWO-DCCA is the best algorithm providing high accuracy of classification and achieves the better computation cost than the conventional approaches. Moreover, the proposed method can effectively identify apnea event in case of multi-signals and single-signal.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยดี ด้วยความกรุณาจากอาจารย์ที่ปรึกษาผู้ช่วยศาสตราจารย์ ดร.ยุทธนา คิดใจเดียว ที่ให้ความรู้ ช่วยเหลือ และชี้แนะแนวทางในการแก้ไขปัญหาที่เกิดขึ้นในงานวิจัย และกระบวนการทำงานวิจัย อีกทั้งฝึกฝนให้ข้าพเจ้ามีกระบวนการ คิด วิเคราะห์ อย่างมีระบบและเป็น เหตุเป็นผล ตลอดจนชี้แนะแนวทางในการดำรงชีวิตนักศึกษา ฝึกสอนให้มีความอดทน ความ เอื้อเฟื้อเผื่อแผ่ หลอมรวมคุณธรรม จริยธรรม เพื่อให้ดำเนินงานอย่างนักวิจัยที่ดีในสังคม

ขอขอบคุณบิดา มารดาที่คอยสนับสนุนเงินทุนในการเล่าเรียนและการทำงานวิจัย ตลอดจนคอย เป็นให้กำลังใจข้าพเจ้าให้เสมอมา ขอขอบคุณเพื่อนๆ พี่ๆ และน้องๆ ทุกคนที่ให้กำลังใจและให้ความ ร่วมมือในการทำงานวิจัย

ขอขอบคุณห้องปฏิบัติการ B408 และภาควิชาอิเล็กทรอนิกส์ คณะวิศวกรรมศาสตร์ สถาบัน เทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่อำนวยความสะดวก เอื้อเฟื้อสถานที่และอุปกรณ์ใน การทำงานวิจัยต่างๆ

สำหรับคุณงามความดีอันใดที่เกิดจากวิทยานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบให้กับบิดา มารดา ซึ่ง เป็นที่รักและเคารพยิ่ง ครูบาอาจารย์ที่เคารพ ตลอดจนผู้มีพระคุณทุกท่าน

พิมพ์พร เหมยนต์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	IX
สารบัญรูป.....	XI
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของงานวิจัย	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	2
1.3 สมมุติฐานของการศึกษา.....	2
1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย.....	3
1.5 ขอบเขตการวิจัย.....	4
1.6 ขั้นตอนของการศึกษา.....	5
บทที่ 2 หลักการ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	6
2.1 ระยะเวลาการนอนหลับ (Sleep stage).....	6
2.2 ภาวะหยุดหายใจขณะนอนหลับ Sleep disorder.....	9
2.3 สัญญาณไฟฟ้าร่างกายที่เกี่ยวข้องกับการนอน.....	11
2.3.1 สัญญาณไฟฟ้าสมองหรือสัญญาณ EEG.....	11
2.3.2 สัญญาณไฟฟ้าการเคลื่อนไหวของกล้ามเนื้อลูกตาหรือสัญญาณ EOG.....	17
2.3.3 สัญญาณไฟฟ้าหัวใจหรือสัญญาณ ECG.....	19
2.3.3.1 การวิเคราะห์ความแปรปรวนของอัตราการเต้นของหัวใจหรือ Heart Rate Variability Analysis (HRV)	20
2.3.3.2 การสกัดสัญญาณการหายใจจากสัญญาณ ECG หรือ Electrocardiography Derived Respiratory (EDR)	27
2.4 งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์การนอนแบบอัตโนมัติ.....	29

สารบัญ(ต่อ)

	หน้า
2.4.1 การวิเคราะห์การนอนแบบอัตโนมัติด้วยชุดสัญญาณเดียว.....	29
2.4.2 การวิเคราะห์การนอนแบบอัตโนมัติด้วยสัญญาณหลายชุด.....	31
2.5 การออกแบบระบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง (Machine learning).....	35
2.6 การลดมิติข้อมูล (Dimensionality reduction).....	36
2.6.1 การแปลงคุณลักษณะ (Feature transformation).....	36
2.6.1.1 การวิเคราะห์หาค่าประกอบหลัก (Principle Component Analysis).....	37
2.6.1.2 การวิเคราะห์แยกแยะเชิงเส้น LDA (Linear Discriminate Analysis).....	37
2.6.1.3 การวิเคราะห์สหสัมพันธ์คาโนนิกอล (Canonical Correlation Analysis)..	37
2.6.1.4 การวิเคราะห์ความสัมพันธ์สหสัมพันธ์จำแนกกลุ่ม (Discriminant Canonical Correlation Analysis)	38
2.6.2 วิธีการเลือกคุณลักษณะ (Feature selection).....	42
2.7 การลดมิติแบบผสมผสาน (Hybrid Dimensionality reduction).....	44
2.8 อัลกอริทึมที่ได้แรงบันดาลใจจากธรรมชาติ (Nature Inspired Algorithms).....	45
2.8.1 อัลกอริทึม Binary Particle Swarm Optimization (BPSO).....	46
2.8.2 อัลกอริทึมการค้นหาแบบนกกาเหว่า Cuckoo Search (CS).....	47
2.8.3 อัลกอริทึม Binary Gray Wolf Optimization (BGWO).....	48
2.9 สัมประสิทธิ์ความแตกต่างของแจ็กการ์ด (Jaccard Dissimilarity Coefficient)	49
2.10 การจำแนกข้อมูลด้วยอัลกอริทึมต่างๆ.....	50
2.10.1 การจำแนกข้อมูลด้วยอัลกอริทึม k-Nearest Neighbors (k-NN).....	50
2.10.2 การจำแนกข้อมูลด้วยอัลกอริทึม Decision Tree (DT).....	51
2.10.3 การจำแนกข้อมูลด้วยอัลกอริทึม Random Forest (RF).....	53
2.10.4 การจำแนกข้อมูลด้วยอัลกอริทึม Support Vector Machine (SVM).....	54
บทที่ 3 การออกแบบระบบการวิเคราะห์การนอนอัตโนมัติ.....	56
3.1 ขั้นตอนการประมวลผลก่อน (Preprocessing)	56
3.1.1 ขั้นตอนการสกัดคุณลักษณะจากสัญญาณ EEG, EOG และ ECG.....	57

สารบัญ(ต่อ)

	หน้า
3.1.2 ขั้นตอนการลดมิติข้อมูล.....	57
3.1.2.1 วิธีการลดมิติข้อมูลด้วยการแปลงข้อมูล.....	57
1. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ PCA.....	58
2. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ LDA.....	58
3. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ CCA.....	59
4. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ DCCA.....	59
3.1.2.2 วิธีการลดมิติข้อมูลด้วยการเลือกคุณลักษณะแบบวิธีแรปเปอร์.....	60
1. วิธีการเลือกคุณลักษณะด้วย BPSO.....	61
2. วิธีการเลือกคุณลักษณะด้วย BCS.....	62
3. วิธีการเลือกคุณลักษณะด้วย EBCS.....	62
4. วิธีการเลือกคุณลักษณะด้วย BGWO.....	64
5. วิธีการเลือกคุณลักษณะด้วย DisBGWO.....	64
3.1.2.3 วิธีการลดมิติแบบผสมผสานด้วยอัลกอริทึมที่นำเสนอ.....	66
1. วิธีการลดมิติแบบผสมผสาน CCA/DCCA-NIAs.....	66
2. วิธีการลดมิติแบบผสมผสาน NIAs-CCA/DCCA.....	68
3.3 ขั้นตอนการประมวลผล.....	70
3.3.1 การประเมินประสิทธิภาพของอัลกอริทึม.....	71
3.3.1.1 ค่าเฉลี่ยความถูกต้องแม่นยำ.....	71
3.3.1.2 ค่าเฉลี่ยความไว.....	71
3.3.1.3 ค่าเฉลี่ยความจำเพาะ.....	72
3.3.1.4 ค่าความแปรปรวนของผลการประเมิน.....	72
บทที่ 4 วิธีการทดลอง ผลการทดลอง และวิเคราะห์ผลการทดลอง.....	73
4.1 การทดลองการจำแนกกระยะการนอนแบบอัตโนมัติ.....	73
4.1.1 ขั้นตอนการเตรียมข้อมูล.....	73
4.1.2 การทดลองและผลการทดลองการทดสอบประสิทธิภาพของวิธีการลดมิติข้อมูลแบบเดี่ยว สำหรับการจำแนกกระยะการนอนด้วยหลายชุดสัญญาณ.....	74

สารบัญ(ต่อ)

หน้า

4.1.2.1 การทดลองและผลการทดลองของการทดสอบประสิทธิภาพของวิธีการลตมิติข้อมูลด้วยการแปลงคุณลักษณะ.....	74
1. การทดลองที่ 1 การทดลองเปรียบเทียบประสิทธิภาพของวิธีการลตมิติด้วยการแปลงคุณลักษณะสำหรับการจำแนกระยะการนอนโดยใช้สัญญาณ EEG EOG และ ECG.....	75
2. ผลการทดลองและวิเคราะห์ผลทดลองขั้นตอนการเรียนรู้การทดลองที่ 1.....	75
3. ผลการทดลองและวิเคราะห์ผลทดลองขั้นตอนการทดสอบของการทดลองที่ 1.....	81
4.1.2.2 การทดลองและผลการทดลองของประสิทธิภาพของวิธีการลตมิติข้อมูลด้วยการเลือกคุณลักษณะ.....	85
1. การทดลองที่ 2 การทดลองเปรียบเทียบประสิทธิภาพของวิธีการลตมิติด้วยการเลือกคุณลักษณะสำหรับการจำแนกระยะการนอนจากสัญญาณ EEG EOG และ ECG...	85
2. ผลการทดลองและวิเคราะห์การทดลองขั้นตอนการเรียนรู้การทดลองที่ 2.....	88
3. ผลการทดลองและวิเคราะห์ผลทดลองของขั้นตอนการทดสอบการทดลองที่ 2.....	90
4.1.2.3 การทดลองและผลการทดลองของประสิทธิภาพของการลตมิติแบบผสมผสานที่นำเสนอ.....	91
1. การทดลองที่ 3 การทดลองเปรียบเทียบประสิทธิภาพของวิธีการลตมิติแบบผสมผสาน CCA/DCCA-NIAs สำหรับการจำแนกระยะการนอนโดยใช้สัญญาณ EEG EOG และ ECG.....	92
2. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการเรียนรู้ของทดลองที่ 3.....	93
3. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการทดสอบทดลองที่ 3.....	95
4. การทดลองที่ 4 การทดลองเปรียบเทียบประสิทธิภาพของการลตมิติแบบผสมผสานด้วยวิธี NIAs-CCA/DCCA สำหรับการจำแนกระยะการนอนโดยใช้สัญญาณ EEG EOG และ ECG.....	97
5. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการเรียนรู้ทดลองที่ 4.....	98
6. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการทดสอบทดลองที่ 4...	101

สารบัญ(ต่อ)

หน้า

7. การทดลองที่ 4.1 การทดลองเปรียบเทียบประสิทธิภาพของการลดมิติแบบผสมผสานของ DisBGWO-DCCA ด้วยเงื่อนไขการแบ่งข้อมูลที่แตกต่างกัน.....	103
8. ผลการทดลองและวิเคราะห์ผลการทดลองที่ 4.1.....	104
9. การทดลองที่ 5 การทดสอบประสิทธิภาพการลดมิติแบบผสมผสานของ DisBGWO-DCCA สำหรับการจำแนกระยะการนอนจากสัญญาณ EEG EOG และ ECG.....	105
10. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการเรียนรู้การทดลองที่ 5....	106
11. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนทดสอบการทดลองที่ 5.....	108
4.2 การทดลองจำแนกภาวะหยุดหายใจขณะนอนหลับแบบอัตโนมัติ.....	111
4.2.1 การทดลองและผลการทดลองประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับ.....	111
4.2.1.1 การทดลองที่ 6 การทดลองเปรียบเทียบประสิทธิภาพของการลดมิติแบบผสมผสานของ DisBGWO-DCCA ด้วยเงื่อนไขการแบ่งข้อมูลที่แตกต่างกันสำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG.....	111
4.2.1.2 ผลการทดลองและวิเคราะห์ผลการทดลองของการทดลองที่ 6.....	112
4.2.1.3 การทดลองที่ 7 การทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว.....	115
4.2.1.4 ผลการทดลองและวิเคราะห์ผลการทดลองของการทดลองที่ 7.....	116
บทที่ 5 สรุปผลการทดลองและข้อเสนอแนะ.....	113
5.1 สรุปผลการทดลอง.....	113
5.2 ข้อเสนอแนะ.....	116
เอกสารอ้างอิง.....	117
ภาคผนวก.....	125
ประวัติผู้วิจัย.....	130

สารบัญตาราง

ตารางที่	หน้า
2.1 ระบบการทำงานของสรีรวิทยาของร่างกายที่มีการเปลี่ยนแปลงของระยะ NREM และ REM.....	9
2.2 ช่วงความถี่และขนาดของแต่ละรูปแบบของคลื่นสมอง.....	12
2.3 การนำเสนอการประเมินสัญญาณ EDR ด้วยเทคนิคต่างๆ.....	27
2.4 สรุปรงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์การนอนแบบอัตโนมัติด้วยสัญญาณหลายชุด.....	34
2.5 วิธีหลอมรวมคุณลักษณะด้วยวิธี PCA, LDA, CCA และ DCCA.....	41
2.6 จุดเด่นและจุดด้อยของวิธีลดมิติด้วยวิธีเลือกคุณลักษณะและวิธีการแปลงคุณลักษณะ.....	43
3.1 จำนวนคุณลักษณะของสัญญาณ EEG, EOG, และ ECG.....	57
4.1 จำนวนตัวอย่างของระยะ Wake, ระยะ NREM1-3 และระยะ REM จากฐานข้อมูล SHHS.....	74
4.2 พารามิเตอร์ที่สำคัญสำหรับอัลกอริทึมการจำแนกข้อมูลต่างๆ.....	75
4.3 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนก ระยะการนอนด้วยอัลกอริทึม kNN.....	76
4.4 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนก ระยะการนอนด้วยอัลกอริทึม DT.....	77
4.5 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนก ระยะการนอนด้วยอัลกอริทึม SVM.....	78
4.6 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนก ระยะการนอนด้วยอัลกอริทึม RF.....	79
4.7 ประสิทธิภาพการจำแนกระยะการนอนด้วยวิธีการแปลงคุณลักษณะ PCA, LDA, CCA, DCCA และ อัลกอริทึมการจำแนกข้อมูล kNN, DT, SVM, และ RF.....	81
4.8 พารามิเตอร์สำหรับวิธีการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO.....	87
4.9 ผลการทดลองของขั้นตอนเรียนรู้ของการลดมิติด้วยวิธีการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO ในขั้นตอนเรียนรู้.....	89
4.10 ผลการทดลองของขั้นตอนทดสอบการจำแนกระยะการนอนด้วยวิธีลดมิติด้วยการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO.....	90
4.11 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองที่ 4 วิธีมิติแบบผสมผสาน CCA-NIAs และ DCCA- NIAs.....	94

สารบัญตาราง(ต่อ)

ตารางที่	หน้า
4.12 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลตมิติแบบผสมผสานด้วยวิธี CCA-NIAs และ DCCA-NIAs.....	95
4.13 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน NIAs-CCA/DCCA.....	100
4.14 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลตมิติแบบผสมผสานด้วยวิธี NIAs-CCA และ NIAs-DCCA.....	101
4.15 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลตมิติแบบผสมผสาน NIAs- CCA/DCCA ด้วยการแบ่งข้อมูลตามวิธี Based on subject.....	104
4.16 จำนวนตัวอย่างของระยะ Wake, ระยะ REMและระยะ NREM1-3 ที่มาจากฐานข้อมูล Dream Database.....	106
4.17 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับ การจำแนกกระยะการนอน.....	107
4.18 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลตมิติแบบผสมผสาน DisBGWO-DCCA.....	108
4.19 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับ การจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG.....	114
4.20 ผลการทดลองของขั้นตอนทดสอบการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG ด้วยวิธีลตมิติแบบผสมผสาน DisBGWO-DCCA.....	114
4.21 ผลการทดลองขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการ จำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว.....	117
4.2 ผลการทดลองของขั้นตอนทดสอบการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียวด้วยวิธีลตมิติ DisBGWO-DCCA, PCA, LDA, DCCA และ DisBGWO.....	117

สารบัญรูป

รูปที่	หน้า
1.1 แผนภาพของกระบวนการดำเนินงานวิจัย.....	5
2.1 วัฏจักรของการนอนตลอดช่วงเวลาการนอนทั้งหมด.....	7
2.2 สัญญาณ EEG ของระยะการนอนต่างๆ ได้แก่ ระยะ Wake (a.), ระยะ NREM1 (b.), ระยะ NREM2 (c.), ระยะ NREM3 (d.), ระยะ NREM4 (e.) และ ระยะ REM (f.) ที่มา: [17].....	7
2.3 สัญญาณ Polysomnography ช่วงระยะเวลา 5 นาที.....	11
2.4 ตำแหน่งการติดอิเล็กโทรดของสัญญาณ EEG มาตรฐาน 10/20 placement system.....	12
2.5 จุดสูงสุดและต่ำสุดของสัญญาณ EEG ในแต่ละ window.....	15
2.6 (a.) ความสัมพันธ์ระหว่างค่าของ $avgMMD_{slope}$ ในแต่ละระยะการนอน, (b.) ความสัมพันธ์ระหว่างค่าของ $sumMMD_{slope}$ ในแต่ละระยะการนอน	16
2.7 (a.) ส่วนประกอบของลูกตาและตำแหน่งของอิเล็กโทรดของการวัดสัญญาณ EOG (b.) ความสัมพันธ์ของการเคลื่อนไหวของลูกตาและสัญญาณ EOG ที่ตำแหน่งอิเล็กโทรด LOC.....	17
2.8 ค่าสัญญาณ EOG ที่มากกว่าและต่ำกว่าระดับอ้างอิง.....	18
2.9 (a.) ตำแหน่งของอิเล็กโทรดตามมาตรฐาน 3 lead electrode placement (b.) สัญญาณ ECG....	20
2.10 แผนภาพของขั้นตอนการตรวจจับคลื่นอาร์.....	21
2.11. (a.) สัญญาณ ECG (b.) RR interval time series.....	22
2.12 ย่านความถี่ของสเปกตรัมสำหรับการวิเคราะห์ HRV.....	23
2.13 (a.) กราฟความสัมพันธ์ระหว่าง $logC(r)$ และ $log(r)$ ของสัญญาณ RR ช่วงการการเกิดภาวะหยุดหายใจขณะนอนหลับ (b.) กราฟ boxplot การเปรียบเทียบค่า CD ของภาวะหยุดหายใจขณะนอนหลับและช่วงหายใจปกติ.....	25
2.14 (a.) สัญญาณ RR time series ที่ถูกนอร์มอไรซ์ (b.) กราฟความสัมพันธ์ระหว่าง logarithm $F(n)$ และ logarithm ของขนาด n ในแต่ละโซน.....	26
2.15 (a.) สัญญาณการหายใจจริง (b.) สัญญาณ EDR จากวิธี EMD (c.) สัญญาณ EDR จากวิธี DWT....	29
2.16 แผนภาพของระบบการเรียนรู้ของเครื่องโดยมีผู้สอน.....	35
2.17 แผนภาพของกระบวนการการลดมิติข้อมูล.....	36
2.18 กระบวนการเลือกตัวแปรด้วย (a) วิธีฟิลเตอร์ (b) วิธีแรปเปอร์ และ (c) วิธีฝังตัว.....	44

สารบัญรูป(ต่อ)

รูปที่	หน้า
2.19 วิธีการลดมิติแบบผสมผสานทั้ง 4 แบบ.....	43
2.20 ตัวอย่างการจำแนกกลุ่มแบบ K nearest neighbor.....	50
2.21 ตารางข้อมูลที่ใช้ในการตัดสินใจการเล่นหรือไม่เล่นกอล์ฟ.....	51
2.22 (a.) ค่า information Gain ของแต่ละคุณลักษณะ (b.) การสร้างต้นไม้ตัดสินใจของโหนดราก.....	52
2.23 ต้นไม้ตัดสินใจความเป็นไปได้ของการเล่นและไม่เล่นกอล์ฟด้วยวิธี ID3.....	53
2.21 กระบวนการจำแนกข้อมูลด้วยอัลกอริทึม RF.....	54
2.22 (a) ระบายเส้นตรงที่เป็นไปได้ทั้งหมดของการแบ่งกลุ่มข้อมูล (b) ระบายเส้นตรงที่เหมาะสมของการแบ่งกลุ่มข้อมูล.....	55
3.1 ขั้นตอนประมวลผลก่อน.....	56
3.2 อัลกอริทึม BPSO สำหรับการเลือกคุณลักษณะ.....	61
3.3 อัลกอริทึม BCS สำหรับการเลือกคุณลักษณะ.....	62
3.4 อัลกอริทึม EBCS สำหรับการเลือกคุณลักษณะ.....	63
3.5 อัลกอริทึม BGWO สำหรับการเลือกคุณลักษณะ.....	64
3.6 อัลกอริทึม DisBGWO สำหรับการเลือกคุณลักษณะ.....	65
3.7 กระบวนการลดมิติแบบผสมผสานของที่นำเสนอ CCA/DCCA-NIAs.....	67
3.8 อัลกอริทึม CCA/DCCA-NIAs.....	68
3.9 กระบวนการลดมิติแบบผสมผสานของที่นำเสนอ NIAs-CCA/DCCA.....	69
3.10 อัลกอริทึม NIAs-CCA/DCCA สำหรับการเลือกคุณลักษณะ.....	70
3.11 ขั้นตอนการประมวลผล.....	71
4.1 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะและค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม kNN.....	77
4.2 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะและค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม DT.....	78
4.3 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะและค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม SVM.....	79
4.4 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะและค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม RF.....	80

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.5 ค่าเฉลี่ยความถูกต้องแม่นยำของจำแนกกระยะการนอนที่ได้จากการลดมิติด้วยวิธีการแปลงคุณลักษณะต่าง ๆ.....	82
4.6 เวลาเฉลี่ยในการประมวลผลของจำแนกกระยะการนอนที่ได้จากการลดมิติด้วยวิธีการแปลงคุณลักษณะต่าง ๆ.....	83
4.7 ค่าฟังก์ชันความเหมาะสมของแต่ละการลดมิติข้อมูลด้วยวิธีการแปลงข้อมูลต่างๆและการจำแนกข้อมูลด้วยอัลกอริทึม KNN, DT, SVM และ RF.....	84
4.8 ค่าเฉลี่ยค่าฟังก์ชันความเหมาะสมของการกำหนดค่าเริ่มต้นของแต่ละวิธีค้นหาคำตอบด้วยการสุ่มเลือกแบบต่างๆ ได้แก่ Large, Random และ Small initializations.....	86
4.9 จำนวนรอบการคำนวณการรู้เข้าคำตอบที่เหมาะสมของการกำหนดค่าเริ่มต้นของวิธีค้นหาคำตอบด้วยการสุ่มเลือกแบบต่างๆ ได้แก่ Large, Random และ Small initializations.....	86
4.10 ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธี BPSO, BCS, EBCS, BGWO และ DisBGWO.....	88
4.11 ความสัมพันธ์ระหว่างค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธี BPSO, BCS, EBCS, BGWO และ DisBGWO.....	89
4.12 กราฟ ROC ของวิธีการเลือกคุณลักษณะต่างๆ (BPSO, BCS, EBCS, BGWO และ DisBGWO) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.).....	91
4.13 ความสัมพันธ์ระหว่างค่าเฉลี่ยของความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน CCA-NIAs และ DCCA-NIAs.....	93
4.14 ความสัมพันธ์ระหว่างค่าเฉลี่ยของค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน CCA-NIAs และ DCCA-NIAs.....	94
4.15 กราฟ ROC ของวิธีการลดมิติแบบผสมผสาน CCA-NIAs (CCA-BPSO, CCA-BCS, CCA-EBCS, CCA-BGWO และ CCA-DisBGWO) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.)	96
4.16 กราฟ ROC ของวิธีการลดมิติแบบผสมผสาน DCCA-NIAs (DCCA-BPSO, DCCA-BCS, DCCA-EBCS, DCCA-BGWO และ DCCA-DisBGWO) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.).....	97

สารบัญรูป(ต่อ)

รูปที่	หน้า
4.17 ความสัมพันธ์ระหว่างค่าเฉลี่ยของความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธี การลดมิติแบบผสมผสาน NIAs-CCA/DCCA.....	99
4.18 ความสัมพันธ์ระหว่างค่าเฉลี่ยของค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธีการลดมิติแบบผสมผสาน NIAs-CCA/DCCA.....	99
4.19 กราฟ ROC ของวิธีการลดมิติแบบผสมผสาน NIAs-CCA (BPSO-CCA, BCS-CCA, EBCS-CCA, BGWO-CCA และ DisBGWO-CCA) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.).....	102
4.20 กราฟ ROC ของวิธีการลดมิติแบบผสมผสาน NIAs-DCCA (BPSO-DCCA, BCS-DCCA, EBCS-DCCA, BGWO-DCCA และ DisBGWO-DCCA) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.).....	103
4.21 กราฟเปรียบเทียบประสิทธิภาพของการลดมิติ DisBGWO-DCCA ของเงื่อนไขการแบ่งข้อมูลสำหรับการออกแบบอัลกอริทึมด้วยการแบ่งข้อมูล Based on event และ Based on subject.....	105
4.22 (a.) ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน DisBGWO-DCCA (b.) ความสัมพันธ์ระหว่างค่าเฉลี่ยค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน DisBGWO-DCCA.....	107
4.23 กราฟ Boxplot ค่าประสิทธิภาพของระยะ wake ของกลุ่มคนปกติ (a.) และกลุ่มคนป่วย (b.)....	109
4.24 กราฟ Boxplot ค่าประสิทธิภาพของระยะ REM ของกลุ่มคนปกติ (a.) และกลุ่มคนป่วย (b.)....	109
4.25 กราฟ Boxplot ค่าประสิทธิภาพของระยะ NREM1 ของกลุ่มคนปกติ(a.)และกลุ่มคนป่วย(b.)....	109
4.26 กราฟ Boxplot ค่าประสิทธิภาพของระยะ NREM2 ของกลุ่มคนปกติ(a.)และกลุ่มคนป่วย(b.)....	110
4.27 กราฟ Boxplot ค่าประสิทธิภาพของระยะ NREM3 ของกลุ่มคนปกติ (a.)และกลุ่มคนป่วย(b.)...110	110
4.28 ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนของการคำนวณของวิธี DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG.....	113
4.29 ความสัมพันธ์ระหว่างค่าเฉลี่ยค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธี DisBGWO-DCCA การจำแนกภาวะหยุดหายใจขณะนอนหลับด้วยสัญญาณ EEG และECG.....	116
4.30 ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีลดมิติ DisBGWO และ DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว.....	116

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของงานวิจัย

งานวิจัยด้านการวิเคราะห์การนอนแบบอัตโนมัติเป็นหัวข้อที่ได้รับความสนใจ เนื่องจากสามารถช่วยประเมินและช่วยลดภาระการวิเคราะห์การนอนแบบเดิมได้ โดยเฉพาะการระบุระยะการนอนหลับ (Sleep stages) และการประเมินความผิดปกติขณะนอนหลับ ได้แก่ ภาวะหยุดหายใจขณะนอนหลับ (Sleep apnea disorder) ปัจจุบันการวิเคราะห์การนอนแบบอัตโนมัตินี้ยังอยู่ในระดับของงานวิจัยและยังมีความต้องการอย่างมากสำหรับการพัฒนา โดยเฉพาะอย่างยิ่งการพัฒนาด้านการออกแบบอัลกอริทึมเพื่อสามารถนำไปประยุกต์ใช้ให้มีประสิทธิภาพสูงสุด การวิเคราะห์การนอนแบบอัตโนมัติมีความเกี่ยวข้องโดยตรงกับทางการแพทย์และความปลอดภัยของมนุษย์ ดังนั้นการศึกษาเพื่อสร้าง ออกแบบ และพัฒนาอัลกอริทึมสำหรับการวิเคราะห์การนอนแบบอัตโนมัติจำเป็นต้องคำนึงถึง ปัจจัยด้านความถูกต้องแม่นยำของอัลกอริทึม ด้านทรัพยากรที่ใช้ในการคำนวณ และความสะดวกในการนำไปพัฒนาประยุกต์ใช้

จากงานวิจัยที่ผ่านมาที่เกี่ยวข้องกับการวิเคราะห์การนอนแบบอัตโนมัติแบ่งออกเป็นสองกลุ่มคือการวิเคราะห์การนอนด้วยสัญญาณเพียงชุดเดียว และการวิเคราะห์การนอนด้วยหลายชุดสัญญาณ ถึงแม้ว่าการวิเคราะห์ด้วยสัญญาณเพียงชุดเดียวมีความสะดวกต่อการใช้งาน แต่ความถูกต้องแม่นยำของระบบอัตโนมัติยังไม่สูงมาก ดังนั้นจึงแก้ไขด้วยการนำสัญญาณหลายชุดมาใช้ร่วมกันในการวิเคราะห์ซึ่งแต่ละสัญญาณมีความสำคัญและมีความสามารถแสดงให้เห็นถึงการเปลี่ยนแปลงขณะนอนหลับ แต่อย่างไรก็ตามการใช้หลายชุดสัญญาณนั้นนำมาซึ่งปัญหาของจำนวนข้อมูลที่มากขึ้นด้วย เป็นที่ทราบกันดีว่า จำนวนข้อมูลหรือจำนวนมิติที่มีจำนวนมากเป็นสาเหตุสำคัญที่ทำให้เกิดปัญหาในกระบวนการเรียนรู้ของเครื่อง (Machine learning) หรือ ปัญหาของมิติข้อมูล (Curse of dimension) ได้แก่ การเกิดโอเวอร์ฟิตติ้ง การลดลงของความถูกต้องของโมเดล และการสิ้นเปลืองทรัพยากรที่ใช้ในการประมวลผล เป็นต้น วิธีการแก้ไขปัญหามิติข้อมูลสูงคือ การออกแบบกระบวนการลดมิติข้อมูลที่มีประสิทธิภาพและความเหมาะสมต่อการประมวลผล ในปัจจุบันงานวิจัยที่นำเสนอวิธีการลดมิติข้อมูลมีจำนวนมากเราสามารถแบ่งออกเป็น 2 แบบได้แก่ วิธีการลดมิติแบบเดี่ยว และวิธีการลดมิติแบบผสมผสาน ซึ่งทั้งสองถูกออกแบบมาเพื่อจุดประสงค์เดียวกันคือ นำมาใช้ในการแก้ไขปัญหามิติข้อมูลที่มีจำนวนมาก แต่อย่างไรก็ตามแต่ละวิธีมีข้อจำกัดของการใช้งาน ดังนั้นการนำไปใช้หรือออกแบบอัลกอริทึมจำเป็นต้องคำนึงถึงความเหมาะสมของการประยุกต์ใช้ด้วย จากการศึกษาของงานวิจัยที่ผ่านมาที่เกี่ยวข้องกับการวิเคราะห์การนอนแบบอัตโนมัติ

พบว่า ส่วนใหญ่นำวิธีการลดมิติแบบเดียวมาใช้แก้ปัญหาเท่านั้น มีเพียงบางงานวิจัยที่นำเสนอการเพิ่มประสิทธิภาพของการวิเคราะห์ด้วยวิธีลดมิติแบบผสมผสาน แต่ประสิทธิภาพของอัลกอริทึมเหล่านั้นยังไม่เป็นที่น่าพอใจมากนัก ดังนั้นหัวข้องานวิจัยนี้จึงยังมีความต้องการการพัฒนาอัลกอริทึมอย่างต่อเนื่อง เพื่อให้โมเดลมีประสิทธิภาพ จากที่มาและความสำคัญข้างต้นจึงเป็นแรงจูงใจให้ผู้วิจัยสนใจ ศึกษา พัฒนา และสร้างอัลกอริทึมเพื่อเพิ่มประสิทธิภาพการวิเคราะห์การนอนอัตโนมัติจากหลายชุดข้อมูลให้มีประสิทธิภาพและมีความเหมาะสม

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

วิทยานิพนธ์นี้มุ่งเน้นการศึกษาเรื่อง การวิเคราะห์การนอนแบบอัตโนมัติ ประกอบด้วย การจำแนกระยะการนอนหลับ และการตรวจจับภาวะหยุดหายใจขณะนอนหลับ ซึ่งสามารถวิเคราะห์จากสัญญาณไฟฟ้าร่างกายที่ถูกบันทึกขณะนอนหลับ ได้แก่ สัญญาณไฟฟ้าสมอง สัญญาณไฟฟ้าเคลื่อนไหวของลูกตา และสัญญาณไฟฟ้าหัวใจ แต่อย่างไรก็ตามการนำหลายชุดสัญญาณเป็นสาเหตุทำให้เกิดปัญหามิติข้อมูล การแก้ไขปัญหาดังกล่าวนี้สามารถแก้ไขได้ด้วยกระบวนการลดมิติข้อมูล ซึ่งเป็นวิธีที่มีความสามารถในการคัดเลือกข้อมูลที่สำคัญและลดจำนวนข้อมูลได้อย่างเหมาะสม เพราะฉะนั้นการศึกษาและออกแบบอัลกอริทึมสำหรับการลดมิติข้อมูลจึงเป็นส่วนที่สำคัญอย่างยิ่งต่อการเพิ่มประสิทธิภาพของระบบการเรียนรู้แบบอัตโนมัติ ด้วยเหตุนี้ผู้วิจัยเล็งเห็นว่า การพัฒนาการวิเคราะห์การนอนแบบอัตโนมัติที่มีประสิทธิภาพนั้นต้องมีการออกแบบในแต่ละขั้นตอนอย่างเหมาะสม ซึ่งกระบวนการออกแบบอัลกอริทึมของงานวิจัยนี้ได้มุ่งเน้นในแต่ละขั้นตอน ประกอบด้วย ขั้นตอนการประมวลผลของสัญญาณจากร่างกาย ขั้นตอนวิธีการสกัดคุณลักษณะ ขั้นตอนวิธีการลดมิติข้อมูล และขั้นตอนจำแนกกลุ่มข้อมูล โดยเฉพาะอย่างยิ่งได้มุ่งเน้นถึง การสร้างและนำเสนออัลกอริทึมของการลดมิติข้อมูลแบบผสมผสานใหม่ เพื่อให้สามารถนำไปใช้ในการระบุระยะการนอนหลับ และจำแนกความผิดปกติของภาวะหยุดหายใจขณะนอนหลับได้อย่างเหมาะสมและมีประสิทธิภาพ

1.3 สมมุติฐานของการศึกษา

ในงานวิจัยที่ผ่านมาได้นำเสนอแนวคิดและพัฒนา การระบุระยะการนอน และการระบุภาวะหยุดหายใจขณะนอนหลับแบบอัตโนมัติ โดยใช้สัญญาณต่างๆ จากร่างกาย ได้แก่ สัญญาณไฟฟ้าสมอง สัญญาณการเคลื่อนไหวของลูกตา และสัญญาณไฟฟ้าหัวใจ มาใช้ร่วมกันในการวิเคราะห์เพื่อเพิ่มความถูกต้องแม่นยำ ประหยัดเวลาทรัพยากร และมีประสิทธิภาพ ซึ่งการสกัดคุณลักษณะที่มาจากหลายชุดสัญญาณถูกสกัดมาจากเทคนิคการวิเคราะห์เชิงเส้นและเทคนิคการวิเคราะห์ไม่เชิงเส้น ซึ่งล้วนแต่มีความสำคัญและเป็นคุณลักษณะแสดงให้เห็นถึงการเปลี่ยนแปลงของร่างกายขณะนอนหลับ อย่างไรก็ตาม

ตามถ้าคุณลักษณะที่ถูกเลือกมีความเหมาะสมจะสามารถทำให้การระบุระยะการนอนต่างๆ และการจำแนกภาวะหยุดหายใจขณะนอนหลับมีความแม่นยำมากขึ้น นอกจากนี้ถ้าเซตของคุณลักษณะที่เลือกมีจำนวนที่เหมาะสมด้วยจะทำให้ประหยัดเวลาและทรัพยากรในการประมวลผล สำหรับการออกแบบและการเลือกใช้อัลกอริทึมการลดมิติแบบต่างๆ ทั้งการลดมิติแบบเดี่ยวและการลดมิติแบบผสมผสานจะขึ้นอยู่กับความต้องการและความเหมาะสมในการนำไปประยุกต์ใช้งาน

1.4 ทฤษฎีหรือแนวความคิดที่ใช้ในการวิจัย

เนื่องจากการวิเคราะห์การนอนแบบอัตโนมัติโดยใช้หลายชุดสัญญาณ ได้แก่ สัญญาณไฟฟ้าสมอง (Electroencephalography: EEG) สัญญาณการเคลื่อนไหวของตา (Electrooculography: EOG) และสัญญาณหัวใจ (Electrocardiography: ECG) ซึ่งได้รับการยอมรับว่ามีประสิทธิภาพ วิธีการสกัดคุณลักษณะจากสัญญาณร่างกายด้วยวิธีเชิงเส้นและวิธีไม่เชิงเส้น โดยวิธีเชิงเส้นประกอบด้วย การคำนวณทางโดเมนของเวลา โดเมนของความถี่ และโดเมนของเวลา-ความถี่ โดยคุณลักษณะเหล่านี้มีความสำคัญต่อการบ่งบอกความเปลี่ยนแปลงของร่างกายขณะนอนหลับสามารถนำไปใช้ในการระบุระยะการนอน และตรวจจับภาวะหยุดหายใจได้อย่างมีประสิทธิภาพ แต่อย่างไรก็ตามการนำหลายชุดสัญญาณมาใช้ในการวิเคราะห์นั้นเป็นสาเหตุสำคัญของปัญหาของมิติข้อมูลสูง โดยมีผลกระทบต่อระบบการเรียนรู้ด้วยเครื่อง ซึ่งสามารถแก้ปัญหาได้ด้วยการลดจำนวนมิติ หรือเรียกว่าการลดมิติข้อมูล ซึ่งวิธีลดมิติข้อมูลสามารถแบ่งออกเป็น 2 ประเภท ได้แก่

1. วิธีลดมิติข้อมูลแบบเดี่ยว (Single approach) คือ การลดจำนวนมิติด้วยวิธีการเลือกคุณลักษณะหรือวิธีการแปลงคุณลักษณะอย่างใดอย่างหนึ่ง ในปัจจุบันวิธีลดมิติแบบเดี่ยวได้ถูกนำมาใช้อย่างแพร่หลายและถูกนำมาใช้ในการวิเคราะห์การนอนด้วยเช่นกัน ตัวอย่างเช่น งานวิจัยของ [1-3] ได้นำวิธีเลือกคุณลักษณะแบบฟิลเตอร์ (Filter) ที่ถูกออกแบบมาเพื่อแก้ไขเฉพาะปัญหานั้นๆ ซึ่งมีข้อดีคือ เป็นวิธีที่ง่ายและมีการประมวลผลอย่างรวดเร็ว แต่มีข้อเสียคือ คุณลักษณะที่ถูกเลือกได้เลือกจากเซตข้อมูลการสอนเพียงเท่านั้น ไม่ครอบคลุมเซตข้อมูลการทดสอบทำให้ประสิทธิภาพความถูกต้องแม่นยำลดลง ทั้งนี้เนื่องจากเซตคุณลักษณะที่ถูกเลือกไม่ถูกประเมินประสิทธิภาพ ดังนั้นการแก้ไขปัญหานี้สามารถแก้ไขได้ด้วยการเพิ่มฟังก์ชันการตรวจสอบสำหรับประเมินประสิทธิภาพของคุณลักษณะที่เลือกก่อนเข้าสู่กระบวนการเรียนรู้หรือเรียกว่า วิธีการเลือกคุณลักษณะแบบแรปเปอร์ (Wrapper) ประกอบด้วย 2 วิธี ได้แก่ วิธีเลือกคุณลักษณะโดยลำดับ (Sequential selection algorithm) และวิธีเลือกคุณลักษณะด้วยการสุ่มเลือกหรือการค้นหาคำตอบด้วยอัลกอริทึมที่ได้แรงบันดาลใจจากธรรมชาติ (Natural Inspired Algorithms: NIAs) โดยงานวิจัยนี้ได้เน้นการเลือกคุณลักษณะด้วยการสุ่มเลือก NIAs ได้แก่ Binary Particle Swarm Optimization (BPSO) [4], Binary Cuckoo Search (BCS) [5], Extended Binary

Cuckoo Search (EBCS) [6], และ Binary Gray Wolf Optimization (BGWO) [7-8] เนื่องจากเป็นวิธีที่ถูกนำเสนอว่ามีความสามารถในการช่วยเพิ่มประสิทธิภาพการจำแนกกลุ่มได้อย่างถูกต้องแม่นยำ และสามารถเลือกคุณลักษณะได้อย่างเหมาะสม นอกจากนี้วิธีการเลือกคุณลักษณะแบบ NIAs ยังเป็นวิธีที่ใช้เวลาในการประมวลผลน้อยกว่าเมื่อเปรียบเทียบกับวิธีการเลือกคุณลักษณะโดยลำดับอีกด้วย [4,6] แต่อย่างไรก็ตามอีกปัญหาหนึ่งที่สำคัญของวิธีเลือกคุณลักษณะในการประยุกต์ใช้กับข้อมูลหลายชุด คือคุณลักษณะที่ถูกเลือกนั้นจะไม่มีความสัมพันธ์ระหว่างคุณลักษณะกันเอง เนื่องจากขั้นตอนการเลือกคุณลักษณะจะพิจารณาจากความสัมพันธ์ระหว่างคุณลักษณะกับข้อมูลเป้าหมายเพียงเท่านั้น จึงมีผลกระทบต่อประสิทธิภาพการวิเคราะห์ของข้อมูลมิติสูง ดังนั้นการแก้ไขปัญหาคือความสัมพันธ์ระหว่างตัวแปรดังกล่าวนี้สามารถแก้ไขด้วย วิธีลดมิติแปลงคุณลักษณะด้วยวิธี Canonical Correlation Analysis (CCA) [9] และ Discriminant Canonical Correlation Analysis (DCCA) [10-11] ซึ่งเป็นเทคนิคที่พิจารณาจากความสัมพันธ์ระหว่างคุณลักษณะกันเอง

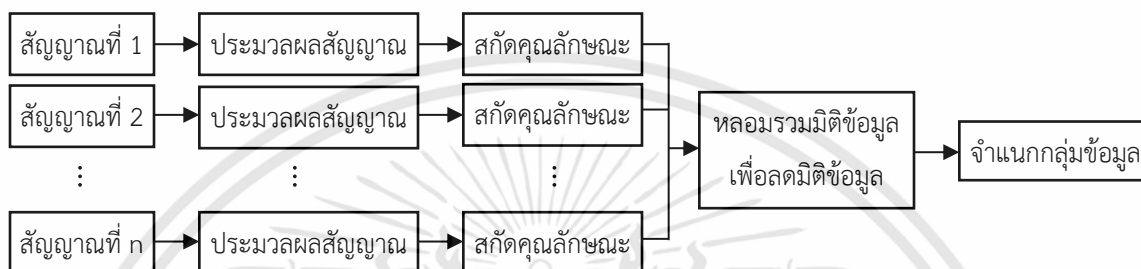
2. วิธีลดมิติข้อมูลแบบผสมผสาน (Hybrid approach) คือ วิธีลดมิติข้อมูลสองวิธีถูกผสมผสานเข้าด้วยกันซึ่งขึ้นอยู่กับการประยุกต์ใช้งาน วิธีลดมิติแบบผสมผสานนอกจากจะนำจุดเด่นของแต่ละวิธีมาใช้แล้วยังสามารถนำจุดเด่นนั้นๆ มาชดเชยแก้ไขของจุดด้อยของอีกวิธีได้อีกด้วย ปัจจุบันการพัฒนาวิธีลดมิติแบบผสมผสานเป็นวิธีใหม่ที่ได้ถูกออกแบบมาเพื่อลดจำนวนมิติและเพิ่มประสิทธิภาพของโมเดลการเรียนรู้ให้มีความถูกต้องแม่นยำมากขึ้น

ดังนั้นแนวคิดข้างต้นถูกนำมาใช้ในงานวิจัยนี้ โดยงานวิจัยนี้นำเสนอวิธีลดมิติแบบผสมผสานด้วยอัลกอริทึมใหม่ที่เกิดจากการผสมผสานกันระหว่าง วิธีลดมิติด้วยวิธีเลือกคุณลักษณะด้วย NIAs และวิธีลดมิติด้วยวิธีแปลงคุณลักษณะของหลายชุดข้อมูลด้วยเทคนิค CCA และ DCCA นอกจากนี้ขั้นตอนการจำแนกกลุ่มของข้อมูลในงานวิจัยนี้ใช้อัลกอริทึมการจำแนกกลุ่มข้อมูล ได้แก่ Support Vector Manchin (SVM), Decision Tree (DT), K-Nearest Neighbor (KNN) และ Random forest (RF) มาใช้ออกแบบและประยุกต์ใช้ในงานวิจัยนี้

1.5 ขอบเขตการวิจัย

งานวิจัยนี้เป็นการศึกษา สร้างอัลกอริทึมสำหรับการวิเคราะห์การนอนแบบอัตโนมัติประกอบด้วย การระบุระยะการนอนหลับแบบอัตโนมัติ และการระบุภาวะหยุดหายใจขณะนอนหลับแบบอัตโนมัติจากหลายชุดสัญญาณด้วยโปรแกรม MATLAB ซึ่งกระบวนการประมวลผลประกอบด้วย 4 ส่วนหลัก ดังนี้ 1) ขั้นตอนการประมวลผลของสัญญาณ 2) ขั้นตอนการสกัดคุณลักษณะจากสัญญาณต่างๆ ของร่างกายที่ถูกบันทึกขณะนอนหลับ ได้แก่ สัญญาณ EEG สัญญาณ EOG สัญญาณ ECG 3) ขั้นตอนการลดมิติข้อมูลด้วยวิธีลดมิติข้อมูลแบบเดี่ยวและวิธีการลดมิติแบบผสมผสาน 4) ขั้นตอนการจำแนกกลุ่มข้อมูล โดยข้อมูลของ

สัญญาณไฟฟ้าจากร่างกายที่ถูกบันทึกขณะนอนหลับหรือ Polysomnography ซึ่งถูกบันทึกจากผู้เชี่ยวชาญและถูกบันทึกด้วยเครื่องมือวัดที่ได้มาตรฐาน จากสถาบัน The National Sleep Research Resource [12] ฐานข้อมูล Dream database [13] และจากฐานข้อมูล Physionet [14-15] ส่วนการสร้างอัลกอริทึมได้ใช้ฟังก์ชันบางส่วนจากไลบรารีของโปรแกรม MATLAB แผนภาพของกระบวนการดำเนินงานวิจัยนี้แสดงได้ดังรูปที่ 1.1



รูปที่ 1.1 แผนภาพของกระบวนการดำเนินงานวิจัย

1.6 ขั้นตอนของการศึกษา

ขั้นตอนของงานวิจัยนี้ เริ่มต้นด้วยการศึกษางานวิจัยที่ผ่านมาที่เกี่ยวข้อง หาข้อดีและข้อด้อยของแต่ละงานวิจัย จากนั้นศึกษาลักษณะทางกายภาพของสัญญาณไฟฟ้าจากร่างกายที่เกี่ยวกับการนอนหลับ ได้แก่ สัญญาณไฟฟ้าสมอง สัญญาณไฟฟ้าการเคลื่อนไหวของตา และสัญญาณไฟฟ้าหัวใจ ซึ่งมีผลต่อการเปลี่ยนแปลงขณะนอนหลับสำหรับนำมาใช้ในการวิเคราะห์การนอน ขั้นตอนที่สองศึกษาและสกัดคุณลักษณะของสัญญาณไฟฟ้าเหล่านั้น และเราได้นำเสนอคุณลักษณะใหม่ของสัญญาณไฟฟ้าสมอง เรียกว่า $avgMMD_{slope}$ และ $sumMMD_{slope}$ [16] ซึ่งเป็นหนึ่งในคุณลักษณะที่แสดงให้เห็นถึงการเปลี่ยนแปลงของระยะเวลาการนอนได้อย่างชัดเจน ขั้นตอนสุดท้ายเราได้ศึกษาวิธีลดมิติแบบเดี่ยวและแบบผสมผสานแบบที่นิยมใช้กันอย่างทั่วไป และนำมาใช้ในการจำแนกระยะเวลาการนอนและจำแนกภาวะหยุดหายใจขณะนอนหลับ ตลอดจนได้นำเสนออัลกอริทึมการลดมิติแบบผสมผสานใหม่ ได้แก่ NIAs-CCA และ NIAs-DCCA [17] มาใช้ประยุกต์และนำมาเปรียบเทียบกับวิธีต่างๆ ขณะเดียวกันได้นำเสนออัลกอริทึมการเลือกคุณลักษณะใหม่เรียกว่า Dissimilarity Binary Gray Wolf Optimization (DisBGWO) รายละเอียดของทฤษฎี หลักการ และงานวิจัยที่เกี่ยวข้องกับวิธีการวิเคราะห์การนอนแบบอัตโนมัติ การลดมิติข้อมูลด้วยวิธีต่างๆ และการจำแนกกลุ่มข้อมูล แสดงในบทที่ 2 และบทที่ 3 กล่าวถึงการออกแบบขั้นตอนดำเนินการ โดยได้อธิบายรายละเอียด ขั้นตอนของการลดมิติด้วยวิธีต่างๆ และขั้นตอนการลดมิติของอัลกอริทึมที่นำเสนอ สำหรับการทดลอง ผลการทดลอง และวิจารณ์ผลการทดลองถูกนำเสนอในบทที่ 4 และสุดท้ายผลการวิจัยและข้อเสนอแนะนำเสนอไว้ในบทที่ 5

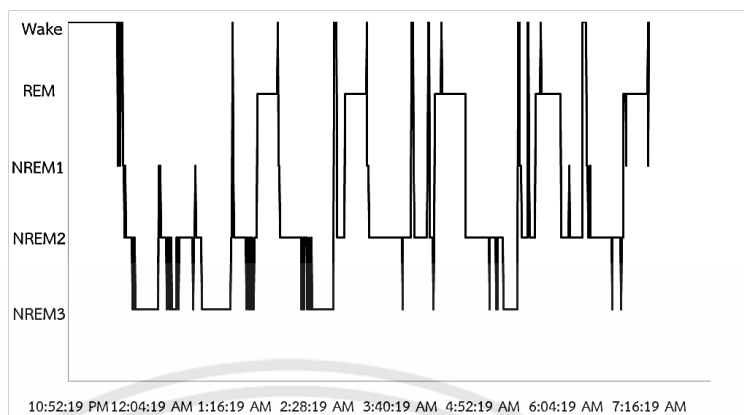
บทที่ 2

หลักการ ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึง หลักการ ทฤษฎี และงานวิจัยที่เกี่ยวข้องกับการออกแบบระบบการวิเคราะห์การนอนแบบอัตโนมัติ โดยเริ่มจากทฤษฎีทางกายภาพของระยะการนอนต่างๆ ภาวะหยุดหายใจขณะนอนหลับ สัญญาณที่ได้จากการบันทึกขณะนอนหลับ และวิธีการสกัดคุณลักษณะของสัญญาณเพื่อนำใช้ในการวิเคราะห์การนอน รวมถึงงานวิจัยที่ผ่านมาที่เกี่ยวข้องกับการออกแบบระบบการวิเคราะห์การนอนแบบอัตโนมัติ ต่อมาได้กล่าวถึงหลักการและวิธีการออกแบบระบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง การลดมิติข้อมูลซึ่งประกอบด้วย การลดมิติด้วยวิธีการเลือกคุณลักษณะ การลดมิติด้วยวิธีการแปลงคุณลักษณะ และการลดมิติข้อมูลแบบผสมผสาน นอกจากนี้ยังได้อธิบายเกี่ยวกับอัลกอริทึมการค้นหาคำตอบที่ได้แรงบรรดาใจมาจากธรรมชาติ ส่วนสุดท้ายได้กล่าวถึงอัลกอริทึมที่ใช้ในการจำแนกกลุ่มข้อมูล

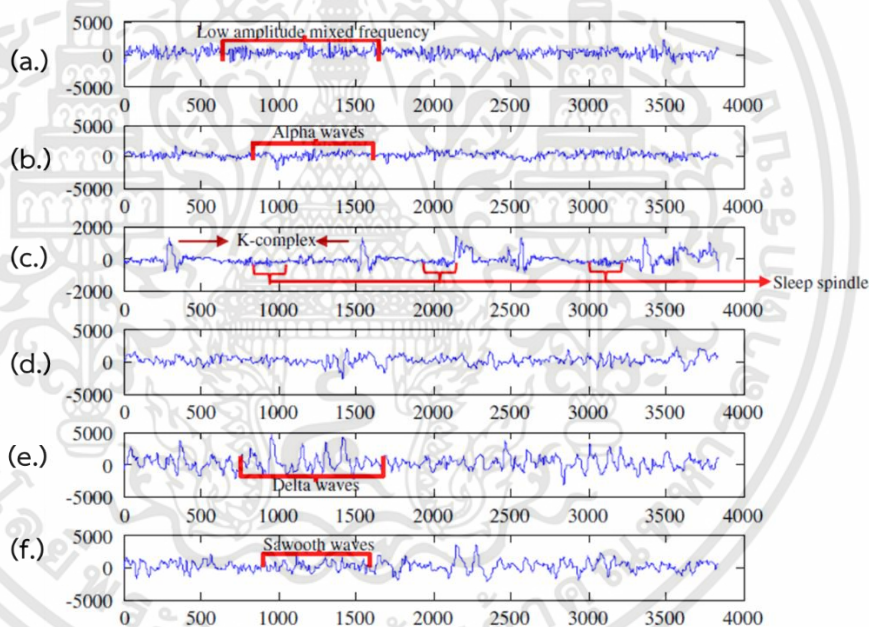
2.1 ระยะการนอนหลับ (Sleep stage)

โครงสร้างพื้นฐานของการนอนปกติประกอบด้วย 2 ระยะหลักของการนอนหลับ ได้แก่ ระยะ Non-rapid eye movement (NREM) และระยะ Rapid eye movement (REM) สำหรับระยะ NREM ถูกแบ่งเป็น 4 ระยะการนอน (Sleep Stage) ได้แก่ NREM1, NREM2, NREM3, และ NREM4 แต่ละระยะการนอนนั้นมีรูปแบบของสัญญาณไฟฟ้าร่างกาย และรูปแบบของการทำงานของสรีระวิทยาร่างกายที่แตกต่างกัน ในขณะที่นอนหลับนั้นมีระยะ NREM และระยะ REM เกิดขึ้นสลับกัน เรียกว่า วัฏจักรของการนอน (Sleep Cycle) ดังแสดงในรูปที่ 2.1 หนึ่งวัฏจักรของการนอน เริ่มต้นด้วยระยะ NREM1 และดำเนินไปสู่ระยะ NREM2 ตามด้วยระยะ NREM3 และ NREM 4 ตามลำดับ และเข้าสู่ระยะสุดท้ายคือระยะ REM ซึ่งตลอดช่วงเวลาการนอนทั้งหมด การเกิดวัฏจักรดังกล่าวนี้จะเกิดขึ้นซ้ำๆ กันประมาณ 7-8 รอบ โดยปกติวัฏจักรของการนอนของผู้ใหญ่พบว่า ในช่วงกลางของเวลาการนอนทั้งหมดจะมีระยะ REM เกิดขึ้นยาวนานกว่าปกติ และระยะ NREM2 เกิดขึ้นเป็นส่วนใหญ่ สำหรับช่วงระยะ NREM3 บางครั้งจะหายไป



รูปที่ 2.1 วัฏจักรของการนอนตลอดช่วงเวลาการนอนทั้งหมด

ระยะการนอนทั้งหมดมีความเกี่ยวข้องกับการทำงานของสมองและระบบสรีรวิทยาทางร่างกายอื่นๆ ความสัมพันธ์ของสัญญาณ EEG ของแต่ละระยะการนอนแสดงดังรูปที่ 2.2



รูปที่ 2.2 สัญญาณ EEG ของระยะการนอนต่างๆ ได้แก่ ระยะ Wake (a), ระยะ NREM1 (b), ระยะ NREM2 (c), ระยะ NREM3 (d), ระยะ NREM4 (e) และ ระยะ REM (f) ที่มา: [17]

ระยะการนอนต่างๆมีความสัมพันธ์กับสัญญาณ EEG มีรายละเอียดดังนี้

1. NREM 1 เป็นระยะที่มีบทบาทในการเปลี่ยนเข้าสู่ช่วงเวลานอนหลับ ในคนปกติระยะนี้เกิดช่วงเริ่มต้นการนอนด้วยระยะเวลาประมาณ 1-7 นาทีหรือประมาณร้อยละ 2-5 ของการนอนหลับทั้งหมด

ระยะ NREM1 เป็นช่วงที่สามารถถูกปลุกให้ตื่นได้อย่างง่ายดาย ซึ่งระยะนี้มีคลื่นสมองที่มีขนาดเล็ก มีช่วงความถี่ต่ำประมาณ 8-13 เฮิร์ตซ์ หรือช่วงคลื่นที่เรียกว่า Alpha wave

2. NREM 2 เป็นระยะที่มีลักษณะของคลื่นสมองที่มีลักษณะเฉพาะ คือ คลื่นไฟฟ้าสมองมีลักษณะเป็นสัญญาณไซน์ (Sinusoidal wave) ซึ่งมีขนาดของแอมพลิจูดสูงตรงกลางคลื่นมากกว่าส่วนเริ่มต้นและส่วนปลาย หรือเรียกสัญญาณนี้ว่า sleep spindles และเป็นระยะที่พบคลื่นไฟฟ้าสมองที่มีขนาดใหญ่ เรียกว่า K-complexes

3. NREM3 และ NREM 4 เป็นระยะที่หลับลึกหรือถูกเรียกว่าเป็นช่วง Slow-wave sleep (SWS) มักจะเกิดขึ้นประมาณหนึ่งในสามของการนอนทั้งหมด เป็นระยะที่สามารถปลุกได้ยากและมีคลื่นไฟฟ้าสมองขนาดใหญ่ที่ปรากฏอย่างเห็นได้ชัด และมีช่วงความถี่ต่ำประมาณ 0.5-4 เฮิร์ตซ์ หรือเรียกว่าคลื่นเดลต้า (delta wave) ส่วนใหญ่ช่วงระยะ NREM 4 จะพบมากกว่าช่วงระยะ NREM 3 จึงทำให้การกำหนดระยะการนอนตามมาตรฐาน AASM ได้รวม NREM 3 และ NREM 4 เป็นระยะเดียวกันและเรียกว่า ช่วงระยะ NREM 3 ของการนอนหลับ

4. ช่วง Rapid Eye Movement (REM) เป็นระยะที่มีคลื่นไฟฟ้าสมองที่มีขนาดเล็กและมีความถี่สูงในช่วงย่านความถี่ของคลื่นเบต้า (beta wave) ที่มีความคล้ายกับคลื่นสมองในขณะที่ตื่น โดยระยะนี้ถูกพบว่า จังหวะของการกรอกตาเป็นจังหวะรวดเร็ว และมีการทำงานของระบบประสาทอัตโนมัติส่วนซิมพาเทติกที่สูงขึ้น

นอกจากนี้ตารางที่ 2.1 เราได้สรุปการทำงานของระบบสรีรวิทยาอื่นๆ ของร่างกายในคนปกติที่มีการเปลี่ยนแปลงขณะนอนหลับช่วงระยะ NREM และ REM จากงานวิจัยของ [18-20] มีรายละเอียดดังนี้

1. ระบบหัวใจและความดันโลหิต การเปลี่ยนแปลงของความดันโลหิตและอัตราการเต้นของหัวใจเกิดขึ้นขณะนอนหลับและถูกกำหนดด้วยระบบประสาทอัตโนมัติ เช่น การเพิ่มขึ้นของความดันโลหิตและอัตราการเต้นของหัวใจจะเกิดขึ้นเพียงช่วงเวลาเล็กๆ ที่เกิดขึ้นพร้อมกับช่วงที่มีคลื่นสมอง K-complexes และเป็นช่วงที่มีการเคลื่อนไหว นอกจากนี้ช่วงเช้าหรือช่วงเวลาก่อนตื่นนอนอัตราการเต้นของหัวใจและความดันโลหิตจะมีการเปลี่ยนแปลงที่เพิ่มขึ้นอย่างรวดเร็ว เนื่องจากร่างกายถูกกระตุ้นและมีความเสี่ยงต่อการเสียชีวิตด้วยหัวใจวายฉับพลันอีกด้วย

2. ระบบประสาทซิมพาเทติก มีการทำงานที่ลดลงในระยะ NREM และจะลดลงเมื่อร่างกายหลับลึกมากขึ้น แต่อย่างไรก็ตามของระยะ NREM อาจจะมีการทำงานที่เพิ่มขึ้น เนื่องจากอัตราการเต้นของหัวใจที่เปลี่ยนแปลงตามคลื่นสมอง K-complexes สำหรับระยะ REM ระบบประสาทซิมพาเทติกจะมีการทำงานที่เพิ่มขึ้นเมื่อเทียบกับระยะ NREM

3. ระบบการหายใจ การไหลเวียนของอากาศในระบบหายใจมีการเปลี่ยนแปลงและมีการทำงานที่เพิ่มขึ้นอย่างรวดเร็วโดยเฉพาะอย่างยิ่งระยะ REM ในระหว่างการนอนสำหรับระยะ NREM ระบบการหายใจจะทำงานลดลงและช้ากว่า และระบบไหลเวียนอากาศทำงานช้าที่สุดช่วง NREM3 และ NREM4

4. ระบบการไหลเวียนของโลหิตของสมอง ช่วงขณะนอนหลับระยะ NREM พบว่ามีการทำงานที่มีความสัมพันธ์โดยตรงกับการลดลงของระบบการไหลเวียนโลหิตและระบบเผาผลาญ ในขณะเดียวกันระยะ REM และระยะ Wake มีความสัมพันธ์โดยตรงกับการเพิ่มขึ้นของการทำงานของระบบการไหลเวียนโลหิตและระบบเผาผลาญ

ตารางที่ 2.1 ระบบการทำงานของสรีรวิทยาของร่างกายที่มีการเปลี่ยนแปลงของระยะ NREM และ REM

กระบวนการทำงานของสรีรวิทยาต่างๆ	ระยะ NREM	ระยะ REM
การทำงานของสมอง	ทำงานลดลงจากระยะตื่นตัว	สมองส่วนควบคุมการเคลื่อนไหวของร่างกายและส่วนควบคุมการสัมผัสมีการทำงานเพิ่มขึ้น สำหรับสมองส่วนอื่นๆมีการทำงานเหมือนกับระยะ NREM
อัตราการเต้นของหัวใจ	ช้าลงจากระยะตื่นตัว	เพิ่มขึ้นและมีความแปรปรวนเมื่อเทียบกับระยะ NREM
ระบบความดันโลหิต	ลดลงจากระยะตื่นตัว	เพิ่มขึ้นประมาณร้อยละ 30 และมีความแปรปรวนเมื่อเทียบกับระยะ NREM
การทำงานของระบบประสาทอัตโนมัติซิมพาเทติก	ลดลงจากระยะตื่นตัว	เพิ่มขึ้นจากระยะตื่นตัว
การทำงานของกล้ามเนื้อ	หยุดเคลื่อนไหว	มีการทำงานคล้ายกับระยะตื่นตัว
การไหลเวียนของเลือดเข้าสู่สมอง	ลดลงจากระยะตื่นตัว	เพิ่มขึ้นจากระยะ NREM
ระบบการหายใจ	ลดลงจากระยะตื่นตัว	เพิ่มขึ้นและมีความแปรปรวนเมื่อเทียบกับระยะ NREM
อุณหภูมิของร่างกาย	ต่ำกว่าจากระยะตื่นตัว มีเหงื่อออก	มีอุณหภูมิที่แปรปรวน

2.2 ภาวะหยุดหายใจขณะนอนหลับ (Sleep disorder)

ภาวะหยุดหายใจขณะนอนหลับ หรือ Sleep apnea disorder เป็นความผิดปกติหนึ่งที่เกิดขึ้นขณะนอนหลับซึ่งพบมากในปัจจุบัน ภาวะหยุดหายใจขณะนอนหลับถูกแบ่งออกเป็น 3 ชนิดตามสาเหตุของการเกิดภาวะนี้ ได้แก่ ภาวะหยุดหายใจขณะหลับชนิดอุดกั้น (Obstructive sleep apnea OSA) เป็นสาเหตุหลักที่พบมากที่สุด ภาวะหยุดหายใจขณะหลับเกิดจากการทำงานผิดปกติของสมอง (Central sleep apnea) และภาวะหยุดหายใจขณะหลับชนิดผสม โดยลักษณะอาการของความผิดปกติเหล่านี้ยากต่อการสังเกตหรือวินิจฉัยได้ด้วยตัวเองต้องอาศัยคนรอบข้างสังเกตอาการขณะนอนหลับ อาการของความผิดปกตินี้คือ เมื่อมีบางสิ่งบางอย่าง เช่น กล้ามเนื้อบริเวณลำคอ ลิ้นไก่ที่มีขนาดใหญ่ เป็นต้น อุดกั้นทางเดินหายใจหรือระบบประสาทอัตโนมัติส่วนกลางที่ควบคุมระบบหายใจทำงานผิดปกติในขณะนอนหลับจะทำให้หายใจลำบากและจนกระทั่งหยุดหายใจชั่วคราว ทำให้ไม่มีก๊าซออกซิเจนไปเลี้ยงสมอง เลือดมีสถานะเป็นกรด และสมองจะปลุกร่างกายตื่นเพื่อหายใจรับก๊าซออกซิเจนเข้าไป ซึ่งลักษณะอาการเช่นนี้จะเกิดขึ้นซ้ำๆกันตลอดทั้งคืน ทำให้ร่างกายไม่ได้รับการพักผ่อนที่เพียงพอหากเกิดติดต่อกันในระยะยาวจะส่งผลกระทบต่อสมองและหัวใจ เช่น โรคหัวใจ โรคความดันสูง โรคกล้ามเนื้ออ่อนแรง โรคหัวใจเต้นผิดจังหวะ และภาวะซึมเศร้า เป็นต้น ในปัจจุบันการตรวจวินิจฉัยภาวะความผิดปกติดังกล่าวนี้แพทย์หรือผู้เชี่ยวชาญจะประเมินและวิเคราะห์จากผลของข้อมูลของสัญญาณร่างกายต่างๆ ที่ถูกบันทึกระหว่างการนอนหลับ หรือเรียกว่าการบันทึกสัญญาณการนอนหลับ (Polysomnography recording) โดยทั่วไปขั้นตอนประเมินผลของการนอนใช้การประเมินจากการวิเคราะห์หรือการระบุด้วยสายตาของมนุษย์ ซึ่งอาศัยประสบการณ์ของผู้เชี่ยวชาญและใช้ระยะเวลาในการวิเคราะห์แต่ละครั้ง

เกณฑ์การระบุช่วงการหยุดหายใจถูกกำหนดโดย American Academy of Sleep Medicine [21,22] มีเงื่อนไขดังนี้

การกำหนดช่วง apnea หรือภาวะหยุดหายใจ มีเงื่อนไขทั้งหมดดังนี้

1. จุดสูงสุดของสัญญาณการหายใจที่วัดได้จากเซนเซอร์ในแต่ละครั้งหรือ สัญญาณที่ได้วัดจาก Nasal pressure ลดลงมากกว่าร้อยละ 90 ของ Baseline

2. ช่วงระยะเวลาการหยุดหายใจยาวนานอย่างน้อย 10 วินาที

การกำหนดช่วง hypopnea หรือภาวะหายใจแผ่ว มีเงื่อนไขทั้งหมดดังนี้

1. จุดสูงสุดของสัญญาณการหายใจที่วัดได้จากเซนเซอร์ในแต่ละครั้งหรือ สัญญาณที่ได้วัดจาก Nasal pressure ลดลงมากกว่าร้อยละ 30 ของ baseline

2. ช่วงระยะเวลาการหายใจแผ่วยาวนานอย่างน้อย 10 วินาที

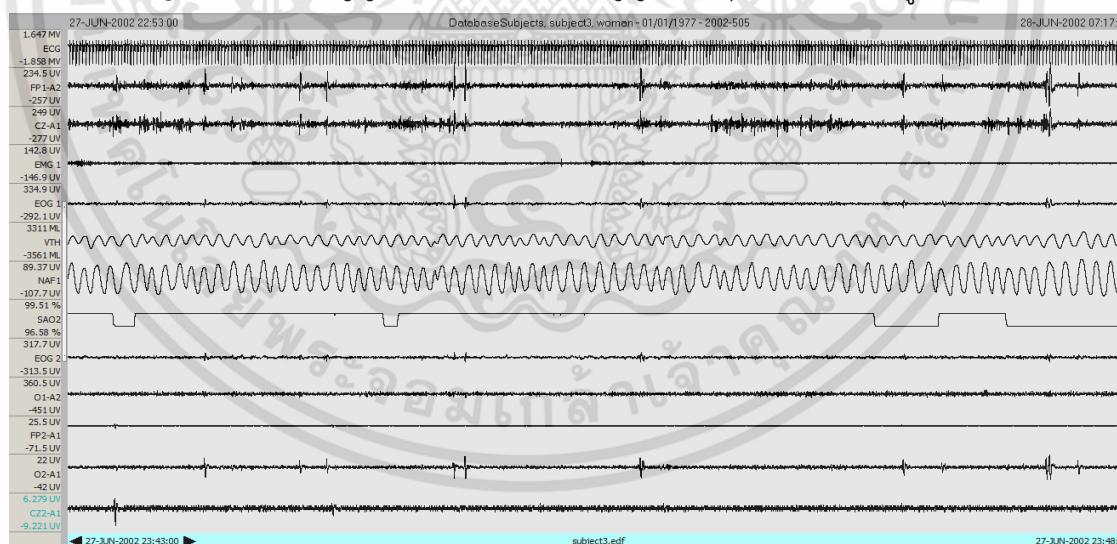
3. ปริมาณความเข้มข้นของออกซิเจนลดลงมากกว่าร้อยละ 4 ของช่วง epoch ก่อนหน้านั้น

จำนวนครั้งของการหยุดหายใจใน 1 ชั่วโมงเรียกว่า Apnea index (AI) สำหรับค่า AI ของการนอนที่เป็นปกติในผู้ใหญ่จะไม่เกิน 5 ครั้งต่อชั่วโมง การถูกรบกวนการนอนอาจเกิดสภาวะหายใจแผ่วร่วมอยู่ด้วย ดังนั้นค่า AHI จึงถือว่าเป็นถือว่าเป็นตัวชี้วัดความผิดปกติดังกล่าวเช่นกัน จำนวนครั้งของการเกิดสภาวะหยุดหายใจและสภาวะหายใจแผ่วในหนึ่งชั่วโมง (AHI หรือ Apnea-hypopnea index) โดยทั่วไปถ้าค่า AHI >5 ในผู้ใหญ่ และมากกว่า 1 ในเด็กถือว่าเป็นความผิดปกติขณะนอนหลับ ความรุนแรงของภาวะหยุดหายใจขณะนอนหลับนี้สามารถแบ่งได้ตามค่า AHI ได้ดังนี้

1. ระดับความรุนแรง Mild : ค่า AHI อยู่ระหว่าง 5-14 ครั้งต่อชั่วโมง
2. ระดับความรุนแรง Moderate : ค่า AHI อยู่ระหว่าง 15-29 ครั้งต่อชั่วโมง
3. ระดับความรุนแรง Severe : ค่า AHI มากกว่า 30 ครั้งต่อชั่วโมง

2.3 สัญญาณไฟฟ้าร่างกายที่เกี่ยวข้องกับการนอน

การประเมินและการวิเคราะห์ขณะนอนหลับถูกประเมินด้วย การวัดสัญญาณไฟฟ้าส่วนต่างๆ ของร่างกายขณะนอนหลับหรือเรียกว่า Polysomnography (PSG) ประกอบด้วย คลื่นไฟฟ้าสมอง (Electroencephalogram: EEG), คลื่นไฟฟ้ากล้ามเนื้อ (Electromyogram: EMG), การเคลื่อนไหวของลูกตา (Electrooculogram: EOG), สัญญาณปริมาณออกซิเจนในเลือด (SaO_2), สัญญาณไฟฟ้าหัวใจ (Electrocardiogram: ECG), สัญญาณการหายใจ และสัญญาณอื่นๆ ที่เกี่ยวข้อง ดังรูปที่ 2.3

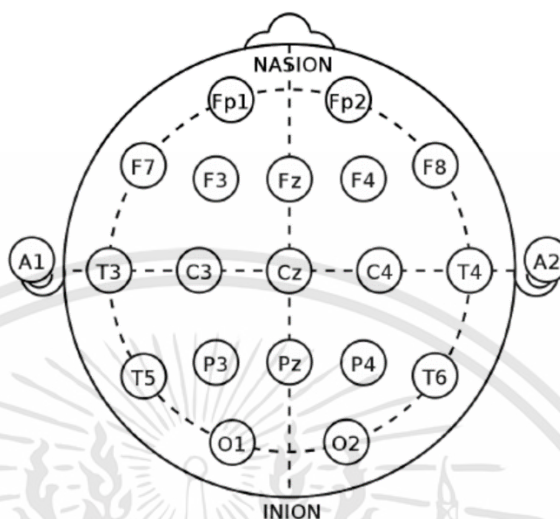


รูปที่ 2.3 สัญญาณ Polysomnography ช่วงระยะเวลา 5 นาที

2.3.1 สัญญาณไฟฟ้าสมองหรือสัญญาณ EEG

สัญญาณ EEG เป็นเครื่องมือแบบ Noninvasive สำหรับการวัดกระแสไฟฟ้าที่เกิดจากการทำงานของสมองมนุษย์ สัญญาณ EEG ถูกวัดด้วยการใช้อิเล็กโทรดจำนวนหลายอิเล็กโทรดจะวัดบริเวณหนึ่ง

ศีรษะในตำแหน่งต่างๆ ตามระบบมาตรฐาน 10/20 (10/20 placement system) ซึ่งสามารถแสดงได้ดังรูปที่ 2.4



รูปที่ 2.4 ตำแหน่งการติดตั้งอิเล็กโทรดของสัญญาณ EEG มาตรฐาน 10/20 placement system ที่มา [22]

สำหรับรูปแบบสัญญาณ EEG สามารถถูกแบ่งเป็น 5 ย่านความถี่ที่แตกต่างกันดังตารางที่ 2.2

ตารางที่ 2.2 ช่วงความถี่และขนาดของแต่ละรูปแบบของคลื่นสมอง

ย่านความถี่	ช่วงความถี่ (เฮิรตซ์)	ขนาดแอมป์ริจูด (ไมโครโวลต์)
เดลต้า (Delta: δ)	0-4	20-100
เทต้า (Theta: θ)	4-8	10
แอลฟา (Alpha: α)	8-13	2-100
เบต้า (Beta: β)	13-22	5-10
แกมมา (Gamma: γ)	>30	-

วิธีการสกัดคุณลักษณะของสัญญาณ EEG เพื่อใช้ในการวิเคราะห์การนอน สามารถแบ่งออกเป็น 4 โดเมน ดังนี้ โดเมนเวลา โดเมนความถี่ โดเมนเวลาความถี่ และการวิเคราะห์แบบไม่เชิงเส้น วิธีการคำนวณทางโดเมนเวลาสามารถทำได้ด้วยวิธีทางสถิติอย่างง่าย ได้แก่ ค่าเฉลี่ย ค่าความแปรปรวน ค่าเบี่ยงเบนมาตรฐาน ค่าสูงสุดต่ำสุดของขนาดแอมป์ริจูด kurtosis skewness สำหรับการวิเคราะห์ทางเวลาอื่นๆ ที่มีความเกี่ยวข้องและถูกนำมาใช้ในการวิเคราะห์ความแตกต่างของระยะการนอนและการจำแนกภาวะหยุดหายใจขณะนอนหลับ มีดังนี้

1. พารามิเตอร์ Hjouth ในงานของ [23] ได้นำเสนอพารามิเตอร์สำหรับใช้ในการจำแนกระยะการนอน พารามิเตอร์นี้ถูกคำนวณมาจากอนุพันธ์เชิงเส้นที่หนึ่งและอนุพันธ์เชิงเส้นที่สอง ประกอบด้วย การวัดค่าเฉลี่ยของพลังงานหรือเรียกว่า activity ดังสมการ 2.1 การประเมินค่าเฉลี่ยของความถี่หรือเรียกว่า Mobility ดังสมการ 2.2 และการประเมินความกว้างของย่านความถี่หรือ Complexity ดังสมการ 2.3

$$Activity = \sigma_0^2 \quad (2.1)$$

$$Mobility = \frac{\sigma_1}{\sigma_0} \quad (2.2)$$

$$Complexity = \sqrt{\left(\frac{\sigma_2}{\sigma_1}\right)^2 - \left(\frac{\sigma_1}{\sigma_0}\right)^2} \quad (2.3)$$

โดย σ_0 คือ ความแปรปรวนของสัญญาณ EEG, σ_1 และ σ_2 ความแปรปรวนของอนุพันธ์ที่หนึ่งและสองตามลำดับ

2. ระยะห่าง Itakura (Itakura distance) งานวิจัยของ [24] นำเสนอพารามิเตอร์ของระยะห่าง Itakura เป็นคุณลักษณะหนึ่งสำหรับการวิเคราะห์การนอนโดยเฉพาะอย่างยิ่งการจำแนกระยะการนอน โดยปกติพารามิเตอร์ระยะห่าง Itakura นิยมใช้กับการประมวลผลของสัญญาณการพูด (Speech Processing) จุดประสงค์ของพารามิเตอร์นี้คือ วัดระยะห่างระหว่างสัญญาณสองชุดโดยประเมินจากค่าสัมประสิทธิ์ (Coefficient) เชิงเส้นของฟังก์ชัน Autoregressive การนำมาประยุกต์ใช้กับสัญญาณ EEG เพื่อวัดความเหมือนระหว่างสัญญาณ EEG แต่ละ epoch และสัญญาณ EEG อ้างอิง (Reference signal) กำหนดให้สัญญาณอ้างอิงเป็น $x[n]$ และสัญญาณ EEG แต่ละ epoch เป็น $y[n]$ เมื่อนำไปผ่านกระบวนการประเมินด้วยฟังก์ชัน Autoregressive จะได้ค่าสัมประสิทธิ์ของสัญญาณ $x[n]$ เป็น $a_x = [1 - a_1 - a_2 \dots - a_p]$ และค่าสัมประสิทธิ์ของสัญญาณ $y[n]$ เป็น $a_y = [1 - a_1 - a_2 \dots - a_p]$ โดยที่ p คือ จำนวนกำลังของโมเดล Autoregressive และให้ R_{xx} เป็นเมตริกซ์ความสัมพันธ์ (Correlation Matrix) ของสัญญาณ $x[n]$ โดยที่ R_{xx} หาได้จากสมการ

$$R_{xx}(m) = \begin{bmatrix} r_{xx}(0) & r_{xx}(1) & \dots & r_{xx}(m) \\ r_{xx}(1) & r_{xx}(0) & \dots & \dots \\ \dots & \dots & \dots & \dots \\ r_{xx}(m) & \dots & \dots & r_{xx}(0) \end{bmatrix} \quad (2.4)$$

โดยที่ r_{xx} เป็นค่า Autocorrelation ของสัญญาณ $x[n]$ ที่มี lag time ที่ m หลังจากนั้นคำนวณค่าความผิดพลาดเฉลี่ยกำลังสอง (Mean square error: MSE) ดังสมการ

$$MSE_{xx} = a_x^T R_{xx}(m) a_x \quad (2.5)$$

ระยะ itakura ถูกอธิบายด้วยการเปรียบเทียบระหว่างสัญญาณอื่นๆกับสัญญาณที่สนใจ ดังนั้น ระยะทาง itakura ของสัญญาณอ้างอิง (d_{xy}) คำนวณจากสมการ

$$d_{xy} = \log \frac{MSE_{xy}}{MSE_{xx}} = \log \frac{a_y^T R_{xx}(m) a_y}{a_x^T R_{xx} a_x} \quad (2.6)$$

และระยะทาง itakura ของสัญญาณ EEG แต่ละ Epoch (d_{yx}) คำนวณจากสมการ

$$d_{yx} = \log \frac{MSE_{yx}}{MSE_{yy}} = \log \frac{a_x^T R_{yy}(m) a_x}{a_y^T R_{yy} a_y} \quad (2.7)$$

ดังนั้นคุณลักษณะ $d(x,y)$ ที่นำไปใช้สำหรับการวิเคราะห์การนอนได้จากการรวมกันของ d_{xy} และ d_{yx} ดังสมการ

$$d(x,y) = \frac{1}{2} \left(\log \left(\frac{a_y^T R_{xx}(m) a_y}{a_x^T R_{xx}(m) a_x} \right) + \log \left(\frac{a_x^T R_{yy}(m) a_x}{a_y^T R_{yy}(m) a_y} \right) \right) \quad (2.8)$$

3. ระยะของจุดสูงสุดและจุดต่ำสุด (Maximum-Minimum distance: MMD)

ระยะของจุดสูงสุดและจุดต่ำสุดเป็นคุณลักษณะทางโดเมนเวลาของสัญญาณ EEG นำเสนอโดย [25] ซึ่งสามารถจำแนกการนอนได้อย่างมีประสิทธิภาพ โดยทั่วไปสัญญาณ EEG เป็นสัญญาณที่ไม่มีรูปแบบที่ชัดเจน (Non-stationary) ดังนั้นสัญญาณ EEG ทางโดเมนเวลาจึงถูกประเมินในวงเวลาสั้นๆ หรือเรียกว่า segment ขั้นตอนการคำนวณคุณลักษณะ MMD เริ่มต้นด้วย สัญญาณ EEG แต่ละ epoch จะถูกแบ่งเป็น window เล็กๆ ที่มีขนาดเท่ากับ ความยาวของรูปคลื่นสัญญาณ EEG หรือความยาวของ window (λ) มีเงื่อนไขดังนี้

$$\lambda = \begin{cases} 100, n < 10,000 \\ 10^{\lfloor \log n \rfloor - 1}, n \geq 10,000 \end{cases} \quad (2.9)$$

ในกรณีของจำนวนตัวอย่างของสัญญาณ (n) ทั้งหมดน้อยกว่า 10,000 ตัวอย่าง λ ถูกกำหนดเท่ากับ 100 แต่ถ้า n มากกว่าหรือเท่ากับ 10,000 ตัวอย่าง λ ถูกกำหนดเท่ากับ $10^{\lfloor \log n \rfloor - 1}$ หลังจากนั้นหาค่า MMD ซึ่งคำนวณได้จากระยะห่างระหว่างจุดสูงสุดและต่ำสุดของสัญญาณ EEG ในแต่ละ window ด้วยหลักการพื้นฐานจากการคำนวณระยะทางพิทาโกรัส (d) ดังสมการ

$$d = \sqrt{\Delta x^2 + \Delta y^2} \quad (2.10)$$

โดยที่ Δx คือ $x_2 - x_1$ และ Δy คือ $y_2 - y_1$ ซึ่งค่า x_1 และ x_2 คือค่าแกน x ณ จุดสูงสุดหรือต่ำสุด และค่า y_1 และ y_2 คือ ค่าแกน y ณ จุดสูงสุดหรือต่ำสุด ดังแสดงดังรูปที่ 2.5 คุณลักษณะ MMD แต่ละ epoch ได้จากผลรวมระยะทางทั้งหมดของ window (w) ดังนี้

$$MMD = \sum_{i=1}^w |d_i| \quad (2.11)$$

4. ค่าเฉลี่ยของอัตราส่วนระหว่างจุดสูงสุดและจุดต่ำสุด และผลรวมของอัตราส่วนระหว่างจุดสูงสุดและจุดต่ำสุด ($avgMMD_{slope}$ และ $sumMMD_{slope}$)

ในงานวิจัย [16] เราได้นำเสนอคุณลักษณะใหม่ที่ได้จากสัญญาณ EEG โดยพัฒนาจากคุณลักษณะ MMD แทนที่จะคำนวณหาค่าผลรวมของระยะทางเพียงอย่างเดียว เราได้นำเสนอคุณลักษณะ ค่าเฉลี่ยของอัตราส่วนระหว่างจุดสูงสุดและต่ำสุดของทุก window แต่ละ epoch ($avgMMD_{slope}$) และผลรวมของอัตราส่วนระหว่างจุดสูงสุดและจุดต่ำสุดของทุก window แต่ละ epoch ($sumMMD_{slope}$) การคำนวณความชันที่ได้จากระยะสูงสุดและต่ำสุดแสดงถึงความเร็วของสัญญาณสูงสุดและต่ำสุด ที่สามารถแสดงให้เห็นถึงความแตกต่างในแต่ละระยะการนอน และมีผลต่อการเพิ่มประสิทธิภาพสำหรับการจำแนกระยะการนอน คุณลักษณะ $avgMMD_{slope}$ และ $sumMMD_{slope}$ สามารถหาได้จากสมการ

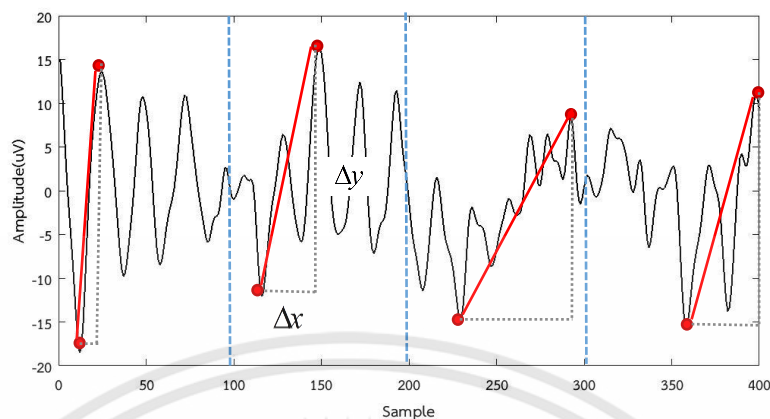
$$avgMMD_{slope} = \frac{\sum_{i=1}^w |m_i|}{w} \quad (2.12)$$

$$sumMMD_{slope} = \sum_{i=1}^w |m_i| \quad (2.13)$$

โดยที่ m ได้มาจากการคำนวณ

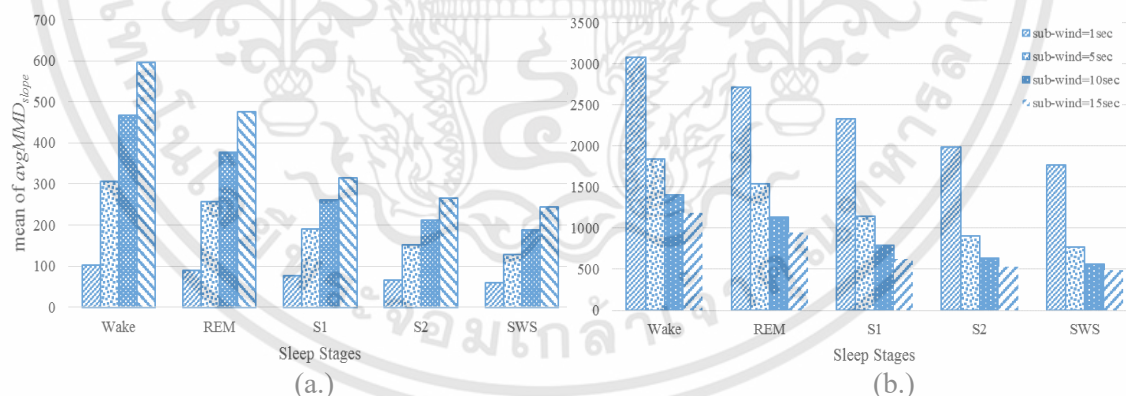
$$m = \sqrt{\frac{|\Delta y|^2}{|\Delta x|^2}} \quad (2.14)$$

โดยที่ Δx และ Δy คือ ผลต่างระหว่างจุดสูงสุดและต่ำสุดของแกน x และแกน y ตามลำดับ ดังรูปที่ 2.4



รูปที่ 2.5 จุดสูงสุดและต่ำสุดของสัญญาณ EEG ในแต่ละ window ที่มา [25]

นอกจากนี้เราได้ทดลองกำหนดขนาดของ Window ต่างๆ ซึ่งเท่ากับ 1 วินาที, 5 วินาที, 10 วินาที และ 15 วินาที โดยกำหนดตามความยาวของรูปคลื่น EEG ขณะนอนหลับ แสดงดังรูปที่ 2.6a และ 2.6b แสดงผลการทดลองการเปรียบเทียบค่าเฉลี่ยของ $avgMMD_{slope}$ และ $sumMMD_{slope}$ ด้วยขนาด window ต่างๆ ที่เกิดขึ้นแต่ละระยะการนอนจะเห็นได้ว่าค่าเฉลี่ยของคุณลักษณะทั้งสองมีความแตกต่างกันอย่างชัดเจน นอกจากนี้เราได้ทดสอบสมมติฐานความแตกต่างของแต่ละกลุ่มด้วยการทดสอบการวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) โดยค่า P-value ที่ได้มีค่าน้อยกว่า 0.05 มากๆ หรือ P-value $\ll 0.05$ ในทั้งสองคุณลักษณะ จึงสามารถสรุปได้ว่า $avgMMD_{slope}$ และ $sumMMD_{slope}$ สามารถแสดงความแตกต่างของแต่ละระยะการนอนได้อย่างมีนัยสำคัญทางสถิติ



รูปที่ 2.6 (a.) ความสัมพันธ์ระหว่างค่าของ $avgMMD_{slope}$ ในแต่ละระยะการนอน, (b.) ความสัมพันธ์ระหว่างค่าของ $sumMMD_{slope}$ ในแต่ละระยะการนอน ที่มา: [16]

คุณลักษณะของสัญญาณ EEG ทางโดเมนความถี่เป็นคุณลักษณะที่ถูกสกัดเพื่อบ่งบอกลักษณะโครงสร้างสเปกตรัมของสัญญาณ EEG วิธีการคำนวณสเปกตรัมของสัญญาณ EEG ที่ถูกนำเสนอในงานวิจัยที่ผ่านมา ได้แก่ งานวิจัยของ [26] ได้นำเสนอคุณลักษณะที่สกัดจากสเปกตรัมแต่ละย่านความถี่

ของสมอง เริ่มต้นด้วยการคำนวณสเปกตรัมในย่านความถี่ของสมองด้วยวิธี Fast Fourier Transform (FFT) และประเมินสเปกตรัมความถี่ด้วยวิธี Welch จากนั้นคำนวณร้อยละอัตราส่วนระหว่างแต่ละย่านความถี่ของสมองและย่านความถี่ทั้งหมดหรือเรียกว่า ความสัมพันธ์ระหว่างร้อยละของค่าพลังงานแต่ละย่านความถี่กับค่าพลังงานรวมของย่านความถี่ทั้งหมด (The relative percent of spectral energy band :RPEB) สามารถคำนวณได้จากสมการ

$$RPEB = \frac{PEB}{TPC} \times 100 \quad (2.15)$$

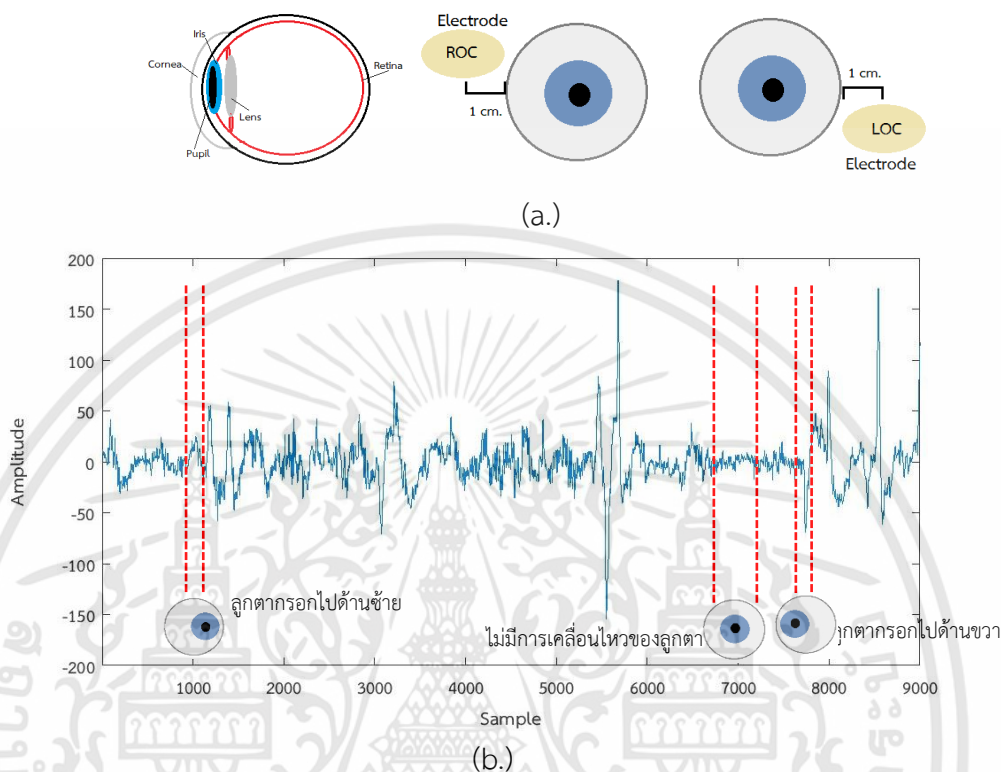
โดยที่ PEB คือ พลังงานของแต่ละย่านความถี่ ประกอบด้วย ย่านความถี่เดลต้า (PEB_{δ}) เทต้า (PEB_{θ}) เอลฟา (PEB_{α}) เบต้า (PEB_{β}) และแกมมา (PEB_{γ}) และ TPC คือ ผลรวมของย่านความถี่ทั้งหมด

การแปลงเวฟเล็ตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform: DWT) เป็นเครื่องมือการประมวลผลสัญญาณที่ถูกนำมาใช้สำหรับการคำนวณคุณลักษณะเพื่อแสดงการกระจายความถี่และเวลานอกจากนี้เป็นเครื่องมือที่มีประสิทธิภาพสำหรับการวิเคราะห์และการแยกแยะคุณลักษณะของสัญญาณแบบ non-stationary ซึ่งเหมาะสมสำหรับสัญญาณ EEG การวิเคราะห์ DWT เป็นวิธีที่แสดงให้เห็นความแตกต่างกันของสัญญาณได้อย่างชัดเจน เมื่อสัญญาณ EEG เข้าสู่กระบวนการแยก DWT สัญญาณจะถูกแบ่งออกเป็น 2 ส่วนคือ ส่วนการประมาณค่า (Approximation) หรือส่วนย่านความถี่ต่ำ และส่วนรายละเอียด (Detail) หรือส่วนย่านความถี่สูง สำหรับจำนวนของระดับการแยกรายละเอียดสัญญาณขึ้นอยู่กับส่วนประกอบความถี่ที่สำคัญของสัญญาณ ในการประยุกต์ใช้กับย่านความถี่ของสมองขณะนอนหลับ ในงานวิจัย [27] ได้กำหนดเวฟเล็ตแม่หรือ Mather wavelet ให้เป็น Daubechies8 (db8) และกำหนดระดับของการจำแนกความถี่ (Level) เท่ากับ 5 ซึ่งผลการทดลองของงานวิจัยดังกล่าวแสดงให้เห็นว่า การกำหนดค่าพารามิเตอร์ของการแปลงเวฟเล็ตเหล่านั้นมีความสัมพันธ์กับแต่ละย่านความถี่ของสัญญาณ EEG ที่ต้องการ คือ ย่านความถี่ช่วง 0-30 เฮิรตซ์

2.3.2 สัญญาณไฟฟ้าการเคลื่อนไหวของกล้ามเนื้อลูกตาหรือสัญญาณ EOG

สัญญาณ EOG เป็นการวัดการเปลี่ยนแปลงความต่างศักย์ทางไฟฟ้าทางด้านบวกลูกตาของกระจกตา (Corneal) ที่มีความสัมพันธ์กับด้านลบที่ได้มาจากเรตินา (Retina) ดังรูป 2.6 การวิเคราะห์การนอนตามมาตรฐาน AASM ตำแหน่งของอิเล็กโทรดขั้วบวกลูกตาอยู่บริเวณตำแหน่งเหนือหางตาด้านซ้าย (Left electro-oculogram: LOC) และบริเวณล่างหางตาด้านขวา (Right electro-oculogram: ROC) ระยะห่างจากตาประมาณ 1 เซนติเมตรดังรูปที่ 2.7 (a.) มีจุดประสงค์เพื่อวัดการเปลี่ยนแปลงของการเคลื่อนไหวของตาจริงระหว่างที่มีการเคลื่อนไหวตา ตัวอย่างเช่น รูปที่ 2.7 (b.) แสดงสัญญาณ EOG ที่วัดจากตำแหน่ง LOC เมื่อไม่มีการเคลื่อนไหวของลูกตาค่าความต่างศักย์อยู่ในระดับศูนย์ หากลูกตาเคลื่อนไหวไปในทิศทาง

ด้านซ้าย สัญญาณ EOG ค่าความต่างศักย์จะมีค่าเป็นบวก และหากเคลื่อนไหวไปในทิศทางด้านขวา สัญญาณ EOG ความต่างศักย์จะมีค่าเป็นลบ



รูปที่ 2.7 (a.) ส่วนประกอบของลูกตาและตำแหน่งของอิเล็กโทรดของการวัดสัญญาณ EOG ที่มา [28]
(b.) ความสัมพันธ์ของการเคลื่อนไหวของลูกตาและสัญญาณ EOG ที่ตำแหน่งอิเล็กโทรด LOC

วิธีการสกัดคุณลักษณะที่ได้จากสัญญาณ EOG ด้วยเทคนิคโดเมนเวลา ประกอบด้วยการคำนวณทางสถิติอย่างง่าย ได้แก่ ค่าสูงสุดและต่ำสุดของขนาดสัญญาณ EOG ค่า kurtosis และค่า skewness นอกจากนี้คุณลักษณะทางโดเมนเวลาอื่นๆ ของสัญญาณ EOG ที่มีผลต่อการวิเคราะห์การนอน มีดังนี้

1. ค่าความแปรปรวนของสัญญาณ EOG เป็นค่าที่ใช้วัดกำลังงานของสัญญาณ สามารถคำนวณได้จากสมการ

$$VAR = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (2.16)$$

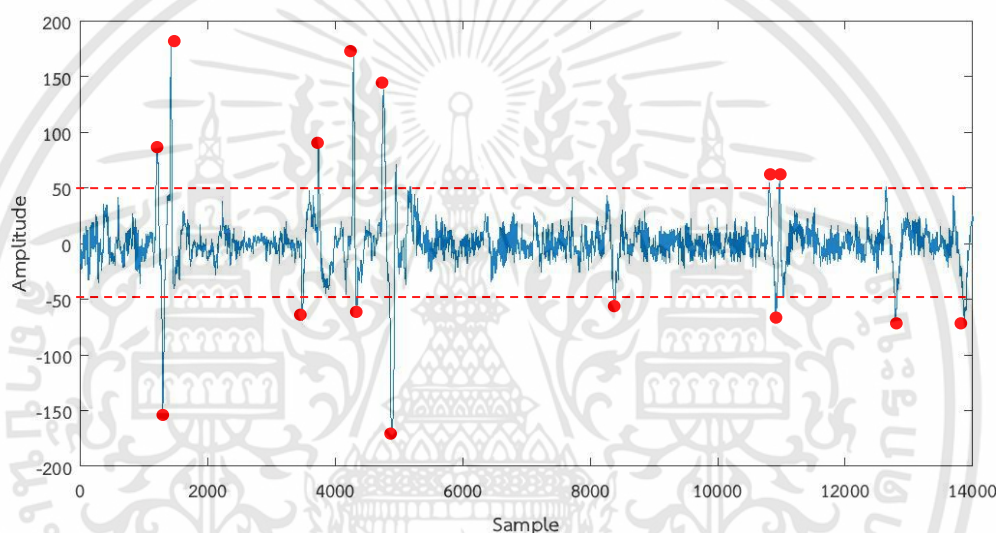
โดยที่ x_i เป็นจำนวนตัวอย่างของสัญญาณ EOG และ N เป็นขนาดของ window

2. ค่าของพื้นที่ใต้กราฟ (Area under curve: AUC) พื้นที่ใต้กราฟของสัญญาณ EOG คือ ผลบวกของค่าสัมบูรณ์ของแอมพลิจูดทั้งกราฟที่มีค่าบวกและค่าลบ สามารถคำนวณจากสมการ

$$AUC = \sum_{i=1}^N |x_i| \quad (2.17)$$

โดยที่ x_i เป็นจำนวนตัวอย่างของสัญญาณ EOG และ N เป็นขนาดของ window

3. จำนวนของค่าตรวจจับผ่านระดับอ้างอิง (Threshold Crossing Value) คือ จำนวนครั้งของสัญญาณ EOG ที่มีขนาดสูงหรือต่ำกว่าระดับอ้างอิงที่กำหนดไว้ดังรูป 2.8 ในผลการทดลองของงานวิจัย [28] ได้นำเสนอว่า คุณลักษณะนี้สามารถแสดงความแตกต่างของแต่ละระยะการนอนได้ โดยเฉพาะความแตกต่างของระยะ REM และระยะ NREM ที่เห็นได้อย่างชัดเจน แต่ความแตกต่างระหว่างระยะ REM และระยะตื่นตัวยังไม่มีความแตกต่างที่เห็นได้อย่างชัดเจน



รูปที่ 2.8 ค่าสัญญาณ EOG ที่มีค่าสูงกว่าและต่ำกว่าระดับอ้างอิง

การคำนวณคุณลักษณะของสัญญาณ EOG ด้วยเทคนิคโดเมนความถี่ และโดเมนเวลา-ความถี่ คล้ายกับวิธีการคำนวณคุณลักษณะของสัญญาณ EEG เนื่องจากมีคุณสมบัติทางความถี่อยู่ในช่วงความถี่ที่ใกล้เคียงกัน

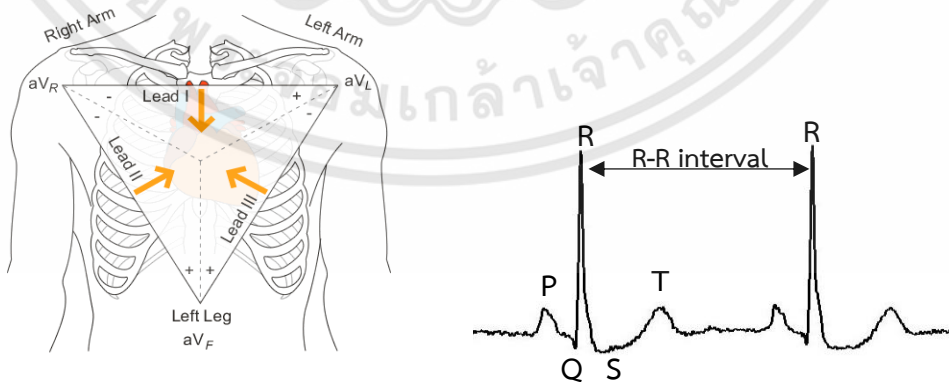
2.3.3 สัญญาณไฟฟ้าหัวใจหรือสัญญาณ ECG

สัญญาณ ECG ได้ถูกยอมรับทางการแพทย์ว่าเป็นสัญญาณที่ถูกนำมาใช้เพื่อระบุและบ่งบอกการเปลี่ยนแปลงของร่างกายในขณะนอนหลับอย่างชัดเจน เนื่องจากการทำงานของสมองและหัวใจมีความสัมพันธ์กันโดยตรง สัญญาณ ECG ถูกสกัดคุณลักษณะด้วย วิธีการวิเคราะห์ความแปรปรวนของอัตราการเต้นของหัวใจ หรือ Heart Rate Variability Analysis (HRV) และวิธีการสกัดสัญญาณการหายใจจากสัญญาณ ECG หรือ Electrocardiography Derived Respiratory (EDR)

2.3.3.1 การวิเคราะห์ความแปรปรวนของอัตราการเต้นของหัวใจหรือ Heart Rate Variability Analysis (HRV)

ความแปรปรวนอัตราการเต้นของหัวใจได้รับการยอมรับทางการแพทย์ว่าเป็นเครื่องมือที่สามารถทำให้เข้าใจการทำงานของระบบประสาทอัตโนมัติ (Autonomic Nervous System: ANS) โดยผลของ HRV สามารถบ่งบอกความสมดุลของการทำงานระหว่างระบบประสาทอัตโนมัติซิมพาเทติก (Sympathetic Nervous System : SNS) และระบบประสาทอัตโนมัติพาราซิมพาเทติก (Parasympathetic Nervous System: PNS) การทำงานของระบบประสาท ANS จะทำงานอย่างอิสระโดยไม่อยู่ภายใต้อำนาจจิตใจหรือการควบคุมของระบบประสาทส่วนกลาง การทำงานของระบบประสาททั้งสองส่วนทำงานในทิศทางตรงกันข้ามกัน ระบบประสาทอัตโนมัติซิมพาเทติกจะทำงานในกรณีที่ร่างกายตกอยู่สภาวะฉุกเฉิน ร่างกายจะเกิดปฏิกิริยาตื่นตัวเพื่อเตรียมพร้อมที่ต่อสู้หรือหนีจากสถานการณ์เหล่านั้น ตัวอย่างปฏิกิริยาของร่างกายที่เกิดขึ้นเมื่อระบบประสาทอัตโนมัติส่วนซิมพาเทติกทำงาน ได้แก่ อัตราการเต้นของหัวใจสูงขึ้น ความดันโลหิตเพิ่มสูงขึ้น ต่อมฮอร์โมนอะดรีนาลินทำงานเพิ่มขึ้น เป็นต้น ในกรณีที่ร่างกายพ้นสภาวะฉุกเฉินแล้วหรือสภาวะนิ่งสงบ ระบบประสาทอัตโนมัติส่วนพาราซิมพาเทติกจะทำงานซึ่งช่วยทำให้ร่างกายกลับคืนสู่สภาวะปกติและปรับสมดุลของร่างกาย ได้แก่ ปรับอัตราการเต้นของหัวใจให้ลดลง ความดันโลหิตกลับสู่สภาวะปกติ เป็นต้น

การวิเคราะห์ HRV คำนวณจากความแปรปรวนของระยะห่างระหว่างคลื่นอาร์ที่อยู่ติดกันหรือเรียกสั้นๆว่า ระยะห่างระหว่างคลื่นอาร์ (RR interval) ซึ่งคลื่นอาร์นี้เป็นคลื่นที่สูงที่สุดของสัญญาณ ECG การบันทึกสัญญาณ ECG เป็นการตรวจทางการแพทย์พื้นฐานเพื่อดูกิจกรรมการทำงานของหัวใจในช่วงเวลาหนึ่งๆเป็นการตรวจแบบภายนอกโดยไม่ต้องเจาะทะลุผิวหนังหรือเข้าไปในร่างกาย (Noninvasive) การติดอิเล็กโทรดตามตำแหน่งต่างๆ ตามมาตรฐาน 3 lead electrode placement แสดงได้ดังรูปที่ 2.9 (a.) สัญญาณ ECG ประกอบด้วย คลื่น P,Q,R,S และ T ดังรูปที่ 2.9 (b.)

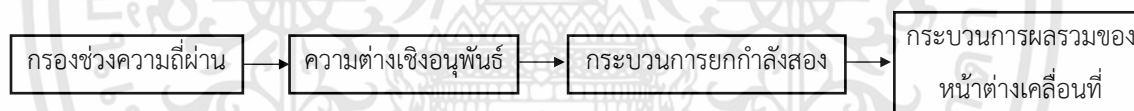


รูปที่ 2.9 (a.) ตำแหน่งของอิเล็กโทรดตามมาตรฐาน 3 lead electrode placement
(b.) สัญญาณ ECG

ขั้นตอนการวิเคราะห์ HRV มีดังนี้

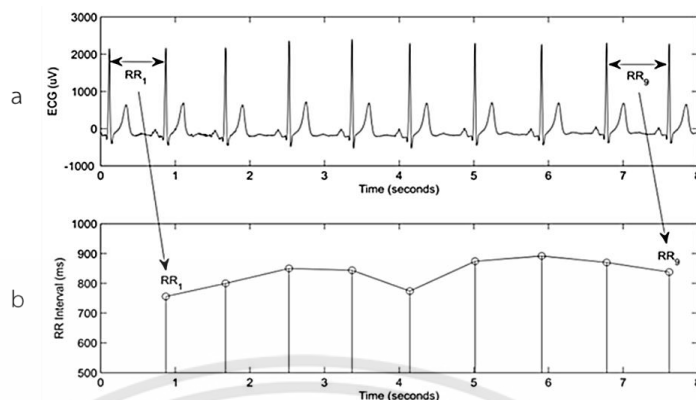
1. ขั้นตอนการประมวลผลก่อนการวิเคราะห์ HRV (Preprocessing)

การวิเคราะห์ HRV มาจากการวิเคราะห์การเปลี่ยนแปลงของ RR interval ดังที่กล่าวไว้ข้างต้น ดังนั้นขั้นตอนการตรวจจับจุดสูงสุดของสัญญาณ ECG หรือคลื่นอาร์ (R peak) มีความสำคัญอย่างยิ่ง เนื่องจากมีผลโดยตรงกับการวิเคราะห์ HRV การตรวจจับคลื่นอาร์ถูกนำเสนอหลากหลายวิธี แต่ในงานวิจัยนี้ได้นำขั้นตอนการตรวจจับคลื่นอาร์ที่นำเสนอโดย Pan และ Tomkins [29] มาใช้ เนื่องจากเป็นวิธีที่มีความแม่นยำในการตรวจจับคลื่นอาร์และนิยมนำมาใช้กันอย่างแพร่หลาย เป็นวิธีที่อ้างอิงจากความชัน แอปปริจูด และความกว้างของสัญญาณ QRS Complex กระบวนการตรวจจับคลื่นอาร์แสดงดังรูปที่ 2.10 ขั้นตอนแรกเริ่มต้นด้วย สัญญาณ ECG ผ่านกระบวนการกรองสัญญาณความถี่ต่ำผ่าน (Low pass frequency) ที่ความถี่ตัด (cutoff frequency) 50 เฮิร์ตซ์ เพื่อกำจัดสัญญาณรบกวน power line และกระบวนการกรองความถี่สูงผ่าน (High pass filter) ที่ความถี่ตัด 5 เฮิร์ตซ์ เนื่องจากเป็นย่านความถี่ของคลื่นอาร์ จากนั้นเข้าสู่กระบวนการหาความต่างเชิงอนุพันธ์ (Differentiator) เพื่อหาค่าความชันของสัญญาณ QRS และผ่านกระบวนการยกกำลังสอง (Squaring) เพื่อให้สัญญาณเป็นด้านบวกทั้งหมด ขั้นตอนสุดท้ายผ่านกระบวนการผลรวมของหน้าต่างเคลื่อนที่ (moving window integration) เพื่อหาจุดสูงสุดหรือคลื่นอาร์ในแต่ละ window



รูปที่ 2.10 แผนภาพของขั้นตอนการตรวจจับคลื่นอาร์

เมื่อได้คลื่นอาร์มาแล้ว จากนั้นคำนวณระยะห่างระหว่างคลื่นอาร์ (RR interval) ด้วยสมการ $RR(i) = RR(i+1) - RR(i)$, $i=1, 2, \dots, n-1$ โดยที่ $RR(i)$ คือ ระยะห่างของคลื่นอาร์ที่ตำแหน่ง i^{th} แสดงดังรูปที่ 2.11a. การพล็อตสัญญาณ HRV กำหนดแกน y คือค่า RR interval และแกน x คือ ตำแหน่งของคลื่นอาร์ที่ i แสดงดังรูปที่ 2.11b. การแปลงค่าระหว่าง RR interval ที่วัดได้ให้เป็นอัตราการเต้นของหัวใจ (Heart rate) คำนวณได้จากสมการ $HR (BPM) = 60 (sec) / RR \text{ interval } l(sec)$ วิธีการตรวจจับคลื่นอาร์บางครั้งยังคงมีความผิดพลาดในการตรวจจับเกิดขึ้นหรือเรียกว่า ectopic beat ซึ่งได้ถูกแก้ไขด้วยวิธีการประมาณค่าในช่วงหรือ Interpolation ซึ่งหากข้อมูลมีค่ามากกว่าหรือต่ำกว่าระดับอ้างอิงที่กำหนดไว้ ค่านั้นจะถูกทดแทนข้อมูลด้วยวิธี interpolation ด้วยฟังก์ชัน Cubic และสุดท้ายสัญญาณที่ได้เรียกว่า NN interval หรือ Normal to Normal interval



รูปที่ 2.11. (a.) สัญญาณ ECG (b.) RR interval time series

2. ขั้นตอนการสกัดคุณลักษณะของสัญญาณ ECG ด้วยเทคนิคการวิเคราะห์ HRV แบบเชิงเส้น

การวิเคราะห์ด้วยเทคนิคเชิงเส้นเป็นที่รู้จักในรูปแบบของการวิเคราะห์ HRV ทางโดเมนเวลา (Time Domain) ทางโดเมนความถี่ (Frequency Domain) และทางโดเมนเวลา - ความถี่ (Time-Frequency Domain) พารามิเตอร์ของการวิเคราะห์ HRV ทางโดเมนเวลาสามารถคำนวณอย่างง่ายด้วยวิธีทางสถิติ ถูกแบ่งเป็นสองกลุ่ม คือ กลุ่มแรกได้จากการคำนวณ NN interval โดยตรง หรือ First order statistic ประกอบด้วย ค่าเฉลี่ยของสัญญาณ NN interval (Mean RR interval: Mean RRI), ค่าเฉลี่ยอัตราการเต้นของหัวใจ (Mean heart rate: Mean HR), ความแตกต่างระหว่างค่ามากที่สุดและค่าน้อยที่สุดของ NN interval ในช่วงเวลาหนึ่งๆ, ค่าเบี่ยงเบนมาตรฐาน NN interval (Standard Deviation of NN: SDNN) และรากที่สองของความแปรปรวนของ NN interval กลุ่มที่สองได้จากการคำนวณความต่างระหว่าง NN interval หรือ Second order statistic ประกอบด้วย รากที่สองของค่าเฉลี่ยกำลังสองของความต่างระหว่าง NN interval (Root mean square value of successive differences: RMSSD), ร้อยละของอัตราส่วนระหว่างผลรวมของจำนวนความต่างของ NN interval ที่มีค่ามากกว่า 50 มิลลิวินาที ต่อจำนวน NN interval ทั้งหมด (The percentage of successive interval of more/less than 50milli second: pNN50), ค่าเบี่ยงเบนมาตรฐานของความต่างระหว่าง NN interval (Standard deviation of successive differences: SDSD) ตัวอย่างการคำนวณพารามิเตอร์ของโดเมนเวลาที่ถูกคำนวณด้วยวิธีทางสถิติ เช่น ค่าเฉลี่ย NN = 680.80 มิลลิวินาที, SDNN=106.93 มิลลิวินาที, SDSD=29.91 มิลลิวินาที, RMSSD=41.74 มิลลิวินาที และ pNN=17.15%

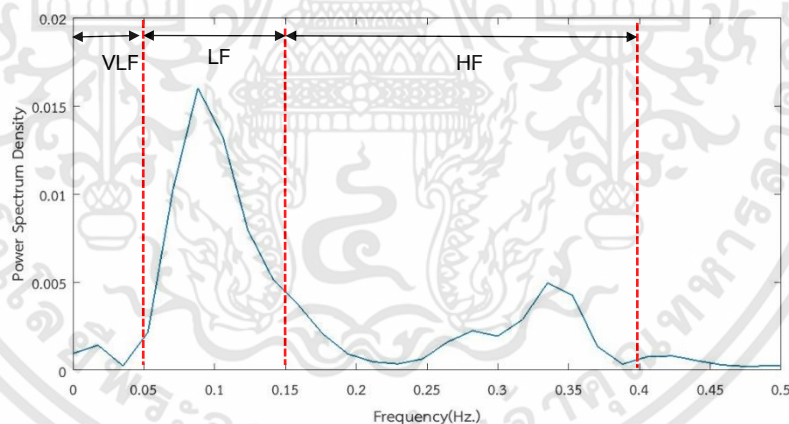
การวิเคราะห์ทางโดเมนความถี่เป็นเทคนิคการประเมินจากความหนาแน่นของพลังงานสเปกตรัม (Power spectrum density: PSD) ในย่านความถี่ต่างๆ การประเมิน PSD ถูกประเมินทั้งวิธี Parametric และวิธี Nonparametric แต่ละวิธีมีข้อดีแตกต่างกัน ข้อดีของวิธี Parametric คือ องค์ประกอบของสเปกตรัมมีความราบเรียบมากกว่าทำให้ง่ายต่อการวิเคราะห์แต่ละช่วงความถี่ของ HRV และสามารถ

แยกแยะความถี่กลางของแต่ละองค์ประกอบได้ง่าย แต่ข้อเสียของวิธี Parametric คือ ความซับซ้อนแต่ละโมเดล เพราะแต่ละโมเดลต้องการฟังก์ชันที่เหมาะสมแตกต่างกัน สำหรับข้อดีของวิธี Non-Parametric คือ เป็นวิธีง่ายและใช้กันอย่างทั่วไป ตัวอย่างเช่น วิธีการประเมินย่านความถี่ด้วยเทคนิค Welch ย่านความถี่ของสเปกตรัมสำหรับการวิเคราะห์ HRV ถูกแบ่งได้ดังรูปที่ 2.12 ประกอบด้วย ช่วงความถี่ต่ำมาก (Very low frequency: VLF ช่วงย่านความถี่ต่ำกว่า 0.04 เฮิรตซ์), ช่วงความถี่ต่ำ (Low frequency: LF ย่านความถี่ช่วง 0.04-0.15 เฮิรตซ์), และช่วงความถี่สูง (High Frequency: HF ย่านความถี่ช่วง 0.15-0.4 เฮิรตซ์) พลังงานสเปกตรัมของแต่ละย่านความถี่ถูกคำนวณจากพื้นที่ใต้กราฟของแต่ละย่านความถี่ และย่านความถี่ LF และ HF ถูกนอร์มอลไลซ์ (Normalized unit) ด้วยสมการ

$$LF(n.u.) = \frac{LF}{LF + HF} \quad (2.18)$$

$$HF(n.u.) = \frac{HF}{LF + HF} \quad (2.19)$$

แต่ละย่านความถี่จะสะท้อนและบ่งบอกถึงการทำงานของระบบประสาทอัตโนมัติ คือ ช่วง HF สะท้อนถึงการทำงานของประสาทอัตโนมัติ Vagal ขณะที่ช่วง LF สะท้อนถึงการทำงานของระบบประสาทอัตโนมัติ ส่วนซิมพาเทติกและอัตราส่วนระหว่าง LF และ HF สะท้อนถึงความสมดุลของระบบประสาท ANS



รูปที่ 2.12 ย่านความถี่ของสเปกตรัมสำหรับการวิเคราะห์ HRV

3. ขั้นตอนการสกัดคุณลักษณะของสัญญาณ ECG ด้วยเทคนิคการวิเคราะห์ HRV แบบไม่เชิงเส้น การทำงานของระบบหัวใจและหลอดเลือดเป็นระบบไม่เชิงเส้นเพราะการเต้นของหัวใจแต่ละครั้งจะขึ้นอยู่กับการกระตุ้นของระบบภายในและภายนอกร่างกาย บางกรณีเทคนิคระบบเชิงเส้นไม่สามารถอธิบายการควบคุมเชิงกลของระบบทำงานของหัวใจและหลอดเลือดได้อย่างครอบคลุม ดังนั้นหลายงานวิจัยได้นำทฤษฎีของระบบไดนามิกแบบไม่เชิงเส้นมาใช้ประยุกต์สำหรับการอธิบายระบบดังกล่าว ในหัวข้อนี้ได้กล่าวถึงการวิเคราะห์ HRV ด้วยเทคนิคแบบไม่เชิงเส้น โดยอ้างอิงจากงานวิจัยของ Meng Xiao และคณะ [30] และจากการศึกษาของ Ayyoob Jafari [31] โดยพวกเขาแนะนำเสนอวิธีการวิเคราะห์ HRV

ด้วยเทคนิคแบบไม่เชิงเส้นมาใช้สำหรับการตรวจจับและจำแนกระยะการนอน และการตรวจจับโรคหยุดหายใจขณะนอนหลับ การวิเคราะห์ HRV ที่ได้จากเทคนิคไม่เชิงเส้นมีรายละเอียดดังนี้

1. Nonlinear Recurrence Space การวิเคราะห์แบบไม่เชิงเส้นเริ่มต้นด้วยการหา ลักษณะเฉพาะของระบบไดนามิกของสัญญาณ time series ก่อน เริ่มจากคำนวณเพื่อหาการเกิดซ้ำของ มิติตัวแทนของสถานะ (state) ที่เป็นไปได้ของระบบ (phase space) เรากำหนดให้สัญญาณ time series ในที่นี้จะป็นสัญญาณ RR interval เป็น $S = \{s_1, s_2, \dots, s_n\}$ และคำนวณ Trajectory ของแต่ละ attractor $X = \{x_1, x_2, \dots, x_n\}$ ได้ดังสมการ

$$X(n) = [S(n), S(n - \tau_d), \dots, S(n - (d - 1)\tau_d)] \quad (2.20)$$

โดยที่ τ_d คือ Time delay ที่ถูกประเมินจากฟังก์ชัน Autocorrelation ของสัญญาณ Time series และ d ได้จากกระบวนการ embedding dimension สำหรับงานวิจัยนี้เราได้ประเมินสัญญาณ RR interval ด้วยฟังก์ชัน autocorrelation ตามงานวิจัยของเราที่ได้นำเสนอใน [32] ซึ่งผลการทดลองที่ได้คือ τ_d เท่ากับ 5 และ d เท่ากับ 4 ดังนั้นการกำหนดพารามิเตอร์ของการทดลองในงานวิจัยนี้เท่ากับค่าดังกล่าว

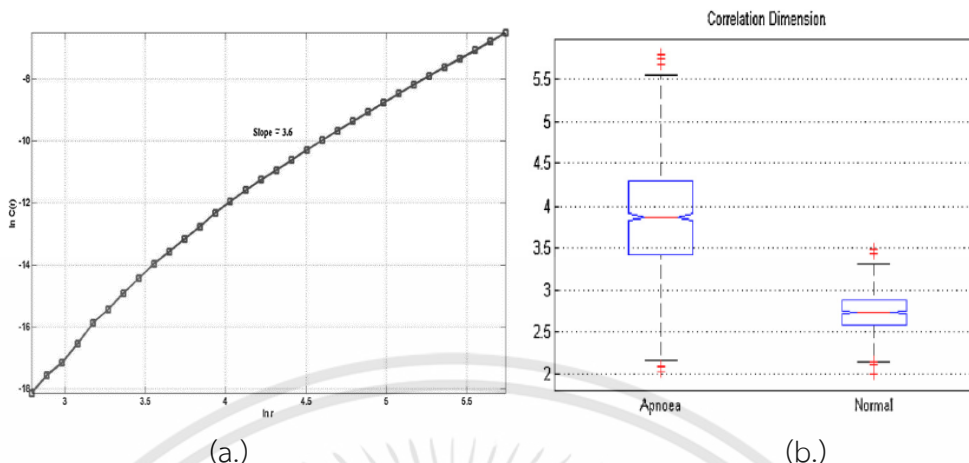
2. Correlation Dimension (CD) เป็นการวัดเชิงปริมาณของธรรมชาติของ trajectory ของการเปลี่ยนแปลงมิติใน Phase space จากการศึกษาของงานวิจัยที่ผ่านมาพบว่าขนาดของ CD มีความสัมพันธ์กับการวิเคราะห์ HRV ในหลายๆโรค โดยค่า CD ที่ถูกนำมาใช้สำหรับการวิเคราะห์ HRV คำนวณจากสมการ

$$CD = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log(r)} \quad (2.21)$$

โดยที่ $C(r)$ คือผลรวมของ correlation จาก

$$C(r) = \frac{1}{N^2} \sum_{X_i=1}^N \sum_{X_j=1}^N \Theta(r - \|X_i - X_j\|) \quad (2.22)$$

โดยที่ X_i และ X_j คือ ตำแหน่งของ trajectory ใน phase space ตามสมการ 2.20, $\|X_i - X_j\|$ คือค่าระยะทาง Euclidean distance ระหว่าง X_i และ X_j , N คือจำนวนทั้งหมดของข้อมูลใน phase space, r คือระยะทางรัศมีรอบจุดอ้างอิง X_i และฟังก์ชัน Θ คือ Heaviside function พารามิเตอร์ที่นำไปใช้สำหรับการวิเคราะห์ HRV คือ ค่าความชันระหว่าง $\log C(r)$ และ $\log(r)$ แสดงดังรูปที่ 2.13(a) ซึ่งเป็นค่าความชันระหว่าง $\log C(r)$ และ $\log(r)$ ของสัญญาณ RR ผลการทดลองของงานวิจัย [31] แสดงให้เห็นว่าค่าพารามิเตอร์ CD ของช่วงภาวะหยุดหายใจขณะนอนหลับมีค่าสูงกว่าช่วงภาวะปกติแสดงดังรูปที่ 2.13(b)



รูปที่ 2.13 (a.) กราฟความสัมพันธ์ระหว่าง $\log C(r)$ และ $\log(r)$ ของสัญญาณ RR ช่วงการเกิดการเกิดภาวะหยุดหายใจขณะนอนหลับ (b.) กราฟ boxplot การเปรียบเทียบค่า CD ของภาวะหยุดหายใจขณะนอนหลับและช่วงหายใจปกติ ที่มา: [31]

3. Detrended Fluctuation Analysis (DFA) เป็นพารามิเตอร์วัดปริมาณของคุณสมบัติสเกลแฟร็กทัลของช่วงสัญญาณ time series สั้นๆ นำสัญญาณ RR time series หรือ $x(i) = \{x(1), x(2), \dots, x(N)\}$ โดยที่ N เท่ากับความยาวของสัญญาณ ผลรวมของสัญญาณ RR time series ที่ถูกนอร์มอไรซ์ด้วยค่าเฉลี่ยคำนวณได้ดังสมการ

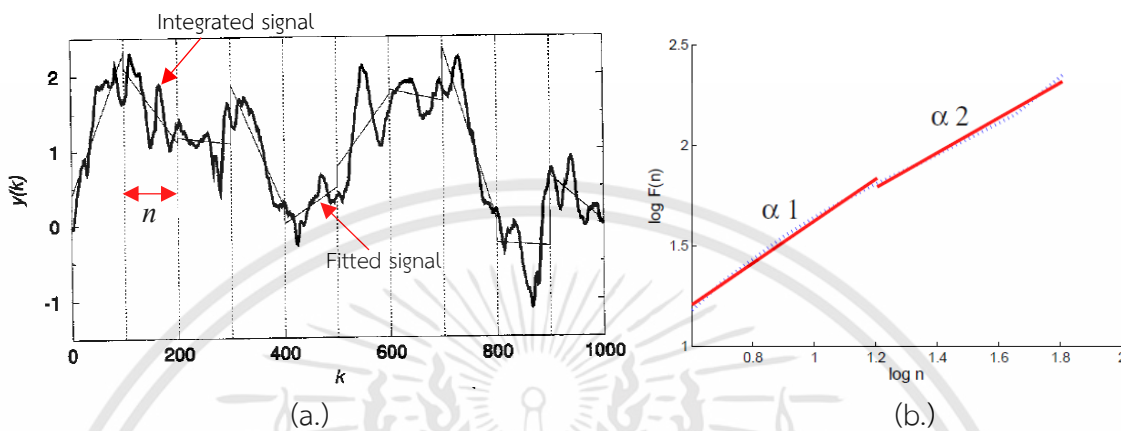
$$y(k) = \sum_{i=1}^N (x(i) - x_{average}) \quad (2.23)$$

จากรูปที่ 2.14(a) สัญญาณเส้นหนาที่บแทนด้วยสัญญาณรวมที่ถูกนอร์มอไรซ์แล้วเท่ากับ $y(k)$ จากนั้นถูกแบ่งเป็นช่วง (segment) ด้วยขนาดความกว้างของวินโดว์ที่กำหนดไว้มีขนาดเท่ากับ n และสัญญาณในแต่ละวินโดว์ถูกนำไปหาแนวโน้มของข้อมูล (fitted trend) ด้วยวิธี least square ของแต่ละช่วงแทนด้วย $y_n(k)$ หรือจากรูปที่ 2.14 (a) แทนด้วยเส้นบางที่บ จากนั้นคำนวณ $F_n(n)$ ด้วยรากที่สองของกำลังสองเฉลี่ยของสัญญาณที่ถูกนอร์มอไรซ์และ fitted trend ดังสมการ

$$F_n = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2} \quad (2.24)$$

ค่าพารามิเตอร์ DFA หาได้จากค่าความชันจากกราฟความสัมพันธ์ระหว่าง $\log F(n)$ และ $\log(n)$ ตัวอย่างแสดงดังรูปที่ 2.14 (b) ในงานวิจัยของ Ayyoob Jafari [31] ได้นำข้อมูลสัญญาณ RR time series ที่ถูกกำหนดขนาดของหน้าต่างด้วยจังหวะการเต้นของหัวใจ 2 ช่วงได้แก่ช่วง short term และช่วง long term จากรูป 2.14 (b) แสดงผลการทดลองของการพล็อต DFA ค่าความชันของช่วง short term α_1 (อ้างอิงจากขนาดของ $1 \leq n \leq 4$) และค่าความชันของช่วง long term α_2 (อ้างอิงจากขนาดของ $4 \leq n$)

≤ 8) และจากงานวิจัยดังกล่าวได้เปรียบเทียบค่าความชันดังกล่าวระหว่างช่วงที่เกิดภาวะหยุดหายใจขณะนอนหลับและช่วงการนอนปกติ ค่าความชันของ DFA ในช่วงที่เกิดภาวะหยุดหายใจขณะนอนหลับมีค่าสูงกว่าช่วงการนอนปกติ



รูปที่ 2.14 (a.) สัญญาณ RR time series ที่ถูกนอร์มอไรซ์ (b.) กราฟความสัมพันธ์ระหว่าง logarithm $F(n)$ และ logarithm ของขนาด n ในแต่ละโซน ที่มา: [31]

4. Mutual information (MI) เป็นการวัดปริมาณของข้อมูลร่วมระหว่างตัวแปรสองตัวแปรที่มีความสัมพันธ์กัน MI ของตัวแปร X และ Y สามารถหาได้จากสมการ

$$MI(x, y) = H(x) + H(y) - H(x, y) \quad (2.25)$$

ซึ่ง $H(x)$ และ $H(y)$ คือ เอนโทรปี (Entropy) ของแต่ละตัวแปร X และ Y และ $H(x, y)$ คือเอนโทรปีร่วมระหว่าง X และ Y จากงานวิจัยของ Meng และคณะ [30] ได้นำวิธีการ Mutual Information มาใช้ในการสกัดคุณลักษณะของการวิเคราะห์ HRV และได้กำหนดคุณลักษณะใหม่จำนวน 2 คุณลักษณะ ได้แก่ Beat Decay ซึ่งเป็นการแสดงค่าความสัมพันธ์ระหว่างข้อมูล RR interval กับสัญญาณ RR ที่ time delay และ Peak Decay เป็นการแสดงค่าความสัมพันธ์ระหว่างข้อมูล RR interval กับสัญญาณ RR ที่ Peak สูงสุด โดยทั้งสองคุณลักษณะคำนวณได้จาก $Beat\ Decay = MI(0) - MI(1)$ และ $Peak\ Decay = MI(0) - MI_{peak}$ โดยที่ $MI(0)$ คือ Mutual information ของ RR interval, $MI(1)$ คือ Mutual information ของ RR interval ที่ Time delay และ MI_{peak} คือ Mutual information ของ RR interval ที่ Peak สูงสุด โดยงานวิจัยของ Meng ได้กล่าวว่าคุณลักษณะทั้งสองสามารถแสดงถึงความซับซ้อนของสัญญาณ RR interval ได้

5. Autocorrelation coefficient จากการทดลองของ Meng และคณะ [30] พบว่า Autocorrelation coefficient ของ RR interval มีความแตกต่างกันระหว่างสภาวะการนอน REM และ NREM สำหรับ Correlation function ของสัญญาณ time series $x(n)$ ถูกคำนวณด้วยสมการ

$$r(k) = \frac{Cov[x(n), x(n+k)]}{D[x(n)] \cdot D[x(n+k)]} \quad (2.26)$$

โดยที่ $Cov(.)$ คือความสัมพันธ์ร่วม, $D(.)$ คือค่าเบี่ยงเบนมาตรฐาน และ $x(n+k)$ คือ time delay series ที่ k

6. Zero crossing analysis เป็นการวิเคราะห์ความแปรปรวนของ RR sequence การพิจารณาข้อมูลในแต่ละ sequence คือ $x(n)$ ที่ถูกลบด้วยค่าเฉลี่ยทั้งหมดของข้อมูล และหาตำแหน่งของ zero crossing ของข้อมูลหรือ Zero crossing index (ZCI) สำหรับพารามิเตอร์ที่ใช้สำหรับการวิเคราะห์ HRV ประกอบด้วย ค่าเฉลี่ยของ ZCI และค่าเบี่ยงเบนมาตรฐานของ ZCI

2.3.3.2 การสกัดสัญญาณการหายใจจากสัญญาณ ECG หรือ Electrocardiography Derived Respiratory (EDR)

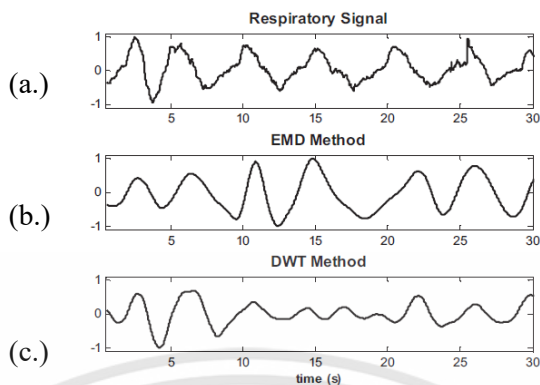
อัตราการหายใจ (Respiratory Rate) คือการตรวจวัดอัตราการหายใจเข้าออกใน 1 นาที อัตราการหายใจของผู้ใหญ่ปกติประมาณ 12-18 ครั้งต่อนาที อัตราการหายใจที่ผิดปกติเป็นพารามิเตอร์หนึ่งที่ถูกนำมาใช้สำหรับบ่งบอกถึงความผิดปกติที่เกิดขึ้นระหว่างการนอนหลับ กระบวนการหายใจทำให้ร่างกายมีการเปลี่ยนแปลงทางกลไกที่เห็นได้อย่างชัดเจน คือ การหายใจเข้าออกทำให้ปริมาณอากาศในปอดเพิ่มขึ้นและลดลงทำให้มีการเปลี่ยนแปลงบริเวณหน้าอก ขณะเดียวกันหากมีการวัดสัญญาณ ECG ขณะที่หายใจเข้าออกก็ทำให้สัญญาณ ECG มีการเปลี่ยนแปลงได้ด้วยเช่นกัน ดังนั้นการวิเคราะห์สัญญาณการหายใจสามารถหาได้จากสัญญาณ ECG ด้วยกระบวนการประมวลผลสัญญาณ หรือ เรียกว่า ECG derived Respiratory (EDR) มีงานวิจัยจำนวนมาก [34-40] ที่ให้ความสนใจนำสัญญาณ EDR มาตรวจวัดสภาวะการหยุดหายใจขณะนอนหลับ การประเมินสัญญาณการหายใจที่ได้จากสัญญาณ ECG ถูกพัฒนาด้วยเทคนิคต่างๆ สามารถสรุปได้ดังตารางที่ 2.3 ดังนี้

ตารางที่ 2.3 การนำเสนอการประเมินสัญญาณ EDR ด้วยเทคนิคต่างๆ

ปี/ผู้แต่ง	เทคนิค	ผลการทดลอง
1992 [34]	การเฝ้าดูการหายใจจากการเปลี่ยนแปลงของ beat to beat ของคลื่น QRS complex และคลื่น T จากสัญญาณ ECG	วิธีการนี้ไม่สามารถทนทานต่อสัญญาณรบกวนที่มีผลโดยตรงกับคลื่น R

2000-2001 [35-36]	เทคนิคการกรองสัญญาณ สำหรับสัญญาณ EDR	ไม่เหมาะสมเนื่องจากย่านความถี่ของสัญญาณรบกวนอยู่ ย่านเดียวกับสัญญาณหายใจ แต่เมื่อเปรียบเทียบกับ วิธีการ[1] เทคนิคนี้มีประสิทธิภาพมากกว่าสำหรับการ กำจัดสัญญาณรบกวนจากสิ่งแวดล้อมภายนอก
2009 [37]	Discrete Wavelet Transform (DWT)	วิธีการนี้สามารถนำไปใช้ได้แต่ยังไม่ดีพอสำหรับการ วิเคราะห์รูปแบบของสัญญาณการหายใจเนื่องจากรูปแบบ สัญญาณจะขึ้นอยู่กับ การเลือกใช้ Mother wavelet ประสิทธิภาพของสัญญาณลดลงและมีผลต่อสัญญาณการ หายใจด้วยเช่นกัน
2011-2013 [38-39]	Empirical Mode Decomposition (EMD)	เทคนิคนี้สามารถสกัดสัญญาณการหายใจที่มี ประสิทธิภาพที่ดีและเหมาะสมที่สุดเมื่อเทียบกับ WT
2012 [40]	Kernel Principle Component Analysis (KPCA)	ใช้เทคนิคการแยกองค์ประกอบหลักที่มีความสัมพันธ์กัน มากที่สุดของสัญญาณ ECG และสัญญาณหายใจ เพื่อสกัด สัญญาณ EDR แต่ขั้นตอนของการ optimize parameter ใน ฟังก์ชัน kernel ที่นำไปใช้ สำหรับการสกัดสัญญาณ การหายใจยังคงยากสำหรับการวิเคราะห์

งานวิจัยของ [41] เสนอว่า การสกัดสัญญาณ EDR ด้วยเทคนิค Empirical Mode Decomposition (EMD) เป็นวิธีที่มีประสิทธิภาพเมื่อเปรียบเทียบกับเทคนิค Discrete Wavelet Transform (DWT) ซึ่งงานวิจัยดังกล่าวได้เปรียบเทียบทั้งสองวิธีด้วยค่าร้อยละสัมพันธ์ผลการทดลองที่ได้คือ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างสัญญาณ EDR ที่ได้จากเทคนิค DWT กับสัญญาณการหายใจจริงเท่ากับร้อยละ 92.14 และ ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างสัญญาณ EDR ที่ได้จากเทคนิค EMD กับสัญญาณการหายใจจริงเท่ากับร้อยละ 93.07 รูปที่ 2.15 ได้แสดงผลการทดลองสัญญาณ EDR ดังกล่าว จากรูปจะเห็นได้ว่าสัญญาณ EDR ที่ได้จากวิธี DWT มีแอมพลิจูดขนาดเล็กกว่าสัญญาณการหายใจจริง เพราะขนาดแอมพลิจูดของสัญญาณ EDR ที่ได้จากเทคนิค DWT ขึ้นอยู่กับ Mother Wavelet ดังนั้นงานวิจัยนี้จึงได้นำเสนอว่าการประเมินสัญญาณ EDR ด้วยวิธี EMD ได้ผลลัพธ์ที่มีประสิทธิภาพดีกว่าวิธี DWT



รูปที่ 2.15 (a.) สัญญาณการหายใจจริง (b.) สัญญาณ EDR จากวิธี EMD
(c.) สัญญาณ EDR จากวิธี DWT ที่มา: [41]

2.4 งานวิจัยที่เกี่ยวข้องกับการวิเคราะห์การนอนแบบอัตโนมัติ

งานวิจัยที่ผ่านมาที่เกี่ยวข้องกับการออกแบบการวิเคราะห์การนอนแบบอัตโนมัติทั้งการจำแนกระยะการนอน และการจำแนกภาวะหยุดหายใจขณะนอนมีจำนวนมาก โดยเราได้แบ่งงานวิจัยเหล่านั้นออกเป็น 2 กลุ่ม ตามจำนวนชุดสัญญาณที่นำมาใช้ในการออกแบบระบบอัตโนมัติ มีดังนี้

2.4.1 การวิเคราะห์การนอนแบบอัตโนมัติด้วยชุดสัญญาณเดียว

วิธีการวิเคราะห์การนอนด้วยชุดสัญญาณเดียวสำหรับระยะการนอนมักจะใช้สัญญาณ EEG หรือสัญญาณ EOG หรือสัญญาณ ECG เพียงสัญญาณใดสัญญาณหนึ่งเท่านั้น ทั้งสามสัญญาณล้วนแสดงให้เห็นถึงการเปลี่ยนแปลงของระยะการนอน ตัวอย่างเช่น ในงานวิจัยของ Liang และคณะ [42] นำเสนอการใช้สัญญาณ EEG เพียงสัญญาณเดียวสำหรับการจำแนกระยะการนอน ซึ่งได้สกัดคุณลักษณะของสัญญาณ EEG ด้วยเทคนิค Multiscale Entropy (MSE) และเทคนิค Autoregressive หลังจากนั้นใช้วิธีการจำแนกระยะการนอนด้วย วิธีการวิเคราะห์แยกแยะกลุ่มเชิงเส้น (Linear discriminant analysis: LDA) ผลการทดลองในงานวิจัยนี้แสดงให้เห็นว่า ค่าความถูกต้องของการจำแนกระยะการนอนเท่ากับร้อยละ 76.91 ในปี 2014 งานวิจัยของ Popovic และคณะ [43] ได้สกัดคุณลักษณะของสัญญาณ EEG ด้วยเทคนิคการวิเคราะห์สเปกตรัม และใช้อัลกอริทึม Decision Tree เพื่อจำแนกระยะการนอน โดยผลการทดลองแสดงค่าความถูกต้องของการจำแนกเท่ากับร้อยละ 80 และงานวิจัยของ Sen และคณะ [27] นำเสนอ การสกัดคุณลักษณะของสัญญาณ EEG หลายช่องสัญญาณ (Multi EEG signal) ด้วยเทคนิคทางโดเมนเวลาและโดเมนความถี่จำนวน 41 คุณลักษณะเพื่อจำแนกระยะการนอน โดยขั้นตอนการลดมิติของงานวิจัยนี้ได้้นำเทคนิคการเลือกคุณลักษณะ ได้แก่ วิธี Minimum Redundancy Maximum Relevance (mRMR), วิธี t-test และ Fisher score algorithm มาใช้เปรียบเทียบ ขั้นตอนการจำแนกข้อมูลได้นำอัลกอริทึม ได้แก่ Random forest, Feed-Forward Neural Network, Decision tree, Support

Vector Machine และ Radial Basis Function Neural Network จากผลการทดลองการเปรียบเทียบ อัลกอริทึมดังกล่าวได้แสดงให้เห็นว่า วิธีการเลือกคุณลักษณะที่ดีที่สุด คือ Fisher score algorithm ที่สามารถเลือกเซตของคุณลักษณะได้อย่างเหมาะสม และอัลกอริทึมที่มีประสิทธิภาพสำหรับการจำแนก ระยะเวลาการนอนได้ดีที่สุด คือ อัลกอริทึมการจำแนกข้อมูลด้วย Random forest ตัวอย่างงานวิจัยที่ได้นำ สัญญาณ EEG เพียงอย่างเดียวมาใช้ในการประเมินภาวะหยุดหายใจขณะนอนหลับ เช่น งานวิจัยของ [44] ได้สกัดคุณลักษณะของสัญญาณ EEG ด้วยเทคนิคการแปลงเวฟเล็ตและใช้อัลกอริทึมการจำแนกข้อมูล ด้วย k-NN สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับของผู้ป่วยที่มีความผิดปกติในระดับความรุนแรง ต่างๆ ซึ่งผลการทดลองได้แสดงค่าความถูกต้องแม่นยำเท่ากับร้อยละ 90.87 และงานวิจัยของ V. Vimala และคณะ [45] ได้สกัดคุณลักษณะ ได้แก่ ค่าพลังงาน ค่าเอ็นโทรปี และค่าความแปรปรวนของย่านความถี่ ต่างๆของสัญญาณ EEG และได้เปรียบเทียบอัลกอริทึมการจำแนกข้อมูล ได้แก่ Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), และ Artificial Neural Network (ANN) ซึ่งผลการ ทดลองของงานวิจัยดังกล่าวได้แสดงให้เห็นว่าอัลกอริทึม SVM สามารถจำแนกภาวะหยุดหายใจขณะนอน หลับ (Apnea) ภาวะหายใจแผ่ว (Hypopnea) และภาวะปกติ (Normal) ได้อย่างมีประสิทธิภาพ

สำหรับงานวิจัยที่ได้นำสัญญาณ EOG มาใช้ในวิเคราะห์การนอนแบบอัตโนมัติ ได้แก่ งานวิจัย ของ [46] นำเสนอการตรวจจับระยะเวลาการนอน ช่วง wake, REM, NREM1-4 ด้วยสัญญาณ EOG โดยได้ กล่าวไว้ว่า ระยะเวลา REM ลูกตาจะมีการเคลื่อนไหวอย่างรวดเร็วและมีการกระตุก ทำให้สัญญาณ EOG ของ ช่วง REM มีความหนาแน่นที่แสดงให้เห็นอย่างเห็นชัดเจน งานวิจัยของ [47] ได้นำสัญญาณ EOG จาก การติดอิเล็กโทรดตำแหน่งหางตาด้านซ้าย ROG และด้านขวา LOG ได้สกัดคุณลักษณะด้วยเทคนิคโดเมน เวลาและโดเมนความถี่จำนวนทั้งสิ้น 21 คุณลักษณะ จากนั้นได้ใช้การเลือกคุณลักษณะโดยขั้นลำดับหรือ เทคนิค SBS เพื่อเลือกเซตของคุณลักษณะที่ดีที่สุด และเข้าสู่การจำแนกระยะเวลาการนอนด้วยอัลกอริทึม ANN ผลการทดลองของอัลกอริทึมได้แสดงค่าความแม่นยำของการจำแนกเท่ากับ ร้อยละ 89.62 และ งานวิจัยของ Md. M. Rahman และคณะ [48] ได้นำเสนอการใช้สัญญาณ EOG เพียงสัญญาณเดียว สำหรับการจำแนกระยะเวลาการนอนต่างๆ ซึ่งได้สกัดคุณลักษณะด้วยโดเมนเวลาและโดเมนเวลาความถี่โดยใช้ Discrete Wavelet Transform นอกจากนี้ได้นำเสนอเทคนิค Neighborhood Component Analysis เพื่อช่วยลดความซับซ้อนของการประมวลผล และนำอัลกอริทึม Under Sampling Boosting เพื่อจำแนก ระยะเวลาการนอนต่างๆ ซึ่งให้ผลการทดลองที่มีค่าประสิทธิภาพการจำแนกสูงสุด อย่างไรก็ตามงานวิจัย ดังกล่าวนี้นี้ได้เสนอแนะว่า การใช้สัญญาณ EOG เพียงอย่างเดียวยังไม่สามารถจำแนกช่วงระยะเวลาการนอน NREM1 ได้อย่างชัดเจน สำหรับงานวิจัยที่นำสัญญาณ EOG เพียงอย่างเดียวมาใช้เพื่อจำแนกภาวะหยุด หายใจขณะนอนหลับมีจำนวนน้อยมาก เนื่องจากสัญญาณ EOG ไม่สามารถแสดงความแตกต่างระหว่าง

ภาวะหยุดหายใจขณะนอนหลับและภาวะหลับที่เป็นปกติได้อย่างชัดเจน ดังนั้นสัญญาณ EOG จึงนิยมนำมาใช้เพื่อการวิเคราะห์การนอนเพียงการจำแนกระยะการนอนเท่านั้น

จากงานวิจัย Togo และ Yamamoto [49] ได้แสดงให้เห็นว่า กลไกของการทำงานของสมองมีความเกี่ยวข้องกับระบบประสาทอัตโนมัติซึ่งมีความสัมพันธ์โดยตรงกับพารามิเตอร์ที่ได้จากการวิเคราะห์ HRV ที่มาจากสัญญาณ ECG โดยแสดงให้เห็นถึงความสัมพันธ์ระหว่างคุณลักษณะทางความถี่ของ HRV หรือค่าผลรวมของพลังงานสเปกตรัมกับระยะเวลาการนอนต่างๆ โดยผลการเปรียบเทียบแสดงให้เห็นว่า ช่วงระยะเวลาการนอน NREM2 และ NREM3 มีค่าผลรวมของพลังงานสเปกตรัมสูงกว่าระยะ REM และระยะ Wake ในขณะเดียวกันค่าขององค์ประกอบของช่วงความถี่ VLF (สะท้อนถึงการทำงานของร่างกายที่ช้ามาก) และค่าขององค์ประกอบช่วงความถี่ LF (สะท้อนถึงการทำงานของระบบประสาทซิมพาเทติก) มีค่าสูงในระยะ NREM2 และ NREM3 จากความสัมพันธ์เหล่านี้งานวิจัยดังกล่าวได้นำพารามิเตอร์ HRV มาใช้ในการจำแนกระยะการนอนได้อย่างมีประสิทธิภาพ นอกจากนี้งานวิจัยที่เกี่ยวกับการนำสัญญาณ ECG เพียงสัญญาณเดียวมาใช้ในจำแนกตรวจจับภาวะหยุดหายใจขณะนอนหลับได้รับความนิยมอย่างมาก เนื่องจากเป็นสัญญาณที่บ่งบอกความผิดปกติเกี่ยวกับการทำงานของระบบหัวใจและหลอดเลือดอย่างชัดเจนและเป็นสัญญาณที่ใช้งานง่าย ราคาไม่สูงมาก จากผลการทดลองของงานวิจัย [31] แสดงให้เห็นว่า อัตราการส่วนระหว่างความถี่ช่วง LF และ HF ของช่วงเวลาที่เกิดภาวะหยุดหายใจขณะนอนหลับมีค่าสูงกว่าช่วงไม่เกิดภาวะหยุดหายใจ และในงานวิจัยของ [33] ได้นำพารามิเตอร์ HRV ที่ถูกคำนวณด้วยเทคนิคเชิงเส้นและไม่เชิงเส้นนำมาใช้สำหรับการจำแนกช่วงการเกิดภาวะหยุดหายใจขณะนอนหลับ (Apnea event) ในงานวิจัยของ Gates และคณะ [50] ได้พบว่าการเปลี่ยนแปลงของระบบประสาทอัตโนมัติและการเปลี่ยนแปลงของ HRV มีความสัมพันธ์กับโรคการนอนกรนและโรคที่เกี่ยวข้องกับความผิดปกติของการนอนอื่นๆ ด้วย นอกจากนี้งานวิจัยของเรา [32] ได้นำเสนอวิธีการจำแนกภาวะหยุดหายใจขณะนอนหลับด้วยสัญญาณ ECG เพียงอย่างเดียว โดยการใช้การหลอมรวมคุณลักษณะด้วยวิธีการวิเคราะห์ความสัมพันธ์คานอนิคอล หรือ Canonical Correlation Analysis ซึ่งคุณลักษณะเหล่านี้ได้มาจากการสกัดคุณลักษณะด้วยวิเคราะห์ HRV และการวิเคราะห์สัญญาณ EDR โดยผลการทดลองแสดงให้เห็นว่าวิธีการดังกล่าวสามารถจำแนกภาวะหยุดหายใจได้อย่างมีประสิทธิภาพด้วยการแสดงค่าความถูกต้องแม่นยำมีค่าเท่ากับร้อยละ 90.42 ดังนั้นจะเห็นได้ว่าการวิเคราะห์สัญญาณ ECG จึงมีความสำคัญต่อการวิเคราะห์การนอน

2.4.2 การวิเคราะห์การนอนแบบอัตโนมัติด้วยสัญญาณหลายชุด

ถึงแม้ว่าการวิเคราะห์การนอนด้วยชุดสัญญาณเดียวมีความสะดวกในการใช้งาน แต่ประสิทธิภาพด้านความแม่นยำยังไม่สูงมาก การแก้ไขปัญหานี้จึงได้นำสัญญาณมากกว่าหนึ่งสัญญาณมาใช้วิเคราะห์ร่วมกันตามมาตรฐานการวิเคราะห์การนอน AASM จากแนวคิดนี้จึงมีหลายงานวิจัยแสดงดังตารางที่ 2.4

ได้นำเสนออัลกอริทึมสำหรับวิเคราะห์การนอนแบบอัตโนมัติจากการพิจารณาสัญญาณหลายชุด จากตารางที่ 2.4 ในปี 2013 งานวิจัยของ Seral Ozsen [51] นำเสนอการจำแนกกระยะการนอนด้วยสัญญาณหลายสัญญาณและได้ใช้เทคนิคการเลือกคุณลักษณะโดยลำดับ (SFS) และจำแนกกระยะการนอนต่างๆ ด้วยอัลกอริทึม ANN ซึ่งผลการทดลองได้แสดงค่าความแม่นยำของการจำแนกเท่ากับร้อยละ 90.93 นอกจากนี้ในงานวิจัยดังกล่าวนี้ให้ข้อเสนอแนะว่า ระยะ NREM2 และ NREM3 คุณลักษณะมีความใกล้เคียงกันมากจึงทำให้ยากต่อการจำแนก และข้อเสียอีกประเด็นของงานวิจัยดังกล่าวนี้ คือ จำนวนตัวอย่างของชุดข้อมูลมีเพียงแค่ 5 ตัวอย่างเท่านั้นดังนั้นผลการวิจัยจึงไม่ครอบคลุม ในปีเดียวกันนั้นงานวิจัยของ Sirvan และคณะ [52] ได้นำเสนอการจำแนกกระยะการนอนด้วยหลายชุดสัญญาณและได้ใช้เทคนิคการลดมิติแบบผสมผสานระหว่าง การเลือกคุณลักษณะด้วยวิธีฟิลเตอร์ในขั้นตอนแรกซึ่งคุณลักษณะจะถูกกรองด้วย ค่าฮิสโตแกรมที่คำนวณด้วยเทคนิค f-test จากความสัมพันธ์ระหว่างคุณลักษณะกับข้อมูลเอาต์พุต ขั้นตอนที่สองได้ใช้วิธีการเลือกคุณลักษณะด้วยเทคนิคของ Minimal Redundancy and Maximum-Relevance(mRMR) หรือการเลือกคุณลักษณะที่พิจารณาจากลำดับความสำคัญ ซึ่งค่าความสำคัญนั้นคำนวณจากผลต่างของค่าความสัมพันธ์ระหว่างคุณลักษณะกับข้อมูลกลุ่ม (Label) และค่าความซ้ำซ้อนของคุณลักษณะ โดยคุณลักษณะที่ถูกเลือกจะมีค่าความสัมพันธ์ระหว่างคุณลักษณะกับข้อมูลกลุ่มสูง และค่าความซ้ำซ้อนของคุณลักษณะต่ำ รายละเอียดของวิธี mRMR สามารถสืบค้นเพิ่มเติมได้จากงานวิจัย [52] ขั้นตอนสุดท้ายได้ใช้อัลกอริทึม SVM ในการจำแนกกระยะการนอน โดยค่าความแม่นยำ (Accuracy) ค่าความไว (Sensitivity) และค่าจำเพาะ (Specificity) ของการจำแนกกระยะการนอนหลับและการตื่น (Sleep-Wake classification) เท่ากับร้อยละ 88.87, 93.74 และ 82.49 ตามลำดับ สำหรับค่าความแม่นยำของการจำแนกกระยะการนอนต่างๆ (Multiclass sleep stage classification) มีค่าความแม่นยำเฉลี่ยเท่ากับร้อยละ 81.77 จากตารางที่ 2.4 จะเห็นได้ว่า จุดด้อยของงานวิจัยนี้คือ ค่าความถูกต้องแม่นยำของการจำแนกกระยะการนอน NREM1 และ NREM2 ที่มีค่าค่อนข้างต่ำ เนื่องจากเป็นช่วงการเปลี่ยนระหว่างระยะการนอนภายใน NREM ด้วยกันเอง ดังนั้นจึงทำให้สัญญาณช่วงการเปลี่ยนเข้าสู่ NREM1 คล้ายกับระยะ NREM2 และ ช่วง NREM2 จะคล้ายกับช่วง NREM3 จึงทำให้การจำแนกกระยะการนอน NREM1 และ NREM2 เกิดความผิดพลาดบ่อย ต่อมาในปี 2018 งานวิจัยของ [53] Gharbali และคณะ ได้พิจารณาหลายชุดสัญญาณ (EEG, EMG, EOG และ ECG) เพื่อจำแนกกระยะการนอน และได้สกัดคุณลักษณะทั้งหมด 32 คุณลักษณะด้วยเทคนิคโดเมนเวลาและโดเมนความถี่ ในขั้นตอนการเลือกคุณลักษณะได้ทำการทดลองเปรียบเทียบวิธีการเลือกคุณลักษณะด้วยวิธีฟิลเตอร์ทั้งหมด 6 วิธี ประกอบด้วย Relief Feature, mRMR, Fisher score Chi-square test และ Information Gain รายละเอียดของวิธีการดังกล่าวสามารถศึกษาได้จากงานวิจัย [53] และขั้นตอนการจำแนกกระยะการนอนได้นำอัลกอริทึม kNN, ANN และ Decision-tree-based multi-SVM (DisSVM) มาใช้เปรียบเทียบ ซึ่งผล

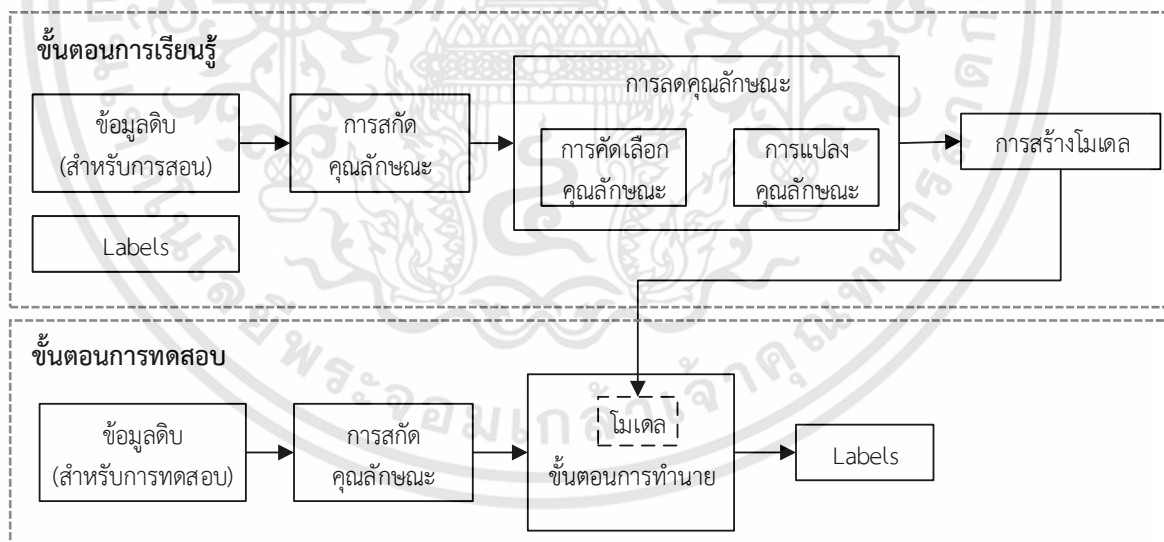
การทดลองของงานวิจัยดังกล่าวได้สรุปว่า วิธี mRMR-DisSVM สามารถจำแนกระยะการนอนได้ดีที่สุด โดยแสดงค่าความถูกต้องแม่นยำของการจำแนกเท่ากับร้อยละ 85.5 และข้อเสนอแนะเพิ่มเติมจากงานวิจัยดังกล่าวนี้คือ คุณลักษณะที่ได้จากสัญญาณ ECG มีผลต่อการเพิ่มประสิทธิภาพการจำแนกระยะการนอน เพราะระยะการนอนช่วง NREM3 มีการทำงานของหัวใจที่ช้าลงอย่างเห็นได้ชัด เพราะฉะนั้นจึงสามารถช่วยแสดงความแตกต่างระหว่าง NREM2 และ NREM3 โดยข้อเสนอแนะดังกล่าวนี้ในงานวิจัยของ [54] ได้สนับสนุนเช่นกัน และงานวิจัยล่าสุดในปี 2018 [55] Rui Yan และคณะ ได้พัฒนาการระบุระยะการนอนแบบอัตโนมัติด้วยการหลอมรวมสัญญาณหลายสัญญาณ และสกัดคุณลักษณะรวมทั้งหมด 232 คุณลักษณะ ขั้นตอนการเลือกคุณลักษณะนั้นประกอบด้วย Relief Feature, Improved distance-based evolution method (IDE), Genetic Algorithm, และ SFS ขั้นตอนการจำแนกข้อมูลได้นำเสนอวิธี Random Forest เป็นวิธีที่แสดงค่าความแม่นยำที่ดีที่สุดเท่ากับร้อยละ 87.24 นอกจากนี้ผลการทดลองของงานวิจัยดังกล่าวได้แสดงให้เห็นว่า การเลือกคุณลักษณะด้วยวิธี Relief feature เป็นวิธีที่เหมาะสมซึ่งเลือกคุณลักษณะด้วยการลำดับความสำคัญ การให้คะแนนความสำคัญขึ้นอยู่กับความสามารถในจำแนกคุณลักษณะนั้นๆ กับคุณลักษณะบริเวณใกล้เคียง หากคุณลักษณะมีระยะห่างจากคุณลักษณะบริเวณใกล้เคียงมากๆ ค่าคะแนนความสำคัญนั้นจะมีค่าสูง จากวิธีการดังกล่าวเราจะเห็นได้ว่าข้อดีของการพิจารณาความสัมพันธ์ระหว่างคุณลักษณะรอบข้างมีผลต่อการเพิ่มประสิทธิภาพของการวิเคราะห์การนอน จากงานวิจัยที่ผ่านมาทั้งหมดที่เกี่ยวข้องกับการวิเคราะห์นอนด้วยหลายชุดสัญญาณได้แสดงผลค่าความถูกต้องแม่นยำในการจำแนกอยู่ในช่วงร้อยละ 85-90 ซึ่งถือว่ายังไม่สูงมาก ดังนั้นงานวิจัยหัวข้อนี้ยังมีความต้องการการพัฒนาอย่างต่อเนื่อง

ตารางที่ 2.4 สรุปงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์การนอนแบบอัตโนมัติด้วยสัญญาณหลายชุด

ผู้แต่ง (ปี)	สัญญาณที่นำมาใช้ในการวิเคราะห์	การสกัดคุณลักษณะและวิธีการเลือกคุณลักษณะ	วิธีการเลือกคุณลักษณะ	อัลกอริทึมจำแนกข้อมูล	ผลการทดลอง
Seral Ozsen [50] (2013)	EEG, EOG, และ EMG	Time, Frequency domain และ Time-frequency domains	เทคนิคการเลือกคุณลักษณะโดยลำดับ (SFS)	ANN	Acc.= 90.93%
Sirvan และคณะ [51] (2013)	EEG 6 channels, EOG 2 channels และ EMG 1 channels	Time, Frequency domain และ Time-frequency domains (Wavelet)	Minimal Redundancy and Maximum-Relevance (mRMR)	SVM	Sleep-wake Classification Acc.= 88.87% Sen.=93.74% Spec. 82.49% Multi-Sleep stage Acc.= 81.77 % (Wake=88.59%, REM=86.89%,NREM1=66.91%, NREM2=79.06%,NREM3=87.13%)
Gharbali และคณะ [52] (2018)	EEG, EOG, EMG และ ECG	32 คุณลักษณะด้วยเทคนิคโดเมนเวลาและโดเมนความถี่	Minimal Redundancy and Maximum-Relevance (mRMR)	Decision-tree-based multi-SVM	Acc.= 85.5%
Rui Yan และคณะ [54] (2018)	EEG, EOG, EMG และ ECG	232 คุณลักษณะด้วยเทคนิคโดเมนเวลา โดเมนความถี่ โดเมนเวลา-ความถี่ และเทคนิคไม่เชิงเส้น	Relief feature	Random Forest	Acc.=87.24%

2.5 การออกแบบระบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง (Machine learning)

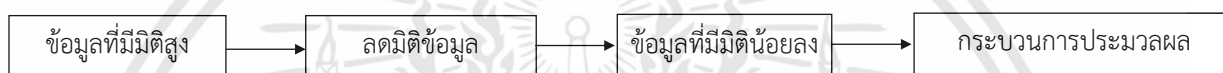
โดยทั่วไประบบการเรียนรู้ของเครื่องแบ่งออกเป็น 3 แบบ คือ การเรียนรู้แบบมีผู้สอน (Supervised machine learning) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised machine learning) การเรียนรู้แบบเสริมกำลัง (Reinforcement learning) ระบบอัตโนมัติของงานวิจัยนี้ได้นำแนวคิดการออกแบบด้วยวิธีการเรียนรู้ของเครื่องโดยมีผู้สอน แผนภาพของระบบการเรียนรู้ของเครื่องโดยมีผู้สอน แสดงดังรูปที่ 2.16 จากรูปแสดงระบบการเรียนรู้ของเครื่องโดยมีผู้สอน เริ่มต้นด้วยการเตรียมข้อมูล ข้อมูลทั้งหมด (Raw data) และข้อมูลที่มีการระบุกลุ่มของข้อมูล (Label) ถูกสุ่มแบ่งออกเป็น 2 ชุดข้อมูล ดังนี้ ชุดข้อมูลสำหรับการเรียนรู้ (Training data) และชุดข้อมูลสำหรับการทดสอบ (Testing data) ชุดข้อมูลสำหรับการเรียนรู้ใช้เพื่อสร้างโมเดลสำหรับจำแนกกลุ่มข้อมูล กระบวนการสร้างโมเดลเริ่มจากขั้นตอนการสกัดคุณลักษณะหรือขั้นตอนการหาตัวแปรเพื่อเป็นตัวแทนของข้อมูลนั้นๆ และเข้าสู่ขั้นตอนลดมิติ ซึ่งเป็นขั้นตอนที่จำเป็นมากสำหรับระบบที่มีคุณลักษณะหรือตัวแปรจำนวนมาก ความหมายของการลดมิติข้อมูล คือ การเลือกเฉพาะคุณลักษณะที่สำคัญ (Selection) หรือการแปลงคุณลักษณะ (Transformation) ให้มีมิติลดลง เมื่อข้อมูลถูกลดมิติแล้วจะเข้าสู่กระบวนการประมวลผลสำหรับการจำแนกกลุ่มข้อมูลเพื่อให้ได้ผลลัพธ์ที่ดีที่สุด ผลการทดสอบประสิทธิภาพของโมเดลได้มาจากการทดสอบโมเดลนั้นๆ กับชุดข้อมูลสำหรับการทดสอบ



รูปที่ 2.16 แผนภาพของระบบการเรียนรู้ของเครื่องโดยมีผู้สอน

2.6 การลดมิติข้อมูล (Dimensionality reduction)

การลดมิติข้อมูลเป็นวิธีที่ถูกรำเสนอเพื่อแก้ไขปัญหาของมิติข้อมูล Curse of dimension ซึ่งได้รับผลกระทบจากข้อมูลที่มีมิติสูง (High dimension) โดยปราศจากการจัดการข้อมูลเบื้องต้น ข้อมูลที่มีจำนวนมากเหล่านั้น มีลักษณะข้อมูลที่กระจัดกระจายและบางข้อมูลไม่ได้นำมาใช้ประโยชน์ในการวิเคราะห์ อาจส่งผลกระทบต่อความถูกต้องของการประมวลผล เนื่องจากอัลกอริทึมการจำแนกข้อมูล บางอัลกอริทึมไม่สามารถรองรับการทำงานของตัวแปรจำนวนมากได้ นอกจากนี้ยังทำให้เกิดการสิ้นเปลืองทรัพยากรในการประมวลผลอีกด้วย ได้แก่ เวลา หรือ หน่วยความจำ ดังนั้นการออกแบบระบบอัตโนมัติที่มีประสิทธิภาพมีความจำเป็นอย่างยิ่งที่ต้องมีการจัดการข้อมูลเบื้องต้นด้วยวิธีการลดมิติ แผนภาพของการลดมิติข้อมูลสามารถแสดงดังรูปที่ 2.17



รูปที่ 2.17 แผนภาพของการลดมิติข้อมูล

เทคนิคการลดมิติที่ใช้ในงานวิจัยที่ผ่านมาเราสามารถแบ่งออกเป็น 2 กลุ่ม ได้แก่ เทคนิคการลดมิติแบบเดี่ยว (Single dimensionality reduction) และเทคนิคการลดมิติแบบผสมผสาน (Hybrid dimensionality reduction) เทคนิคการลดมิติแบบเดี่ยวคือ การแปลงคุณลักษณะ (Feature Transformation) หรือ การเลือกคุณลักษณะที่สำคัญ (Feature Selection) เพียงอย่างใดอย่างหนึ่ง และเทคนิคการลดมิติแบบผสมผสานคือ การรวมกันของวิธีการลดมิติหลายวิธี โดยรายละเอียดของวิธีการลดมิติข้อมูลสามารถอธิบายได้ดังนี้

2.6.1 การแปลงคุณลักษณะ (Feature transformation)

การแปลงคุณลักษณะที่ใช้สำหรับการลดมิติข้อมูลเป็น การแปลงคุณลักษณะจากเบสิสเดิมไปยังเบสิสใหม่ซึ่งมีผลทำให้ขนาดของจำนวนข้อมูลลดลง การแปลงคุณลักษณะของเบสิสเดิม X ไปยังเบสิสใหม่สามารถดำเนินการได้ตามสมการ

$$Y = wX \quad (2.27)$$

โดยที่ w คือ ค่าน้ำหนักหรือสัมประสิทธิ์ และฟังก์ชัน Y คือ ข้อมูล X ที่ถูกลดมิติให้ที่มีขนาดเล็กลง วิธีการแปลงคุณลักษณะของชุดข้อมูลที่มีมากกว่า 2 ชุดข้อมูลขึ้นไป ถูกเรียกอีกอย่างว่า วิธีหลอมรวมคุณลักษณะ (Feature fusion) การวิธีหลอมรวมคุณลักษณะมีหลายวิธีดังนี้

2.6.1.1 การวิเคราะห์หาองค์ประกอบหลัก (Principle Component Analysis: PCA)

วิธี PCA เป็นวิธีทางสถิติที่วิเคราะห์หาองค์ประกอบหลักของข้อมูลภายใต้เงื่อนไขค่าความแปรปรวน (Variance) สูงสุด หรือเป็นกระบวนการแปลงข้อมูล X บนเมทริกซ์การแปลง P เพื่อเป็นข้อมูลเชิงเส้น Y แสดงความสัมพันธ์ดังสมการ

$$Y=PX \quad (2.28)$$

เนื่องจากวิธี PCA เป็นวิธีที่พิจารณาความสัมพันธ์ร่วมกันของข้อมูล X หรือข้อมูลเพียงชุดเดียว ดังนั้นกระบวนการหลอมรวมหลายชุดข้อมูลด้วยวิธี PCA นี้ ข้อมูลแต่ละชุดถูกจัดเรียงแบบอนุกรมดังตาราง 2.5 เพื่อหาความสัมพันธ์ร่วมและและคำนวณหาองค์ประกอบหลัก

2.6.1.2 การวิเคราะห์แยกแยะเชิงเส้น (Linear Discriminant Analysis: LDA)

วิธี LDA เป็นวิธีการทางสถิติที่ใช้การวิเคราะห์จำแนกกลุ่มตั้งแต่ 2 กลุ่มขึ้นไป หรือเรียกอีกอย่างหนึ่งว่า กระบวนการวิเคราะห์ LDA เป็นกระบวนการแยกแยะกลุ่มด้วยเทคนิคการเรียนรู้แบบมีผู้สอน กระบวนการหลอมรวมหลายชุดข้อมูลด้วยวิธี LDA มีการจัดเรียงข้อมูลคล้ายกับวิธี PCA แต่วิธี LDA จะนำข้อมูลของ label มาพิจารณาร่วมด้วยและเงื่อนไขสำหรับการหาเมทริกซ์การแปลงของวิธี LDA มีความแตกต่างกันด้วย เงื่อนไขของวิธี LDA หาจากอัตราส่วนระหว่างค่าความแปรปรวนระหว่างกลุ่ม (Covariance between group, S_b) กับค่าความแปรปรวนร่วมภายในกลุ่ม (Covariance within group, S_w) สูงสุด ดังสมการ

$$S_w = \sum_{i=1}^c \sum_{x \in D_i} (x - m_i)(x - m_i)^T \quad (2.29)$$

$$S_b = \sum_{x \in D_i} (m_i - m)(m_i - m)^T \quad (2.30)$$

โดยที่ c คือ จำนวนกลุ่ม, m_i คือ ค่าเฉลี่ยของข้อมูลแต่ละกลุ่ม และ m คือ ค่าเฉลี่ยของข้อมูลทั้งหมด

2.6.1.3 การวิเคราะห์สหสัมพันธ์คาโนนิคอล (Canonical Correlation Analysis : CCA)

วิธี CCA เป็นวิธีการทางสถิติมีจุดประสงค์เพื่อหาความสัมพันธ์เชิงเส้นของตัวแปรระหว่างแต่ละชุดข้อมูลตั้งแต่ 2 ชุดข้อมูลเป็นต้นไป ซึ่งตัวแปรทั้งหมดอยู่ในระดับเดียวกันหรือตัวอย่างเดียวกันและข้อมูลต้องมีการกระจายแบบปกติ (Gaussian distribution) เรากำหนดให้สองชุดข้อมูลที่มาจกตัวอย่าง

ผู้ป่วยเดียวกันเป็น $\{(X_1, X_2)\}'$ เงื่อนไขของวิธี CCA คือ การหาค่าสัมประสิทธิ์สหสัมพันธ์คาโนนิคอลที่ได้จากค่าความสัมพันธ์ที่เป็นไปได้สูงสุดระหว่างกลุ่มข้อมูลต่างๆ ดังสมการ

$$\rho = \max \frac{w_{X_1}^T X_1 X_2^T w_{X_2}}{\sqrt{w_{X_1}^T X_1 X_1^T w_{X_1} w_{X_2}^T X_2 X_2^T w_{X_2}}} \quad (2.31)$$

โดยที่ w_{X_1} และ w_{X_2} เป็นค่าสัมประสิทธิ์สหสัมพันธ์คาโนนิคอลของข้อมูลชุดที่ 1 และข้อมูลชุดที่ 2 ซึ่งเป็นค่านำหนักที่แสดงความสำคัญของตัวแปรต่างๆ

จากสมการ 2.31 สามารถเขียนเป็นเมทริกซ์เพื่อแก้สมการหาค่าไอเกนได้ดังนี้

$$\begin{bmatrix} 0 & X_1 X_2^T \\ X_2 X_1^T & 0 \end{bmatrix} \begin{bmatrix} w_{X_1} \\ w_{X_2} \end{bmatrix} = \lambda \begin{bmatrix} X_1 X_1^T & 0 \\ 0 & X_2 X_2^T \end{bmatrix} \begin{bmatrix} w_{X_1} \\ w_{X_2} \end{bmatrix} \quad (2.32)$$

λ คือค่าไอเกนที่สอดคล้องกับไอเกนเวกเตอร์ w_{X_1} และ w_{X_2}

2.6.1.4 การวิเคราะห์ความสัมพันธ์สหสัมพันธ์จำแนกกลุ่ม (Discriminant Canonical Correlation Analysis: DCCA)

เนื่องจากวิธี CCA ซึ่งเป็นวิธีการแปลงคุณลักษณะแบบไม่มีผู้สอน และไม่ใช้ประโยชน์จากข้อมูล label หรือข้อมูลกลุ่ม จึงทำให้มีข้อจำกัดในการประยุกต์ใช้งานกับประสิทธิภาพของการรู้จำของกลุ่มข้อมูลมิติสูง ดังนั้นงานวิจัย [10] ได้นำเสนอวิธีการ DCCA ซึ่งได้การนำข้อมูลกลุ่มมาใช้ในโครงสร้างของ CCA เพื่อเพิ่มประสิทธิภาพของระบบรู้จำที่ข้อมูลมาจากหลายแหล่งข้อมูล ในงานวิจัย [10-11] ได้ทำการทดลองเปรียบเทียบวิธี DCCA และวิธี CCA ในระบบการรู้จำต่างๆ ได้แก่ การจัดหมวดหมู่ของคำ ระบบรู้จำใบหน้า ระบบรู้จำการเขียน ซึ่งผลการทดลองแสดงให้เห็นว่ามีวิธีการแปลงคุณลักษณะด้วยวิธี DCCA สามารถเพิ่มประสิทธิภาพการจำแนกกลุ่มข้อมูลได้

แนวคิดของวิธี DCCA ได้มาจากวิธี CCA เรากำหนดสองชุดข้อมูลที่มาจกตัวอย่างผู้ป่วยเดียวกันเป็น $\{(X_1, X_2)\}'$ ซึ่งมาจากกลุ่มข้อมูล C กลุ่ม จากสมการ 2.30 ความสัมพันธ์ร่วมระหว่างชุดข้อมูลทั้งสอง $X_1 X_2^T$ ของฟังก์ชันจุดประสงค์ของ CCA ถูกแก้ไขเป็น $C_w - \eta C_b$ ดังนั้นฟังก์ชันจุดประสงค์ของ DCCA จะได้

$$\rho = \max \frac{w_{X_1}^T (C_w - \eta C_b) w_{X_2}}{\sqrt{w_{X_1}^T X_1 X_1^T w_{X_1} w_{X_2}^T X_2 X_2^T w_{X_2}}} \quad (2.33)$$

$$C_B = -X_1 \mathbf{A} X_2^T \quad (2.41)$$

แทนสมการ 2.34 และสมการ 2.41 ในเทอมของ $C_W - \eta C_B$ จะได้สมการ

$$\begin{aligned} C_W - \eta C_B &= X_1 \mathbf{A} X_2^T + \eta X_1 \mathbf{A} X_2^T \\ &= (1 + \eta) X_1 \mathbf{A} X_2^T \end{aligned} \quad (2.42)$$

ในงานวิจัย [10] ได้พิสูจน์สมการ 2.42 และละทิ้ง η ได้ ดังนั้นเมทริกซ์ของ DCCA สามารถเขียนได้ดังนี้

$$C_W - \eta C_B = X_1 \mathbf{A} X_2^T \quad (2.43)$$

จากสมการ 2.33 สามารถเขียนเป็นเมทริกซ์เพื่อแก้สมการหาค่าไอเกนได้ดังนี้

$$\begin{bmatrix} 0 & X_1 \mathbf{A} X_2^T \\ X_2 \mathbf{A} X_1^T & 0 \end{bmatrix} \begin{bmatrix} w_{X_1} \\ w_{X_2} \end{bmatrix} = \lambda \begin{bmatrix} X_1 X_1^T & 0 \\ 0 & X_2 X_2^T \end{bmatrix} \begin{bmatrix} w_{X_1} \\ w_{X_2} \end{bmatrix} \quad (2.44)$$

λ คือค่าไอเกนที่สอดคล้องกับไอเกนเวกเตอร์ w_{X_1} และ w_{X_2}

จากการศึกษาวิธีการแปลงคุณลักษณะของแต่ละวิธีข้างต้น เราได้สรุปข้อแตกต่างของการหลอมรวมข้อมูลด้วยวิธีการแปลงคุณลักษณะของแต่ละวิธีไว้ดังตารางที่ 2.5 มีรายละเอียดดังนี้ วิธี PCA เป็นวิธีการแปลงคุณลักษณะภายใต้เงื่อนไขความแปรปรวนของข้อมูลมากที่สุดของข้อมูลเพียงข้อมูลเดียว โดยจุดเด่นของวิธี PCA คือ เป็นวิธีที่มีขั้นตอนการดำเนินการง่าย และเป็นวิธีที่ถูกรับไปใช้อย่างทั่วไป ตัวอย่างเช่น ในงานวิจัยของ [56] นำเสนอการลดคุณลักษณะของสัญญาณ ECG เพื่อใช้ในการช่วยวินิจฉัยภาวะหยุดหายใจขณะนอนหลับ และงานวิจัยของ [57] ลดมิติข้อมูลของสัญญาณ EEG ด้วยวิธี PCA เพื่อปรับปรุงประสิทธิภาพของการจำแนกระยะการนอน อย่างไรก็ตามวิธี PCA มีจุดด้อยคือ มีประสิทธิภาพต่ำเมื่อนำมาใช้กับข้อมูลที่มีลักษณะเป็นข้อมูลแบบไม่เชิงเส้น วิธีหลอมรวมมิติข้อมูลด้วยวิธี LDA มีขั้นตอนการจัดเรียงข้อมูลที่เหมือนกับวิธี PCA แต่วิธี LDA ได้นำข้อมูลกลุ่มของตัวแปรมาร่วมพิจารณาด้วยจึงทำให้การหลอมรวมข้อมูลหลายชุดมีประสิทธิภาพดีขึ้นเมื่อเทียบกับวิธี PCA เพราะเมทริกซ์การแปลงของวิธี LDA จะพิจารณาความสัมพันธ์ระหว่างข้อมูลที่อยู่กลุ่มเดียวกัน ขณะเดียวกันก็ได้พิจารณาความสัมพันธ์ของข้อมูลระหว่างกลุ่มด้วย สำหรับวิธี CCA และ DCCA เป็นวิธีการหาเมทริกซ์การแปลงจากความสัมพันธ์ระหว่างสองชุดข้อมูล จากตารางจะเห็นว่าวิธี CCA และ DCCA พิจารณาความสัมพันธ์ของแต่ละชุดข้อมูลต่างจากวิธี PCA และ LDA โดยทั่วไปวิธี CCA และ DCCA เหมาะสำหรับ

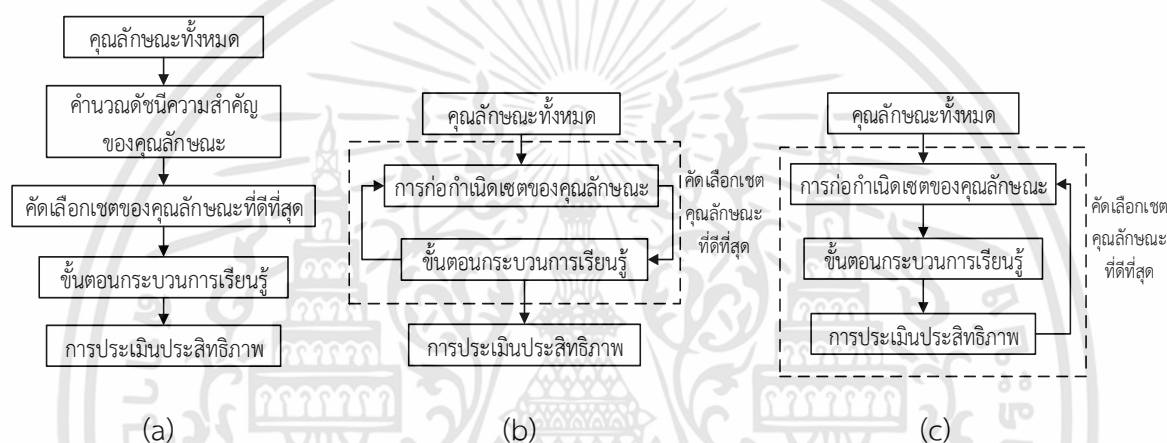
การวิเคราะห์ข้อมูลที่มาจกหลายชุดข้อมูล แต่อย่างไรก็ตามวิธี CCA เป็นการวิเคราะห์โดยไม่ได้คำนึงถึงความเป็นกลุ่มข้อมูล (Label) ดังนั้นในกรณีการนำไปประยุกต์ใช้กับข้อมูลที่มีขนาดใหญ่หรือจำนวนมากทำให้ประสิทธิภาพลดลง ดังนั้นจึงได้มีการนำแนวคิดข้อมูลกลุ่มร่วมพิจารณาด้วย ดังเช่นในงานวิจัยของ [10-11] ได้ทดลองเปรียบเทียบประสิทธิภาพการหลอมรวมมิติของหลายชุดข้อมูลด้วยวิธี CCA และ DCCA ผลการทดลองของวิจัยเหล่านั้นพบว่าวิธี DCCA มีความสามารถที่ช่วยเพิ่มประสิทธิภาพของระบบการเรียนรู้ของเครื่องได้

ตารางที่ 2.5 วิธีหลอมรวมคุณลักษณะด้วยวิธี PCA, LDA, CCA และ DCCA

วิธี	การแปลงคุณลักษณะ	ภาพคำอธิบาย
PCA	เมทริกซ์การแปลงคุณลักษณะหาจากความแปรปรวนร่วมที่มากที่สุด	<p>Common profiles</p> <p>ข้อมูลชุดที่ 1 ข้อมูลชุดที่ 2 = A PC</p> <p>Principle components</p>
LDA	เมทริกซ์การแปลงคุณลักษณะได้จากการพิจารณาจาก ข้อมูล Labeled Class และหาความสัมพันธ์ระหว่างกลุ่มและความสัมพันธ์ภายในกลุ่ม	<p>Common profiles</p> <p>Labeled class ข้อมูลชุดที่ 1 ข้อมูลชุดที่ 2 = A $S_W^{-1} S_B$</p> <p>Principle components</p>
CCA	เมทริกซ์การแปลงคุณลักษณะได้จากการหาความสัมพันธ์ของระหว่างสองชุดข้อมูล	<p>Mixing profiles</p> <p>ข้อมูลชุดที่ 1 = A1 C1</p> <p>ข้อมูลชุดที่ 2 = A2 C2</p> <p>Canonical variants</p> <p>Associated components</p>
DCCA	เมทริกซ์การแปลงคุณลักษณะได้จากการพิจารณาจาก ข้อมูล Labeled Class และความสัมพันธ์ของระหว่างสองชุดข้อมูล	<p>Mixing profiles</p> <p>Labeled class ข้อมูลชุดที่ 1 = A1 C1</p> <p>Labeled class ข้อมูลชุดที่ 2 = A2 C2</p> <p>Canonical variants</p> <p>Associated components</p>

2.6.2 วิธีการเลือกคุณลักษณะ (Feature selection)

วิธีการเลือกคุณลักษณะเป็นวิธีการเลือกคุณลักษณะหรือเลือกเซตย่อยของคุณลักษณะที่มีความสำคัญต่อการวิเคราะห์ข้อมูล วิธีการเลือกคุณลักษณะถูกแบ่งออกได้เป็น 3 วิธี ได้แก่ วิธีฟิลเตอร์ (Filter method) วิธีแรปเปอร์ (Wrapper method) และวิธีฝังตัว (Embed method)



รูปที่ 2.18 กระบวนการเลือกตัวแปรด้วย (a) วิธีฟิลเตอร์ (b) วิธีแรปเปอร์ และ (c) วิธีฝังตัว

รูปที่ 2.18 แสดงกระบวนการเลือกตัวแปรของทั้งสามวิธีมีรายละเอียดดังนี้ วิธีฟิลเตอร์เป็นวิธีที่ใช้เทคนิคการจัดเรียงลำดับของคุณลักษณะ (Feature Ranking technique) เพื่อคัดเลือกคุณลักษณะ จากรูปที่ 2.18(a) กระบวนการเลือกตัวแปรด้วยวิธีฟิลเตอร์เริ่มจาก คุณลักษณะทั้งหมดจะถูกวัดค่าความสำคัญ จากนั้นกำหนดระดับอ้างอิง (Threshold) ขึ้นมาเพื่อกำจัดคุณลักษณะที่มีค่าดัชนีของความสำคัญน้อยกว่าระดับที่อ้างอิงออกไป และเข้าสู่กระบวนการประมวลผล ข้อดีของวิธีเลือกคุณลักษณะแบบฟิลเตอร์ คือ เป็นเทคนิคที่คำนวณได้ง่าย รวดเร็ว และหลีกเลี่ยงการเกิดโอเวอร์ฟิตติ้ง (Overfitting) เพราะวิธีฟิลเตอร์ไม่นำผลทดสอบประสิทธิภาพการเรียนรู้ของเครื่องมาพิจารณาพร้อมด้วยซึ่งคุณลักษณะที่ถูกคัดเลือกจะไม่ถูกไบแอส ข้อเสียของวิธีนี้ คือ คุณลักษณะที่ถูกคัดเลือกเป็นคุณลักษณะที่เป็นอิสระต่อกัน เพราะขั้นตอนการคำนวณค่าความสำคัญจะพิจารณาความสัมพันธ์เพียงด้านเดียว คือ ความสัมพันธ์ระหว่างคุณลักษณะนั้นๆกับเอาต์พุต ไม่ได้คำนึงถึงความสัมพันธ์ระหว่างคุณลักษณะกันเอง ดังนั้นเซตของคุณลักษณะที่ถูกเลือกอาจมีความสัมพันธ์ระหว่างกันหรือไม่มีความสัมพันธ์ระหว่างกันก็ได้ เมื่อนำเซตคุณลักษณะที่ถูกเลือกเหล่านี้มาใช้ในการจำแนกกลุ่มข้อมูล จึงส่งผลทำให้ค่าความถูกต้องของ

การเรียนรู้ลดลง ดังนั้นการเลือกคุณลักษณะแบบฟิลเตอร์จึงเหมาะสมเฉพาะการวิเคราะห์ข้อมูลที่มีจำนวนมิติไม่สูงมาก

วิธีเลือกคุณลักษณะแบบแรปเปอร์เป็นวิธีที่ถูกพัฒนามาเพื่อแก้ไขวิธีฟิลเตอร์ ซึ่งวิธีแรปเปอร์มีกระบวนการทำงานที่เพิ่มขึ้นคือ คุณลักษณะทั้งหมดจะถูกจัดให้อยู่ในรูปของเซตคุณลักษณะ หลังจากนั้นดำเนินการค้นหาเซตของคุณลักษณะที่เหมาะสมด้วยการประเมินด้วยฟังก์ชันความเหมาะสม (Fitness function) เพื่อค้นหาเซตคุณลักษณะที่เหมาะสมที่สุดก่อนเข้าสู่กระบวนการจำแนกข้อมูล วิธีค้นหาเซตคุณลักษณะที่เหมาะสมประกอบด้วย 2 วิธี ได้แก่ 1) วิธีค้นหาเลือกคุณลักษณะโดยลำดับ (Sequential selection algorithm) เป็นวิธีการที่เลือกคุณลักษณะด้วยการพิจารณาคุณลักษณะที่เพิ่มทีละตัวตามลำดับ หรือลดคุณลักษณะลงทีละตัวตามลำดับจนกว่าจะได้เซตคุณลักษณะที่เหมาะสม วิธีการนี้เป็นวิธีการที่มีกระบวนการดำเนินงานง่ายแต่มีข้อเสียคือ ใช้เวลาในการประมวลผลนาน 2) วิธีการค้นหาเซตของคุณลักษณะด้วยวิธีการสุ่มเลือกหรือ วิธีค้นหาคำตอบที่เหมาะสมด้วยอัลกอริทึมของการแก้ไขปัญหาคิววิริสติก (Heuristic search algorithm) ได้แก่ Genetic algorithm (GA), Particle Swarm Optimization (PSO) เป็นต้น รายละเอียดของวิธีนี้ถูกนำเสนออย่างละเอียดในหัวข้อ 2.8 กระบวนการค้นหาเซตของคุณลักษณะจะเริ่มต้นด้วย กำหนดจำนวนของเซตของคุณลักษณะและจำนวนของคุณลักษณะที่ถูกเลือก แต่ละเซตของคุณลักษณะจะถูกประเมินด้วยฟังก์ชันความเหมาะสม และค่าฟังก์ชันความเหมาะสมของแต่ละเซตคุณลักษณะถูกนำมาเปรียบเทียบเพื่อค้นหาเซตของคุณลักษณะที่เหมาะสมที่สุด จากนั้นเซตของคุณลักษณะที่เหมาะสมที่สุดจะถูกนำไปใช้กระบวนการจำแนกข้อมูลต่อไป ข้อดีของวิธีการค้นหาแบบคิววิริสติกเมื่อเทียบกับวิธี SSA เป็นวิธีที่ใช้เวลาน้อยกว่า เนื่องจากวิธีการค้นหาคิววิริสติกเป็นการค้นหาข้อมูลแบบครอบคลุม (Global Search)

วิธีเลือกคุณลักษณะวิธีการฝังตัวเป็นวิธีที่ถูกนำเสนอเพื่อแก้ไขข้อเสียของวิธีฟิลเตอร์และวิธีแรปเปอร์ ใช้เวลาน้อยกว่าวิธีแรปเปอร์ กระบวนการเลือกคุณลักษณะของวิธีนี้ได้รวมการเลือกคุณลักษณะไว้เป็นส่วนหนึ่งของกระบวนการเรียนรู้ ซึ่งมีข้อดีคือ มีการค้นหาเซตของคุณลักษณะทั้งปริภูมิครอบคลุม (Global space) และปริภูมิเฉพาะที่ (Local space) จึงทำให้มีประสิทธิภาพในการค้นหามีประสิทธิภาพที่ดีขึ้น แต่อย่างไรก็ตามวิธีฝังตัวมีข้อเสีย คือ การเลือกเซตคุณลักษณะไม่มีความยืดหยุ่นเนื่องจากขึ้นอยู่กับอัลกอริทึมการจำแนกกลุ่มข้อมูล

จากหลักการวิธีลดมิติทั้งสองแบบ ประกอบด้วย วิธีแปลงคุณลักษณะและวิธีการเลือกคุณลักษณะ เราสามารถสรุปจุดเด่นและจุดด้อยของแต่ละวิธีได้ดังตารางที่ 2.6

ตารางที่ 2.6 จุดเด่นและจุดด้อยของวิธีลดมิติด้วยวิธีเลือกคุณลักษณะและวิธีการแปลงคุณลักษณะ

วิธีการ	กระบวนการ	จุดเด่น	จุดด้อย
---------	-----------	---------	---------

วิธีเลือก คุณลักษณะ	เลือกคุณลักษณะที่มี อิทธิพลต่อ Output	- คุณลักษณะที่ไม่มี ความสำคัญจะถูกตัดทิ้ง - เป็นวิธีที่ง่าย - ข้อมูลไม่สูญหาย	- คุณลักษณะที่ถูกเลือกเป็น อิสระต่อกัน ไม่มีความสัมพันธ์ ระหว่างกัน และไม่ครอบคลุม โครงสร้างคุณลักษณะเดิม
วิธีการแปลง คุณลักษณะ	ชุดข้อมูลทั้งหมดถูก แปลงไปยังเบสใหม่ สร้างจากความสัมพันธ์ ของคุณลักษณะ	- ครอบคลุมโครงสร้างของ คุณลักษณะเดิม - แต่ละคุณลักษณะมี ความสัมพันธ์ระหว่างกัน	- ประสิทธิภาพต่ำเมื่อใช้กับ ข้อมูลมิติสูง - มีข้อมูลสูญหายระหว่างการ แปลงคุณลักษณะ

2.7 การลดมิติแบบผสมผสาน (Hybrid Dimensionality reduction)

จากหัวข้อที่ 2.6 จะเห็นได้ว่าวิธีการลดมิติแบบเดียวจะพิจารณามุมมองของคุณลักษณะเพียงด้านเดียว ดังนั้นการพิจารณาสองมุมมองของคุณลักษณะจึงถูกนำเสนอในวิธีการลดมิติแบบผสมผสาน การลดมิติแบบผสมผสานเป็นวิธีที่ผสมผสานระหว่างวิธีการลดมิติแบบเดียวหลายวิธี โดยได้นำข้อดีของวิธีหนึ่งเพื่อชดเชยแก้ไขข้อเสียของอีกวิธีหนึ่ง หรือนำข้อดีของแต่ละวิธีมาผสมผสานกันเพื่อสามารถเพิ่มประสิทธิภาพได้ดีมากขึ้น จากงานวิจัยที่ผ่านมาที่นำเสนอการออกแบบวิธีการลดมิติแบบผสมผสานมีหลากหลายวิธี ซึ่งเราได้แบ่งออกเป็น 4 กลุ่มตามลำดับขั้นตอนของกระบวนการผสมผสานดังรูปที่ 2.19



รูปที่ 2.19 วิธีการลดมิติแบบผสมผสานทั้ง 4 แบบ

จากรูป 2.19 แสดงวิธีการลดมิติแบบผสมผสานทั้ง 4 แบบ ได้แก่ แบบที่ 1 การลดมิติแบบผสมผสานที่เกิดจากการแปลงคุณลักษณะทั้งสองวิธี แบบที่ 2 การลดมิติแบบผสมผสานที่เกิดจากการเลือกคุณลักษณะทั้งสองวิธี แบบที่ 3 การลดมิติแบบผสมผสานที่ขั้นตอนแรกเริ่มต้นด้วยวิธีการ

แปลงมิติตามด้วยวิธีการเลือกคุณลักษณะ และแบบที่ 4 การลดมิติแบบผสมผสานขั้นตอนแรกเริ่มต้นวิธีการเลือกคุณลักษณะตามด้วยวิธีการแปลงคุณลักษณะ

งานวิจัยที่ผ่านมาที่เกี่ยวข้องกับการนำเสนอการลดมิติแบบผสมผสาน ตัวอย่างเช่น งานวิจัย [58-59] ได้นำเสนอวิธีผสมผสานระหว่างวิธีการเลือกคุณลักษณะด้วยวิธีฟิลเตอร์และวิธีการเลือกคุณลักษณะด้วยวิธีแรปเปอร์แบบ SFS ซึ่งวิธีการผสมผสานนี้นิยมนำมาพิจารณาร่วมกันเพราะสามารถช่วยแก้ไขปัญหาเวลาที่ใช้การคำนวณได้ เนื่องจากข้อดีของวิธีการฟิลเตอร์ คือ เลือกเฉพาะคุณลักษณะที่สำคัญ จากนั้นคุณลักษณะที่ถูกเลือกนี้จะถูกเลือกอีกครั้งด้วยฟังก์ชันเป้าหมายหรือผลการประเมินประสิทธิภาพ แต่วิธีการนี้ยังมีจุดด้อยคือ คุณลักษณะที่ถูกเลือกขาดความสัมพันธ์ระหว่างคุณลักษณะด้วยตัวเองซึ่งมีผลกระทบต่อการใช้งานไปประยุกต์ใช้กับข้อมูลที่มีมิติสูง สำหรับงานวิจัยของ [60] นำเสนอวิธีหลอมรวมคุณลักษณะที่ได้จากสัญญาณสมองและภาพ MRI เพื่อใช้ในการจำแนกผู้ป่วยโรคจิตเภท โดยใช้วิธี multi-CCA ซึ่งวิธีนี้สามารถหลอมรวมคุณลักษณะและยังคงแสดงความสัมพันธ์ระหว่างคุณลักษณะ แต่วิธีการแปลงคุณลักษณะทั้งสองยังคงมีข้อจำกัดในการจำแนกกลุ่มข้อมูล สำหรับตัวอย่างงานวิจัยการลดมิติแบบผสมผสานแบบที่ 3 ที่เริ่มต้นขั้นตอนแรกด้วยวิธีการแปลงคุณลักษณะและตามด้วยวิธีการเลือกคุณลักษณะ ได้แก่ งานวิจัยของ [61] ได้ผสมผสานระหว่างวิธี PCA และการเลือกคุณลักษณะด้วยวิธี GA และ PSO งานวิจัยของ [62] ผสมผสานระหว่างวิธี Probabilistic PCA และ GA และงานวิจัยของ [63] ได้นำเสนอการหลอมรวมระหว่างวิธี CCA และ GA แม้ว่าการลดมิติข้อมูลการผสมผสานแบบที่ 3 นี้ข้อมูลเดิมจะถูกแปลงและถูกคัดเลือกคุณลักษณะที่เหมาะสม ซึ่งสามารถช่วยให้ข้อมูลลดลง แต่มีจุดด้อยคือใช้เวลานานในการประมวลผลของขั้นตอนของการทดสอบ เนื่องจากการคุณลักษณะที่ถูกแปลงในขั้นตอนแรกนั้นจะเปลี่ยนไป ดังนั้นขั้นตอนการเลือกคุณลักษณะจะถูกเปลี่ยนไปใหม่ทุกครั้ง สำหรับวิธีการลดมิติผสมผสานแบบที่ 4 ตามที่ผู้วิจัยได้สืบค้นงานวิจัยที่ผ่านมาอย่างไม่พบงานวิจัยใดที่นำเสนอวิธีการดังกล่าวนี้

2.8 อัลกอริทึมที่ได้แรงบันดาลใจมาจากธรรมชาติ (Nature Inspired Algorithms: NIAs)

อัลกอริทึมที่ได้แรงบันดาลใจมาจากธรรมชาติเป็นส่วนหนึ่งของการคำนวณแบบชาญฉลาด จุดประสงค์ของการพัฒนาอัลกอริทึมเหล่านี้เพื่อแก้ไขปัญหาที่ซับซ้อนทางวิศวกรรมต่างๆ รวมถึงปัญหาคำนวณมิติสูงด้วย อัลกอริทึมที่ได้แรงบันดาลใจมาจากธรรมชาติแบ่งออกเป็นสองกลุ่มใหญ่ๆ ได้แก่ อัลกอริทึมเชิงวิวัฒนาการ (Evolution algorithm) และ อัลกอริทึมปัญญาเชิงกลุ่ม (Swarm intelligence algorithm) โดยอัลกอริทึมเชิงวิวัฒนาการได้รับแรงบันดาลใจมาจากการวิวัฒนาการเชิงชีววิทยา ประกอบด้วย ขั้นตอนการสืบพันธุ์ (Reproduction) ขั้นตอนการกลายพันธุ์ (Mutation) ขั้นตอนการแลกเปลี่ยนยีน (Recombination) และขั้นตอนการคัดเลือก (Selection) โดยผลลัพธ์ที่ได้จากอัลกอริทึมเชิงวิวัฒนาการนี้ขึ้นอยู่กับสภาพแวดล้อมหรือเงื่อนไขที่ถูกกำหนดไว้ สามารถนำไปใช้ในการแก้ไขปัญหาที่

ซับซ้อนได้อย่างรวดเร็วและสามารถค้นหาคำตอบได้อย่างเหมาะสม สำหรับอัลกอริทึมปัญญาเชิงกลุ่มหรือถูกเรียกว่า เทคนิคการหาคำตอบที่เหมาะสมเชิงกลุ่มเป็นอัลกอริทึมที่ได้แรงบันดาลใจมาจากเลียนแบบพฤติกรรมกรอยู่รวมกันเป็นกลุ่มอย่างฉลาด โดยปกติปัญญาเชิงกลุ่มประกอบด้วย กลุ่มประชากรที่มีสมาชิกที่สามารถสื่อสารกันภายในกลุ่มและสามารถสื่อสารกับสภาพแวดล้อมได้ ไม่มีการควบคุมจากส่วนกลาง และมีการจัดระบบด้วยตัวเอง ขั้นตอนการค้นหาของ NIAs มีสองปัจจัยสำคัญที่ควรคำนึงถึงเพื่อให้สามารถค้นหาผลเฉลย (Solution) ที่เหมาะสมที่สุด คือ ความสมดุลของการค้นหาสิ่งใหม่ๆจากปริภูมิครอบคลุมหรือเรียกว่า Exploration และการแสวงหาผลประโยชน์จากวิธีการหาคำตอบที่มีอยู่ในปริภูมิเฉพาะที่เพื่อสามารถใช้ประโยชน์ได้อย่างคุ้มค่าที่สุดหรือเรียกว่า Exploitation เนื่องจากอัลกอริทึม NIA มีจำนวนมาก ดังนั้นวิทยานิพนธ์นี้จะกล่าวถึงบางอัลกอริทึมที่มีความเกี่ยวข้องกับการนำไปประยุกต์ใช้สำหรับการหาเซตคุณลักษณะที่เหมาะสม ได้แก่ BPSO , BCS, EBCS, และ BGWO

2.8.1 อัลกอริทึม Binary Particle Swarm Optimization (BPSO)

อัลกอริทึม BPSO ถูกพัฒนาโดย [4] มีพื้นฐานจากอัลกอริทึม PSO โดยทั่วไป อัลกอริทึม PSO ได้แรงบันดาลใจมาจากพฤติกรรมกรอพยพของสิ่งมีชีวิตที่อาศัยอยู่เป็นฝูง กระบวนการของ PSO เริ่มต้นด้วยการกำหนดพารามิเตอร์ ดังนี้ ตำแหน่งของแต่ละอนุภาค i เป็น $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$, และความเร็วของแต่ละอนุภาค $v_i = (v_{i1}, v_{i2}, \dots, v_{id})$ โดยที่ d คือ จำนวนอนุภาคทั้งหมด ระหว่างการอพยพตำแหน่งที่ดีที่สุดของแต่ละอนุภาคที่เกิดขึ้นก่อนหน้าถูกบันทึกเป็น $pbest$ และตำแหน่งที่ดีที่สุดของประชากรทั้งหมดถูกบันทึกเป็น $gbest$ การค้นหาวิธีการที่เหมาะสมที่สุดของอัลกอริทึม PSO จะถูกพิจารณาจากการอัปเดตตำแหน่งด้วยสมการ 2.45 และความเร็วของแต่ละอนุภาคที่มีความเกี่ยวข้องกัน $pbest$ และ $gbest$ ด้วยสมการ 2.46

$$x_i(t+1) = x_i(t) + v(t+1) \quad (2.45)$$

$$v_i(t+1) = w \times v_i(t) + C_1 r_1 (pbest - x_i(t)) + C_2 r_2 (gbest - x_i(t)) \quad (2.46)$$

โดยที่ t คือ จำนวนของ iteration w คือ ค่าน้ำหนักคงที่ C_1 และ C_2 คือค่าคงที่ของ acceleration และ r_1 และ r_2 คือค่าการสุ่มที่อยู่ในช่วง $[0,1]$ นอกจากนี้อัลกอริทึม PSO สามารถแบ่งเป็น continuous PSO และ binary PSO เมื่อเปรียบเทียบทั้งสองแบบแล้ว พบว่า BPSO มีความเหมาะสมสำหรับการค้นหาเซตของคุณลักษณะมากกว่า continuous PSO เนื่องจากจำนวนตัวแปรมีความยาวคงที่ ในกระบวนการของ BPSO ความเร็วจะถูกอัปเดตด้วยความเร็วและตำแหน่งของอนุภาคนั้นๆ และค่า $pbest$ และค่า $gbest$ ที่ถูกบันทึกไว้ จากนั้นความเร็วที่ถูกอัปเดตถูกแปลงด้วยฟังก์ชัน Sigmoid และนำไประบุตำแหน่งใหม่ด้วยเงื่อนไขการสุ่มด้วยสมการ

$$x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1+e^{-v_i(t+1)}} \geq \text{rand} \\ 0, & \text{otherwise} \end{cases} \quad (2.47)$$

สมาชิกของตำแหน่งแต่ละอนุภาคประกอบด้วยไบนารี 1 หรือ 0 หากไบนารีเป็น 1 แทนตำแหน่งของคุณลักษณะนั้นจะถูกเลือก และหากเป็นไบนารีเป็น 0 แทนตำแหน่งของคุณลักษณะนั้นจะไม่ถูกเลือก

2.8.2 อัลกอริทึม Binary Cuckoo Search Optimization (BCSO)

การค้นหาแบบนกกาเหว่าถูกพัฒนาขึ้นในปี 2009 โดย yang และ Deb [64] อัลกอริทึมการค้นหาแบบนกกาเหว่าได้แนวคิดมาจากพฤติกรรมการฝากไข่อย่างชาญฉลาดของนกกาเหว่า โดยนกกาเหว่าจะฝากไข่ไว้ยังรังของนกตัวอื่น (Host nest) แม่นกกาเหว่าจะเลียนแบบการวางไข่ของตัวเองให้เหมือนกับไข่ของเจ้าของรังให้มากที่สุด โดยพยายามไม่ให้นกเจ้าของรังค้นหาเจอว่าในรังมีไข่ของนกกาเหว่าอยู่ กรณีที่นกเจ้าของรังหาไข่ของนกกาเหว่าเจอจะทำลายด้วยการทิ้งไข่หรือทิ้งรังทั้งหมด เรากำหนดให้แต่ละ iteration ของการคำนวณเป็น t สำหรับ $x_i(t)$ คือไข่ฟองที่ i^{th} ในรังไข่ซึ่งการอพยพไข่หรือการปรับปรุงคำตอบใหม่ $x_i(t+1)$ ถูกปรับปรุงด้วยหลักการ Levy flights ดังสมการ

$$x_i(t+1) = x_i(t) + \alpha \oplus \text{Levy}(\lambda) \quad (2.48)$$

โดยที่ $\alpha \oplus \text{Levy}(\lambda)$ คือช่วงระยะห่างหรือระยะทางที่เกิดการเปลี่ยนแปลงของตำแหน่งของคำตอบ α คือ พารามิเตอร์ที่เกี่ยวกับขนาดการปรับปรุงคำตอบหรือ Step size ค่าพารามิเตอร์นี้ขึ้นอยู่กับปัญหาที่ต้องการค้นหาคำตอบและถูกกำหนดให้มีค่ามากกว่า 0 ($\alpha > 0$) สำหรับหลักการของ $\text{Levy}(\lambda)$ ถูกอธิบายด้วยขนาดของ step size (u) ดังสมการ

$$\text{Levy} \sim u = t^{(-\lambda)}, 1 < \lambda \leq 3 \quad (2.49)$$

สำหรับปัญหาของการเลือกเซตคุณลักษณะที่เหมาะสมในงานวิจัย [5] นำเสนออัลกอริทึมการค้นหาแบบนกกาเหว่าในรูปแบบไบนารีหรือ Binary Cuckoo Search (BCS) นอกจากนี้ในงานวิจัยดังกล่าวนี้ ได้เปรียบเทียบกับ BCS กับวิธีอื่นๆ เช่น PSO, GA, ABC พบว่าอัลกอริทึม BCS มีประสิทธิภาพในการหาเซตของคุณลักษณะมีประสิทธิภาพที่ดีกว่าวิธีอื่นๆ ขั้นตอนกระบวนการของ BCS แตกต่างจากอัลกอริทึม CS เพียงเล็กน้อย การปรับปรุงคำตอบใหม่ของ BCS $x_i(t+1)$ ถูกปรับด้วยมีเงื่อนไขการเปรียบเทียบการสุ่มด้วยการแปลงค่า $x_i(t)$ ด้วยฟังก์ชัน sigmoid ดังสมการ

$$x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1+e^{-x_i(t)}} \geq rand \\ 0, & \text{otherwise} \end{cases} \quad (2.50)$$

โดยที่ $rand$ คือค่าการสุ่มที่อยู่ช่วง $[0,1]$

ถึงแม้ว่า CS จะมีถูกนำเสนอว่าเป็นวิธีที่มีประสิทธิภาพเมื่อเทียบกับวิธีอื่นๆ แต่วิธี CS ที่ค้นหาด้วยหลักการของการกระจายแบบ Levy flight ทำให้ค่าการสุ่มกระโดดจากพื้นที่ไปยังอีกพื้นที่หนึ่ง ดังนั้นการลู่เข้าสู่คำตอบที่เหมาะสมจึงใช้ระยะเวลาานาน ดังนั้นในงานวิจัยของ [6] ได้ปรับปรุงวิธีการของ BCS เรียกว่า Extended Binary Cuckoo Search (EBCS) โดยได้เพิ่มกระบวนการกลายพันธุ์ (Mutation) ในอัลกอริทึมของ BCS โดยมีจุดประสงค์เพื่อสามารถค้นหาคำตอบในบริเวณใกล้เคียงกับคำตอบที่ดีที่สุด และยังเก็บคำตอบที่ดีที่สุดไว้ในซับเซตของคุณลักษณะที่ถูกเลือกด้วย กระบวนการดำเนินการ EBCS ซึ่งในแต่ละ iteration วิธีการหาคำตอบจะถูกสุ่มเลือกเพื่อเป็นอินพุตของการกระบวนการกลายพันธุ์ หลังจากนั้นตำแหน่งคำตอบของวิธีการหาคำตอบนั้นจะถูกสุ่มเลือก และตำแหน่งนั้นจะถูกเปลี่ยนค่าไปบนารี่ในตำแหน่งนั้นให้เป็นไบนารี่ตรงข้าม เช่น ถ้าตำแหน่งของคำตอบนั้นเป็นไบนารี่หนึ่งจะถูกเปลี่ยนเป็นศูนย์ และหากเป็นไบนารี่ศูนย์จะถูกเปลี่ยนเป็นหนึ่ง

2.8.3 อัลกอริทึม Binary Gray Wolf Optimization (BGWO)

อัลกอริทึม BGWO เป็นอัลกอริทึมที่ใช้ในการเลือกคุณลักษณะ ซึ่งปรับปรุงมาจากอัลกอริทึม GWO ดังนั้นเราจะอธิบายอัลกอริทึม GWO ก่อน อัลกอริทึม GWO เป็นอัลกอริทึมใหม่ของอัลกอริทึมวิวัฒนาการซึ่งถูกนำเสนอโดย Mirjalili และคณะในปี 2014 [65] เป็นอัลกอริทึมที่เกิดจากการเลียนแบบพฤติกรรมที่อยู่ร่วมกันของฝูงหมาป่าสีเทาในธรรมชาติ ในฝูงหมาป่าหนึ่งๆ จะมีกฎที่เข้มงวดในลำดับชั้นของการอาศัยอยู่ร่วมกันมีหัวหน้าฝูงเรียกว่า อัลฟา (α) มีหน้าที่รับผิดชอบในการตัดสินใจ และหัวหน้ารองเรียกว่า เบต้า (β) มีหน้าที่ในการช่วยหัวหน้าฝูงในการตัดสินใจ ส่วนผู้ควบคุมฝูงหมาป่าหรือหัวหน้าลำดับที่สาม เรียกว่า เดลต้า (δ) มีหน้าที่รายงานการปกครองและดูแลหมาป่าสมาชิกที่เหลือถูกเรียกว่า โอเมก้า (ω) กระบวนการล่าเหยื่อแต่ละครั้งจะมีหัวหน้าทั้งสามเป็นจำฝูง และส่งสัญญาณตำแหน่งของเหยื่อให้กับฝูงหมาป่า ดังนั้นการปรับของตำแหน่งของหมาป่า $x_i(t+1)$ หรือตำแหน่งการหาคำตอบถูกปรับจากสมการ

$$x_i(t+1) = \frac{x_1(t) + x_2(t) + x_3(t)}{3} \quad (2.51)$$

โดยที่ x_1, x_2 และ x_3 หาได้จากสมการ

$$x_1(t) = x_\alpha(t) + A_1(C_1 x_\alpha(t) - x(t)) \quad (2.52)$$

$$x_2(t) = x_\beta(t) + A_2(C_2 x_\beta(t) - x(t)) \quad (2.53)$$

$$x_3(t) = x_\delta(t) + A_3(C_3 x_\delta(t) - x(t)) \quad (2.54)$$

โดยที่ $A_i = 2ar_i - \alpha$ และ $C = 2r_2$ ซึ่ง r_1 และ r_2 คือค่าการสุ่มที่มีช่วงระหว่าง $[0, 1]$ และ α คือค่าที่ลดลงเชิงเส้นของแต่ละรอบการวนซ้ำมีค่าตั้งแต่ 2 ถึง 0 สำหรับกระบวนการ BGWO มีการอัปเดตตำแหน่งของหมาป่าแตกต่างจากอัลกอริทึม GWO แบบต่อเนื่องเพียงเล็กน้อย โดยการปรับตำแหน่งของการหาค่าตอบ $x_i(t+1)$ หาได้จากการแปลงฟังก์ชันของ sigmoid ดังสมการ

$$x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1 + e^{-m\left(\frac{x_1 + x_2 + x_3}{3} - n\right)}} \geq rand \\ 0, & \text{otherwise} \end{cases} \quad (2.55)$$

โดยที่ m และ n เป็นพารามิเตอร์ของฟังก์ชัน sigmoid

2.9 สัมประสิทธิ์ความแตกต่างของแจ๊คคาร์ต (Jaccard Dissimilarity Coefficient)

สัมประสิทธิ์ความแตกต่างของแจ๊คคาร์ตเป็นเทคนิคที่ใช้สำหรับการวัดความแตกต่างของข้อมูลเชิงบรรยาย (Attribute data) ซึ่งถูกตัดแปลงมาจากสมการความเหมือนของแจ๊คคาร์ตเพื่อใช้วัดความเหมือนกันของเวกเตอร์สองเวกเตอร์ (S_u และ S_v) ประสิทธิภาพความเหมือนของแจ๊คคาร์ตสามารถอธิบายได้ด้วยสมการ

$$\text{Similarity}(S_u, S_v) = \frac{J_{11}}{J_{11} + J_{10} + J_{01}} \quad (2.56)$$

โดยที่ J_{11} แทนจำนวนบิตของเวกเตอร์ S_u และ S_v ที่มีค่าไบนารีเป็น 1 J_{10} แทนจำนวนบิตของเวกเตอร์ S_u มีค่าไบนารีเป็น 1 และ S_v ที่มีค่าไบนารีเป็น 0 J_{01} แทนจำนวนบิตของเวกเตอร์ S_u มีค่าไบนารีเป็น 0 และ S_v ที่มีค่าไบนารีเป็น 1

สำหรับสัมประสิทธิ์ความแตกต่างของแจ๊คคาร์ตเป็นเครื่องมือที่ใช้วัดความแตกต่างของสองเวกเตอร์ (S_u และ S_v) สามารถอธิบายได้ด้วยสมการ

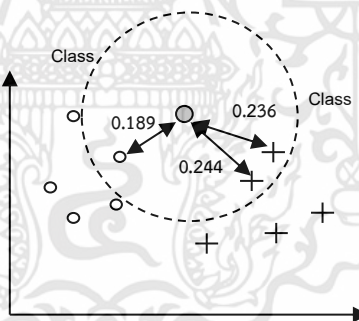
$$\text{Dissimilarity}(S_u, S_v) = 1 - \frac{J_{11}}{J_{11} + J_{10} + J_{01}} \quad (2.57)$$

งานวิจัยนี้เราได้นำเสนอ วิธีการเลือกคุณลักษณะแบบใหม่ด้วยการนำสัมประสิทธิ์ความแตกต่างของแเจ็คคาร์ตมาปรับปรุงขั้นตอนของการปรับค่าคำตอบของอัลกอริทึม BGWO แบบเดิมด้วยแนวคิดการกลายพันธุ์เรียกว่า DisBGWO ซึ่งนำเสนอในบทความ [17] รายละเอียดของอัลกอริทึมนี้หัวข้อ 3.1.2.2

2.10 การจำแนกข้อมูลด้วยอัลกอริทึมต่างๆ

2.10.1 การจำแนกข้อมูลด้วยอัลกอริทึม k- Nearest Neighbors (k-NN)

การจำแนกข้อมูลด้วยอัลกอริทึม kNN เป็นอัลกอริทึมที่ง่ายที่ใช้ในการจำแนกกลุ่มข้อมูลและเป็นวิธีการหนึ่งที่เหมาะสมกับข้อมูลที่กระจัดกระจาย โดยหลักการของวิธีการนี้คือ จะจำแนกกลุ่มข้อมูลด้วยระยะทางที่มีค่าใกล้เคียงที่สุดกับข้อมูลในแต่ละกลุ่มซึ่งกำหนดโดยค่าของข้อมูลในรอบการวนซ้ำก่อนหน้าเมื่อได้ระยะทางที่ถูกคำนวณแล้วจะถูกจัดเรียงลำดับและกำหนดข้อมูลเพื่อนบ้านตามระยะทางที่ใกล้ที่สุดจำนวน k ตามที่กำหนดไว้ การคำนวณระยะทางมีหลากหลายวิธีซึ่งการนำมาใช้งานขึ้นอยู่กับลักษณะของตัวแปร ได้แก่ การคำนวณระยะทาง Euclidean และการคำนวณระยะทาง Manhattan เหมาะสำหรับตัวแปรเป็นตัวเลข (continuous variable) ซึ่งเป็นวิธีที่นำมาใช้มากที่สุด การคำนวณระยะทาง Hamming เหมาะสำหรับตัวแปรที่ไม่ใช่ตัวเลข



รูปที่ 2.20 ตัวอย่างการจำแนกกลุ่มแบบ K nearest neighbor

รูปที่ 2.20 แสดงตัวอย่างของการจำแนกด้วยวิธี K-nearest Neighbor ข้อมูลที่ต้องการจำแนกกลุ่มคือ (○) โดยกำหนดจำนวนของข้อมูลที่นำมาพิจารณาหรือข้อมูลเพื่อนบ้านเท่ากับ $k = 3$ และการคำนวณระยะทางระหว่างข้อมูลด้วย Euclidean Distance ดังสมการ

$$D_{Euclidean}(p, q) = D_{Euclidean}(q, p) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (2.58)$$

$$= \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.59)$$

ข้อมูลที่ต้องการจำแนกคือ p และข้อมูลในแต่ละกลุ่มคือ q จากรูปที่ 2.20 ข้อมูลที่ต้องการจำแนกกลุ่ม (○) จะถูกคำนวณระยะทางจากข้อมูลอื่นๆ จากนั้นนำค่าระยะทางเหล่านั้นมาจัดเรียงจากน้อยไปมาก และกำหนดข้อมูลเพื่อนบ้านตามจำนวนที่กำหนดไว้เท่ากับ 3 ตัว จากรูปจะเห็นได้ว่าภายในขอบเขตประกอบด้วย ข้อมูลกลุ่ม (+) จำนวน 2 ตัว และข้อมูลกลุ่ม (○) จำนวน 1 ตัว การจำแนกกลุ่มของข้อมูล (○) จะพิจารณาจากจำนวนสมาชิกของกลุ่มที่มากที่สุด ดังนั้นทำให้ข้อมูล (○) ถูกจัดให้อยู่กลุ่มของ (+) ตัวอย่างงานวิจัยที่นำอัลกอริทึม kNN ไปใช้ในการจำแนกระยะการนอน ได้แก่ งานวิจัย [50] โดยได้ใช้การวัดระยะทางด้วยเทคนิค Euclidean distance และได้เปรียบเทียบกับอัลกอริทึมการจำแนกข้อมูลอื่นๆ ซึ่งผลการทดลองของงานวิจัยดังกล่าวแสดงค่าความถูกต้องแม่นยำเท่ากับร้อยละ 81.8

2.10.2 การจำแนกข้อมูลด้วยอัลกอริทึม Decision Tree (DT)

การจำแนกข้อมูลด้วยอัลกอริทึม DT หรืออัลกอริทึมต้นไม้ตัดสินใจเป็นการตัดสินใจโดยใช้โครงสร้างลำดับชั้น ที่มีลักษณะโครงสร้างคล้ายกับต้นไม้กลับหัวที่เริ่มต้นจากส่วนบนคือราก (Root) ไปสู่ส่วนล่างสุดคือใบ (left node) ภายในต้นไม้ประกอบด้วยโหนด (node) และมีคุณลักษณะ (attribute) เป็นตัวแปรที่ใช้ในการทดสอบ กิ่งของต้นไม้ (branch) แสดงถึงค่าที่เป็นไปได้ของคุณลักษณะที่ถูกทดสอบ และใบแสดงถึงกลุ่มของข้อมูล (class) หรือผลลัพธ์ของการตัดสินใจ ขั้นตอนการสร้างต้นไม้อย่างง่ายสามารถสร้างได้จากคุณลักษณะที่เป็นไปได้ สามารถอธิบายได้ด้วยตัวอย่างการตัดสินใจในการเล่นกอล์ฟ จากรูปที่ 2.21 แสดงข้อมูลของแต่ละตัวแปรที่ใช้ในการตัดสินใจ

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

รูปที่ 2.21 ตารางข้อมูลที่ใช้ในการตัดสินใจการเล่นหรือไม่เล่นกอล์ฟ

การสร้างต้นไม้ตัดสินใจมีหลายวิธีโดยวิธีเหล่านั้นได้พัฒนามาจากวิธีพื้นฐานคือวิธี ID3 โดยขั้นตอนการสร้างต้นไม้ตัดสินใจด้วยวิธี ID3 นั้นจะสร้างต้นไม้ตัดสินใจจากบนลงล่างด้วยการตัดสินใจว่าคุณลักษณะใดควรที่จะเป็นรากของต้นไม้ด้วยค่า Information Gain และค่า Entropy การคำนวณค่า Entropy ของเซตของข้อมูลหรือ $E(S)$ คำนวณได้จากสมการ

$$E(S) = \sum_{i=1}^c -p_i \log_2 p_i \quad (2.60)$$

ซึ่งชุดข้อมูลของ S ประกอบด้วย $\{s_1, s_2, \dots, s_m\}$ และความน่าจะเป็นที่จะเกิดเหตุการณ์ s_i มีค่าเท่ากับ p_i จากข้อมูลของตัวอย่างรูปที่ 2.21 จะเห็นได้ว่า จำนวนของเหตุการณ์ทั้งหมดเท่ากับ 14 เหตุการณ์ โดยผลลัพธ์คือ เล่นกอล์ฟหรือ Yes จำนวนการเกิดเหตุการณ์ทั้งหมดเท่ากับ 9 ครั้ง และไม่เล่นกอล์ฟหรือ No จำนวนการเกิดเหตุการณ์ทั้งหมดเท่ากับ 5 ครั้ง ดังนั้นจะคำนวณค่า Entropy เท่ากับ $E(S) = \text{Entropy}(5,9) = -(0.36 \log_2 0.36) - (0.64 \log_2 0.64) = 0.94$ สำหรับตัวอย่างการคำนวณการสร้างต้นไม้ตัดสินใจเพื่อเริ่มสร้างรากของต้นไม้โดยพิจารณาจากคุณลักษณะของ outlook สามารถคำนวณค่า Entropy เท่ากับ $E(\text{PlayGolf}, \text{Outlook}) = P(\text{Sunny}) \times E(3,2) + P(\text{Overcast}) \times E(4,0) + P(\text{Rainy}) \times E(2,3) = 0.693$ สำหรับค่า Information Gain หาได้จากสมการ

$$\text{Gain}(S, A) = E(S) - E(S, A) \quad (2.61)$$

ดังนั้นค่า Information Gain ของคุณลักษณะ Outlook เท่ากับ $0.94 - 0.693 = 0.247$ จากรูปที่ 2.22(a) แสดงค่า Information Gain ของแต่ละคุณลักษณะ

		Play Golf	
		Yes	No
Outlook	Sunny	3	2
	Overcast	4	0
	Rainy	2	3
Information Gain = 0.247			

		Play Golf	
		Yes	No
Temp.	Hot	2	2
	Mild	4	2
	Cool	3	1
Information Gain = 0.029			

		Play Golf	
		Yes	No
Humidity	High	3	4
	Normal	6	1
Information Gain = 0.152			

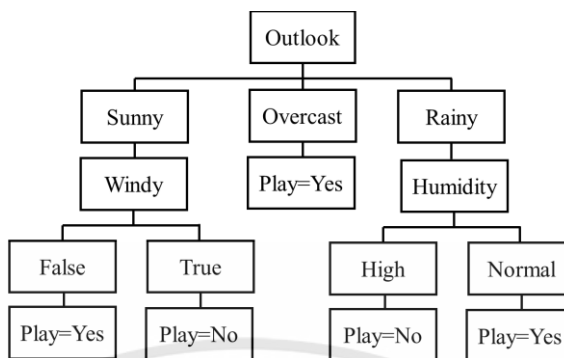
		Play Golf	
		Yes	No
Windy	False	6	2
	Ture	3	3
Information Gain = 0.048			

```

graph TD
    Outlook[Outlook] --> Sunny[Sunny]
    Outlook --> Overcast[Overcast]
    Outlook --> Rainy[Rainy]
    Sunny --> PlayYes1[Play=Yes]
    Overcast --> PlayYes2[Play=Yes]
    Rainy --> PlayYes3[Play=Yes]
  
```

รูปที่ 2.22 (a.) ค่า information Gain ของแต่ละคุณลักษณะ (b.) การสร้างต้นไม้ตัดสินใจของโหนดราก

การเลือกคุณลักษณะที่ใช้สำหรับเป็นโหนดรากจะเลือกจากค่า Information Gain ที่มากที่สุด จากรูปที่ 2.22(a) จะเห็นได้ว่า ค่า information Gain สูงสุด คือ คุณลักษณะ outlook ดังนั้นจึงถูกเลือกเป็นคุณลักษณะในการตัดสินใจสำหรับการเลือกเป็นโหนดรากแสดงดังรูปที่ 2.22(b.) โดยคุณลักษณะ Overcast มีค่า Entropy เท่ากับ 0 คือเป็นการตัดสินใจเพียงในทิศทางเดียวหรือความเป็นไปได้ที่จะเล่นกอล์ฟทั้งหมด สำหรับคุณลักษณะอื่นๆ ที่มีค่า Entropy ไม่เท่ากับศูนย์จะถูกคำนวณค่า information Gain และเลือกค่าที่มากที่สุดเพื่อกำหนดเป็นโหนดรากต่อมา และจนกระทั่งมีค่า Entropy เท่ากับศูนย์โดยข้อมูลทั้งหมดถูกจำแนกกลุ่มและสามารถสร้างต้นไม้ตัดสินใจได้ดังรูปที่ 2.23



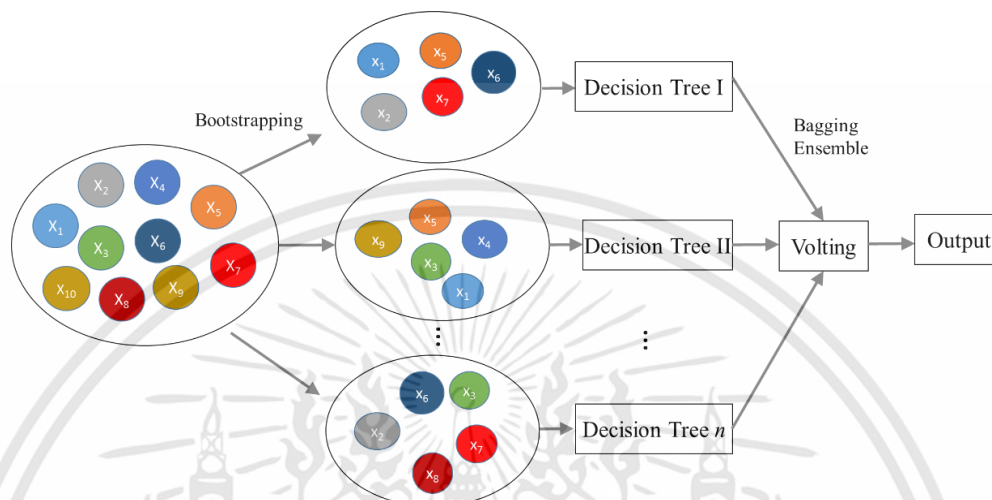
รูปที่ 2.23 ต้นไม้ตัดสินใจความเป็นไปได้ของการเล่นและไม่เล่นกอล์ฟด้วยวิธี ID3

งานวิจัยของ [43] ได้นำอัลกอริทึมต้นไม้ตัดสินใจมาใช้ในการจำแนกระยะการนอนโดยใช้คุณลักษณะที่มาจากสัญญาณ EEG เพียงอย่างเดียว ซึ่งได้แสดงค่าความถูกต้องแม่นยำของจำแนกระยะการนอนต่างๆ เท่ากับร้อยละ 80.03

2.10.3 การจำแนกข้อมูลด้วยอัลกอริทึม Random Forest (RF)

การจำแนกข้อมูลด้วยอัลกอริทึมป่าไม้สุ่มเป็นอัลกอริทึมที่ถูกนำเสนอมาเพื่อแก้ไขปัญหาการเกิด overfitting ของอัลกอริทึมต้นไม้ตัดสินใจที่เกิดจากความไม่ยืดหยุ่นของการตัดสินใจ โดยค่าความถูกต้องแม่นยำจะสูงเมื่อใช้กับเซตข้อมูลการเรียนรู้ แต่ค่าจะลดลงเมื่อนำไปใช้กับข้อมูลการทดสอบ อย่างไรก็ตาม อัลกอริทึมต้นไม้ตัดสินใจได้ปรับปรุงประสิทธิภาพของโมเดลด้วยวิธีการเล็มต้นไม้ (Pruning) แต่ยังไม่มียุทธศาสตร์มากนักเมื่อนำไปใช้กับการจำแนกข้อมูลที่มีขนาดใหญ่ สำหรับอัลกอริทึมป่าไม้สุ่มได้แก้ไขปัญหาดังกล่าวด้วยวิธีการที่เรียกว่า Bagging Ensemble หรือการสร้างผลลัพธ์ของคำตอบสุดท้ายด้วยผลลัพธ์ย่อยต่างๆ ซึ่งอัลกอริทึมป่าไม้สุ่มจะพิจารณาผลลัพธ์สุดท้ายด้วยผลลัพธ์ย่อยของโมเดลต้นไม้ตัดสินใจจำนวนหลายโมเดล ขั้นตอนของอัลกอริทึมป่าไม้สุ่มแสดงดังรูปที่ 2.24 เริ่มต้นด้วย การสุ่มเลือกข้อมูลและตัวแปรด้วยวิธี Bootstrapping จำนวนทั้งหมด n ชุด ตามจำนวนของโมเดลต้นไม้ตัดสินใจ ซึ่งแต่ละข้อมูลที่สุ่มเลือกแต่ละชุดตัวแปรจะไม่เหมือนกันแต่ตัวแปรซ้ำกันได้ จากนั้นเซตข้อมูลที่ถูกสุ่มเลือกจะนำไปสร้างต้นไม้ตัดสินใจ ผลลัพธ์สุดท้ายของคำตอบเอาท์พุทจะพิจารณาจากผลลัพธ์ย่อยที่มารวมกันหรือ Bagging ensemble โดยทั่วไปวิธีการพิจารณาผลลัพธ์ของคำตอบสุดท้ายมีหลายวิธี สำหรับวิธีที่ง่ายและใช้มากที่สุด คือ จำนวนผลลัพธ์ย่อยที่มีจำนวนที่ซ้ำกันมากที่สุดหรือ Voting ในกระบวนการออกแบบอัลกอริทึมป่าไม้สุ่มจะต้องกำหนดจำนวนของต้นไม้ซึ่งขึ้นอยู่กับความเหมาะสมของการนำไปใช้งาน การจำแนกด้วยอัลกอริทึมป่าไม้สุ่มเป็นวิธีการที่สามารถแก้ไขปัญหาหรือข้อมูลขนาดใหญ่ที่มีความซับซ้อนได้ แต่มีข้อเสียคือ เวลาในการประมวลผลจะใช้เวลานานกว่าเมื่อเปรียบเทียบกับอัลกอริทึมต้นไม้ตัดสินใจ จากงานวิจัยของ Rui Yan และคณะ [54] ได้นำเสนอการจำแนกระยะการนอนด้วยหลายชุด

สัญญาณ ซึ่งได้ทดลองเปรียบเทียบอัลกอริทึมการจำแนกข้อมูลอื่นๆ ซึ่งผลการทดลองแสดงให้เห็นว่าการจำแนกกระยะการนอนด้วยอัลกอริทึม RF ให้ค่าความถูกต้องแม่นยำสูงสุดเมื่อเทียบกับอัลกอริทึมการจำแนกข้อมูลอื่นๆ ได้แก่ binary decision tree, naïve Bayes และ Support Vector Machine



รูปที่ 2.21 กระบวนการจำแนกข้อมูลด้วยอัลกอริทึม RF

2.10.4 การจำแนกข้อมูลด้วยอัลกอริทึม Support Vector Machine (SVM)

แนวคิดของวิธีนี้คือการใช้ระนาบเพื่อแบ่งหรือจำแนกกลุ่มข้อมูลเป็นวิธีที่มีประสิทธิภาพในการจำแนกข้อมูลที่มีความซับซ้อน แต่ข้อเสียของวิธีการนี้คือเวลาที่ใช้ในการประมวลผลนาน ความยืดหยุ่นของอัลกอริทึมนี้ขึ้นอยู่กับฟังก์ชัน kernel กระบวนการจำแนกข้อมูลด้วย SVM เริ่มจากการสร้างระนาบเส้นตรง (Hyperplane) ซึ่งใช้ในการแบ่งกลุ่มข้อมูล ฟังก์ชันระนาบเส้นตรงของการตัดสินใจที่เหมาะสมแทนได้ดังสมการ

$$f(x) = w^T x + b \quad (2.62)$$

ตัวอย่างเช่น $x \in \mathcal{H}^n$ คือข้อมูลอินพุตสำหรับแบ่งกลุ่มข้อมูล $f(x)$ คือฟังก์ชันระนาบเส้นตรงของการตัดสินใจของข้อมูลดังกล่าว สำหรับกลุ่มของข้อมูลหรือ classifier label ถูกกำหนดให้เท่ากับ $y \in \{\pm 1\}$ ประกอบด้วยกลุ่ม +1 และ -1 ดังนั้นเซตข้อมูลอินพุตเขียนได้เป็น $\{(x_i, y_i), i=1, 2, \dots, l\}$ โดยที่ l คือจำนวนข้อมูลทั้งหมด ทำให้ฟังก์ชันของระนาบเส้นตรงการตัดสินใจกลุ่มสามารถเขียนแทนดังนี้

$$w^T x_i + b = +1 \quad (2.63)$$

$$w^T x_i + b = -1 \quad (2.64)$$

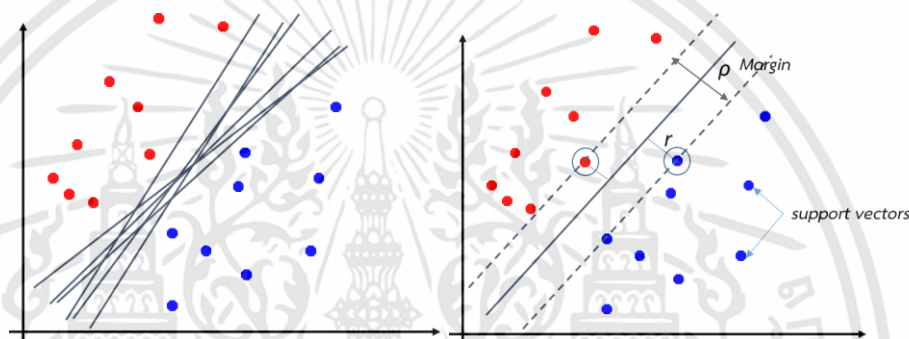
ขั้นตอนการแบ่งกลุ่มของข้อมูลที่น่ามาทดสอบจะถูกแบ่งด้วยระนาบการตัดสินใจดังนี้ ถ้าหาก

$$w^T x_i + b \geq +1 \quad (2.65)$$

ข้อมูลการทดสอบนั้นจะถูกจัดให้อยู่ในกลุ่ม +1 แต่ถ้า

$$w^T x_i + b \leq -1 \quad (2.66)$$

ข้อมูลการทดสอบนั้นจะถูกจัดให้อยู่ในกลุ่ม -1 ค่า coefficient ของ w มีความเป็นไปได้หลายค่า ดังนั้นทำให้เกิดระนาบเส้นตรงหลายเส้นดังรูป 2.22(a) การพิจารณาเลือกฟังก์ชันในการตัดสินใจจะเลือกจากความเหมาะสม ค่า w จะถูกเลือกจากค่า Margin (ρ) หรือผลรวมระยะห่างของเส้นตรงระนาบตัดสินใจถึงเส้นตรงที่ผ่านอินพุตที่ใกล้ที่สุดกับระนาบเส้นตรงของข้อมูลทั้งสองกลุ่มที่มากที่สุดดังรูป 2.22(b)



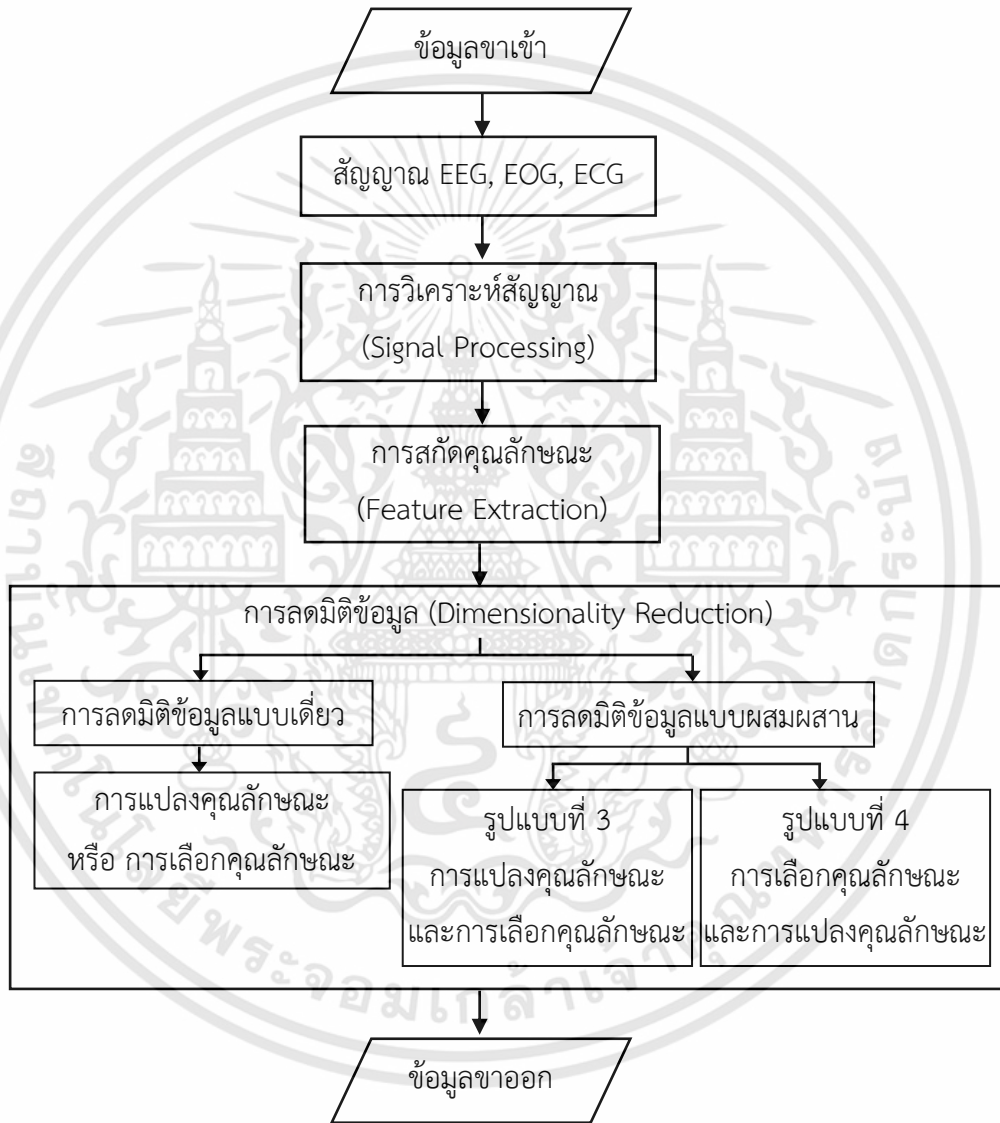
รูปที่ 2.22 (a) ระนาบเส้นตรงที่เป็นไปได้ทั้งหมดของการแบ่งกลุ่มข้อมูล
(b) ระนาบเส้นตรงที่เหมาะสมของการแบ่งกลุ่มข้อมูล

วิธี SVM แบบเชิงเส้นบางครั้งไม่เพียงพอต่อความซับซ้อนของข้อมูล ดังนั้นจึงมีการแก้ไขด้วยการแปลงข้อมูลอินพุตไปยังปริภูมิที่มีมิติที่สูงขึ้น การใช้ฟังก์ชัน kernel และการหาระนาบการตัดสินใจที่เป็นไปได้ด้วยค่าที่มากที่สุดของ Margin ฟังก์ชันที่ถูกนำมาใช้เป็น kernel ได้แก่ Polynomial, radial based function (RBF), normalized polynomial kernel เป็นต้น การเลือก kernel จะเลือกได้ตามความเหมาะสมของข้อมูล จากงานวิจัยของ Ahsan และคณะ [66] ได้ศึกษาวิธีการจำแนกกลุ่มข้อมูลของนาที่ช่วงเวลาการเกิดภาวะหยุดหายใจขณะนอนหลับ (Apnea minute) และช่วงเวลาที่ไม่เกิดภาวะหยุดหายใจขณะนอนหลับ (Non-Apnea minute) ด้วยวิธี SVM และเปรียบเทียบ kernel ในแบบต่างๆ ค่าความถูกต้องที่ได้จาก kernel แบบ Polynomial เท่ากับร้อยละ 92.85 ถือว่ามีความเหมาะสมสำหรับการจำแนก sleep apnea minute

บทที่ 3

การออกแบบระบบการวิเคราะห์การนอนอัตโนมัติ

3.1 ขั้นตอนการประมวลผลก่อน (Preprocessing)



รูปที่ 3.1 ขั้นตอนประมวลผลก่อน

จากรูปที่ 3.1 ข้อมูลขาเข้าของขั้นตอนนี้คือสัญญาณที่ได้จากการบันทึกขณะนอนหลับ ประกอบด้วย สัญญาณ EEG, EOG และ ECG จากนั้นเข้าสู่กระบวนการวิเคราะห์สัญญาณ สัญญาณ EEG และสัญญาณ EOG ถูกกรองด้วยบัตเตอร์เวิร์ด (Butterworth filter) ที่ย่านความถี่ช่วง 0.3-30 เฮิรตซ์

สำหรับสัญญาณ EEG และที่ย่านความถี่ช่วง 0.3-10 เฮิรตซ์ สำหรับสัญญาณ ECG ถูกกำจัดสัญญาณรบกวนและตรวจจับคลื่นอาร์ติงขึ้นตอนในหัวข้อ 2.3.3.1 หลังจากนั้นคำนวณหาพารามิเตอร์ต่างๆ และเข้าสู่กระบวนการลดมิติข้อมูลด้วยวิธีต่างๆ เพื่อหาเซตคุณลักษณะที่เหมาะสมเพื่อนำเป็นคุณลักษณะสำหรับขั้นตอนการประมวลผลเพื่อแบ่งกลุ่มต่อไป

3.1.1 ขั้นตอนการสกัดคุณลักษณะจากสัญญาณ EEG, EOG และ ECG

คุณลักษณะที่ใช้ในออกแบบการวิเคราะห์การนอนแบบอัตโนมัติ ทั้งการจำแนกระยะการนอนและการจำแนกภาวะหยุดหายใจขณะนอนหลับ ประกอบด้วย การสกัดคุณลักษณะด้วยเทคนิคทางโดเมนเวลา เทคนิคโดเมนความถี่ โดเมนเวลาความถี่ และเทคนิคไม่เชิงเส้น จำนวนของคุณลักษณะในแต่ละสัญญาณแสดงได้ดังตารางที่ 3.1 และรายละเอียดของแต่ละคุณลักษณะอธิบายไว้ในภาคผนวกคุณลักษณะเหล่านี้ได้ถูกเสนอจากงานวิจัยที่ผ่านมาว่ามีความสำคัญและสามารถแยกแยะความแตกต่างของระยะการนอนและบ่งบอกการเกิดภาวะหยุดหายใจขณะนอนหลับได้

ตารางที่ 3.1 จำนวนคุณลักษณะของสัญญาณ EEG, EOG, และ ECG

คุณลักษณะที่ได้จากสัญญาณไฟฟ้าของร่างกาย	จำนวน (คุณลักษณะ)
คุณลักษณะที่ได้จากสัญญาณ EEG	81
คุณลักษณะที่ได้จากสัญญาณ EOG	73
คุณลักษณะที่ได้จากสัญญาณ ECG	
คุณลักษณะที่คำนวณจากการวิเคราะห์ HRV	56
คุณลักษณะที่คำนวณจากการวิเคราะห์ EDR	31

3.1.2 ขั้นตอนการลดมิติข้อมูล

เมื่อสัญญาณแต่ละแหล่งกำเนิดถูกสกัดคุณลักษณะแล้วจะนำเข้าสู่กระบวนการลดมิติข้อมูลด้วยวิธีต่างๆ ตามหัวข้อที่ 2.6 ซึ่งแบ่งออกเป็น 2 วิธี ได้แก่ วิธีการลดมิติข้อมูลแบบเดี่ยว และวิธีการลดมิติแบบผสมผสาน ซึ่งวิธีการลดมิติแบบเดี่ยวคือ การลดมิติด้วยวิธีการแปลงคุณลักษณะหรือการลดมิติด้วยการเลือกคุณลักษณะอย่างใดอย่างหนึ่ง สำหรับวิธีการลดมิติแบบผสมผสาน ในงานวิจัยนี้เราได้มุ่งเน้นการนำเสนอและทดลองด้วยวิธีการลดมิติผสมผสานแบบที่ 3 และวิธีการลดมิติแบบผสมผสานแบบที่ 4

3.1.2.1 วิธีการลดมิติข้อมูลด้วยการแปลงข้อมูล

การลดมิติข้อมูลด้วยวิธีการแปลงคุณลักษณะนี้มีจุดประสงค์ เพื่อกำจัดคุณลักษณะที่ไม่มีความสำคัญหรือมีความสำคัญน้อยต่อการวิเคราะห์ข้อมูลออกไปและให้มีจำนวนข้อมูลลดลง ซึ่งคุณลักษณะจะถูกแปลงไปยังฟังก์ชันใหม่ที่มีค่าความสัมพันธ์ร่วมกัน วิธีการแปลงคุณลักษณะที่ใช้ในการ

ทดลองเปรียบเทียบในงานวิจัยนี้ ได้แก่ PCA, LDA, CCA และ DCCA เนื่องจากเป็นวิธีการลดมิติที่ถูกนำไปใช้งานวิจัยที่ผ่านมาที่เกี่ยวข้องกับการเพิ่มประสิทธิภาพในระบบการเรียนรู้ของเครื่องซึ่งแต่ละวิธีมีขั้นตอนดังนี้

1. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ PCA

คุณลักษณะที่ได้จากหลายแหล่งข้อมูลจะถูกหลอมรวมคุณลักษณะเพื่อลดมิติข้อมูล วิธีหลอมรวมคุณลักษณะของหลายชุดข้อมูลด้วยวิธี PCA มีขั้นตอนดังนี้

ขั้นตอนที่ 1. กำหนดเซตของคุณลักษณะทั้งหมดเท่ากับ X ประกอบด้วย ตัวอย่างจำนวน N ตัวอย่าง ซึ่งประกอบด้วยคุณลักษณะที่มาจากแหล่งกำเนิดที่ 1 ถึง m ที่เรียงต่อกันแบบอนุกรม $X = \{X_{N1}, X_{N2}, \dots, X_{Nm}\}$ ดังตาราง 2.5 หลังจากนั้นเซตคุณลักษณะ X ถูก centralize ด้วยค่ากลางของข้อมูล $E\{X\}$ หรือ $X = X - E\{X\}$

ขั้นตอนที่ 2. เซตคุณลักษณะ X ถูกคำนวณหาเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) ซึ่งเป็นความสัมพันธ์ร่วมของเซตคุณลักษณะ X โดย $cov(X) = X^T X$ ที่มีขนาด $m \times m$

ขั้นตอนที่ 3 จากนั้นแตกองค์ประกอบของเมตริกซ์ $cov(X)$ ดังสมการ

$$Av = \lambda v \quad (3.1)$$

โดยที่ v คือเวกเตอร์ไอเกน (Eigen Vector) และ λ คือ ค่าไอเกน (Eigen Value) ที่มีความสอดคล้องกัน

ขั้นตอนที่ 4. จัดเรียง $diagonal(\lambda)$ จากค่าสูงสุดไปต่ำสุด จากนั้นจำนวนองค์ประกอบหลักถูกเลือกมีขนาด d ขึ้นอยู่เงื่อนไขการเลือกในงานวิจัยนี้เราได้เลือกตามค่าร้อยละความเชื่อมั่น (Confidence) และเมตริกซ์ของไอเกนจะมีขนาดเท่ากับจำนวนองค์ประกอบที่เลือก ดังนั้นการลดมิติข้อมูลด้วยการแปลงด้วยวิธี PCA จะมีขนาดเมตริกซ์ของเซตคุณลักษณะจะถูกลดลงเหลือเพียง $d \times N$ โดยที่ $d < m$

2. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ LDA

วิธีหลอมรวมคุณลักษณะของหลายชุดข้อมูลเพื่อลดมิติด้วยวิธี LDA มีขั้นตอนดังนี้

ขั้นตอนที่ 1. กำหนดเซตของคุณลักษณะทั้งหมดเท่ากับ X ประกอบด้วย ตัวอย่างจำนวน N ตัวอย่าง ซึ่งประกอบด้วยคุณลักษณะที่มาจากแหล่งกำเนิดที่ 1 ถึง m ที่เรียงต่อกันแบบอนุกรม $X = \{X_{N1}, X_{N2}, \dots, X_{Nm}\}$ ดังตารางที่ 2.5 โดยข้อมูลมีจำนวนกลุ่ม C กลุ่มหรือ $c = \{1, 2, \dots, C\}$ หลังจากนั้นคำนวณเวกเตอร์เฉลี่ยของตัวอย่างในแต่ละกลุ่ม m_i

ขั้นตอนที่ 2. คำนวณเมตริกซ์การกระจายภายในกลุ่ม (Within-class Scatter Matrices) และเมตริกซ์การกระจายระหว่างกลุ่ม (Between-class Scatter Matrices) ดังสมการที่ 2.29 และ 2.30

ขั้นตอนที่ 3. แก้มการหาไอเกนเวกเตอร์และค่าไอเกนของเมตริกซ์ $S_w^{-1}S_b$ ด้วยวิธี Singular Value Decomposition (SVD)

ขั้นตอนที่ 4. ขั้นตอนของการเลือกจำนวนมิติที่ลดลงเหมือนกับขั้นตอนที่ 4 ของวิธี PCA

3. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ CCA

วิธีหลอมรวมคุณลักษณะของหลายชุดข้อมูลเพื่อลดมิติด้วยวิธี CCA มีขั้นตอนดังนี้

ขั้นตอนที่ 1. แต่ละเซตของคุณลักษณะที่มาจากแหล่งที่มาต่างๆ ถูกหลอมรวมคุณลักษณะ โดยจะดำเนินการทีละคู่ รวมทั้งหมด $K(K-1)/(2 \times I)$ โดยที่ K คือ จำนวนของเซตคุณลักษณะทั้งหมด กำหนดให้คุณลักษณะคู่แรก เป็น เซตคุณลักษณะ X_1 และเซตคุณลักษณะ X_2 คุณลักษณะของแต่ละเซตจะเท่ากันหรือไม่ก็ได้ มีขนาดเป็น $N \times m$ และ $N \times p$ ตามลำดับ ข้อมูลทั้งสองถูก centralize ด้วยค่ากลางของข้อมูลจะได้ $X_1 = X_1 - E\{X_1\}$ และ $X_2 = X_2 - E\{X_2\}$

ขั้นตอนที่ 2. หาเมตริกซ์ความสัมพันธ์ร่วมภายในเซตคุณลักษณะ $X_1^T X_1$ และ $X_2^T X_2$ และหาเมตริกซ์ความสัมพันธ์ร่วมระหว่างเซตคุณลักษณะ ทั้งสอง $X_1^T X_2$

ขั้นตอนที่ 3. คำนวณหาค่า w_{X1} และ w_{X2} ที่สอดคล้องกับค่าไอเกน λ ตามสมการ 2.32

ขั้นตอนที่ 4. ค่าไอเกนถูกเรียงลำดับจากมากไปหาน้อย โดยขนาดของไอเกนเวกเตอร์จะเท่ากับ $\min(m, p)$ หลังจากนั้นลดจำนวนคุณลักษณะเหลือ $d < \min(m, p)$ โดยจำนวนตัวแปรที่ถูกเลือกตามค่าความเชื่อมั่นที่เหมาะสม ดังนั้นค่าไอเกนที่ถูกลดลงจะได้ $\lambda_k, k=1, \dots, d, \Omega_{X1}=[w_{X1}^1, \dots, w_{X1}^d]$ และ $\Omega_{X2}=[w_{X2}^1, \dots, w_{X2}^d]$

ขั้นตอนที่ 5. ดำเนินการหาค่าไอเกนและเวกเตอร์ไอเกนของแต่ละคู่ที่เหลือจนครบ เมื่อได้ค่าไอเกนของแต่ละคู่แล้ว ขั้นตอนการหลอมรวมคุณลักษณะเข้าด้วยกันสามารถหลอมรวมได้ทั้งแบบอนุกรม $\Omega_{X1}^T X_1 + \Omega_{X2}^T X_1$ หรือแบบขนาน $\begin{bmatrix} \Omega_{X1}^T X_1 \\ \Omega_{X2}^T X_2 \end{bmatrix}$ จากนั้นจะเข้าสู่การบวนการจำแนกกลุ่มข้อมูลต่อไป

4. วิธีการลดมิติด้วยวิธีแปลงคุณลักษณะ DCCA

วิธีหลอมรวมคุณลักษณะของหลายชุดข้อมูลเพื่อลดมิติด้วยวิธี DCCA มีขั้นตอนดังนี้

ขั้นตอนที่ 1. แต่ละเซตของคุณลักษณะที่มาจากแหล่งที่มาต่างๆ ถูกหลอมรวมคุณลักษณะ โดยจะดำเนินการทีละคู่ รวมทั้งหมด $K(K-1)/(2 \times I)$ โดยที่ K คือ จำนวนของเซตคุณลักษณะทั้งหมด กำหนดให้คุณลักษณะคู่แรก เป็นเซตคุณลักษณะ X_1 และเซตคุณลักษณะ X_2 คุณลักษณะของแต่ละเซตจะเท่ากันหรือไม่ก็ได้ มีขนาดเป็น $N \times m$ และ $N \times p$ ตามลำดับ โดยมีข้อมูลกลุ่ม C ข้อมูลทั้งสองถูก centralize ด้วยค่ากลางของข้อมูลจะได้ $X_1 = X_1 - E\{X_1\}$ และ $X_2 = X_2 - E\{X_2\}$

ขั้นตอนที่ 2. คำนวณหาเมตริกซ์ A ตามสมการ 2.38 และใช้ในคำนวณสมการ 2.43

ขั้นตอนที่ 3. คำนวณหาค่า w_{X1} และ w_{X2} ที่สอดคล้องกับค่าไอเกน λ ตามสมการที่ 2.44

ขั้นตอนที่ 4. ค่าไอเกนถูกเรียงลำดับจากมากไปหาน้อย โดยขนาดของไอเกนเวกเตอร์จะเท่ากับ $\min(m,p)$ หลังจากนั้นลดจำนวนคุณลักษณะเหลือ $d < \min(m,p)$ โดยจำนวนตัวแปรที่ถูกเลือกตามค่าความเชื่อมั่นที่เหมาะสม ดังนั้นค่าไอเกนที่ถูกกลดลงจะได้ $\lambda_k, k=1, \dots, d, \Omega_{X1}=[w_{X1}^1, \dots, w_{X1}^d]$ และ $\Omega_{X2}=[w_{X2}^1, \dots, w_{X2}^d]$

ขั้นตอนที่ 5. ดำเนินการหาค่าไอเกนและเวกเตอร์ไอเกนของแต่ละคู่ที่เหลือจนครบ เมื่อได้ค่าไอเกนของแต่ละคู่แล้ว ขั้นตอนการหลอมรวมคุณลักษณะเข้าด้วยกันสามารถหลอมรวมได้ทั้งแบบอนุกรม $\Omega_{X1}^T X_1 + \Omega_{X2}^T X_2$ หรือแบบขนาน $\begin{bmatrix} \Omega_{X1}^T X_1 \\ \Omega_{X2}^T X_2 \end{bmatrix}$ จากนั้นจะเข้าสู่การบวนการจำแนกกลุ่มข้อมูลต่อไป

3.1.2.2 วิธีการลดมิติข้อมูลด้วยการเลือกคุณลักษณะแบบวิธีแรปเปอร์

จากหัวข้อ 2.6.2 ได้อธิบายว่าการลดมิติด้วยวิธีการเลือกคุณลักษณะด้วยวิธีแรปเปอร์มี 2 วิธี ได้แก่ วิธีการเลือกคุณลักษณะโดยลำดับ และวิธีการเลือกคุณลักษณะด้วยวิธีการสุ่มเลือก โดยใช้ อัลกอริทึมการค้นหาคำตอบด้วยวิธีการแก้ปัญหาแบบฮิวริสติกซึ่งเป็นอัลกอริทึมที่ได้แรงบันดาลใจมาจากธรรมชาติ ได้แก่ BPSO, BCS, EBCS, BGWO การเลือกเซตของคุณลักษณะด้วยวิธีแรปเปอร์จะเลือกเซตคุณลักษณะที่เหมาะสมด้วยฟังก์ชันความเหมาะสม (Fitness function) งานวิจัยนี้ได้มุ่งเน้นการออกแบบอัลกอริทึมการลดมิติข้อมูลที่มีประสิทธิภาพสำหรับการจำแนกข้อมูลด้วยความแม่นยำและขณะเดียวกันจำนวนคุณลักษณะน้อย ดังนั้นขั้นตอนลดมิติด้วยการเลือกคุณลักษณะด้วยวิธีแรปเปอร์ เซตของคุณลักษณะสำหรับการเรียนรู้จะถูกแบ่งออกเป็น 2 ส่วน เพื่อใช้คำนวณค่าของฟังก์ชันความเหมาะสม ได้แก่ เซตคุณลักษณะสำหรับสร้างโมเดล และเซตคุณลักษณะสำหรับการสอบทวน (Validation) ดังนั้นเรากำหนดฟังก์ชันความเหมาะสมของงานวิจัยนี้ด้วยการคำนวณจากค่าความถูกต้องของอัลกอริทึมการจำแนกกลุ่มและจำนวนคุณลักษณะเพื่อใช้ในการเลือกคุณลักษณะที่เหมาะสม ฟังก์ชันความเหมาะสม (F) สามารถคำนวณได้จาก

$$F = \left(\gamma \left(\frac{t}{T} \right) + (1 - \gamma) \left(\frac{f_a - f_s}{f_a} \right) \right) \times 100 \quad (3.2)$$

โดยที่ γ คือค่าน้ำหนักของวัตถุประสงค์ที่ต้องการให้ความสำคัญ กรณีที่ต้องการให้ความสำคัญกับค่าความถูกต้องแม่นยำมากกว่าจำนวนคุณลักษณะค่าคงที่ดังกล่าวจะมีค่ามากกว่า t คือจำนวนของตัวอย่างที่ตอบถูก T คือ จำนวนตัวอย่างทั้งหมด f_s คือ จำนวนของคุณลักษณะที่ถูกเลือก และ f_a คือจำนวนคุณลักษณะทั้งหมด

1. วิธีการเลือกคุณลักษณะด้วย BPSO

วิธีเลือกคุณลักษณะของหลายชุดข้อมูลเพื่อลดมิติด้วยอัลกอริทึม BPSO เริ่มต้นด้วย กำหนดค่าเริ่มต้นต่างๆ ได้แก่ จำนวนรอบการค้นหา ค่าตอบ กำหนดจำนวนของกลุ่มอนุภาคหรือโซลูชันการค้นหา ค่าตอบ จำนวนประชากรในกลุ่มอนุภาคหรือค่าคำตอบเท่ากับจำนวนคุณลักษณะทั้งหมด และค่าคงที่ C_1 , C_2 และ w ตามสมการ 2.46 จากนั้นเซตของคุณลักษณะทั้งหมด $X=\{X_1, X_2, \dots, X_m\}$, m คือ จำนวนเซตของคุณลักษณะที่มาจากแต่ละแหล่งกำเนิดถูกแบ่งออกเป็น 2 เซต คือเซตคุณลักษณะสำหรับสร้างโมเดล และเซตคุณลักษณะสำหรับการสอบทวน และดำเนินการตามอัลกอริทึม BPSO ตามรูปที่ 3.2 ในการเลือกคุณลักษณะของแต่ละวิธีหาคำตอบที่เหมาะสมจะเลือกตามค่าโบนารีที่ถูกแปลงด้วยฟังก์ชัน sigmoid ค่าของคำตอบที่ได้หากเป็น 1 ตำแหน่งคุณลักษณะนั้นถูกเลือก หากค่าของคำตอบเป็น 0 ตำแหน่งคุณลักษณะนั้นไม่ถูกเลือก

อัลกอริทึม BPSO

Input : P : the population size, T : maximum iteration, D : Dimensionality of search space (total number of features), The parameters C_1 , C_2 , w

Output: $gbest$, training and validation accuracies of the selected features

Begin

```

1  divide dataset into a training set and a validation set
2  randomly initialize the position each particle;
3  for  $t=1$  to  $T$  do
4  evaluate fitness (3.2) of each particle;
5  for  $i=1$  to  $P$  do
6  update the  $pbest$  of particle  $i$ ;
7  update the  $gbest$  of particle  $i$ ;
8  end
9  for  $i=1$  to  $P$  do
10 for  $d=1$  to  $D$  do
11 update the velocity of particle  $i$  as
12  $v_i(t+1) = w \times v_i(t) + C_1 r_1 (pbest - x_i(t)) + C_2 r_2 (gbest - x_i(t));$ 
13 update the position of particle  $i$  as  $x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1+e^{-v_i(t+1)}} \geq rand \\ 0, & \text{otherwise} \end{cases};$ 
14 end
15 end
16 end

```

13 return the position of $gbest$ and training and validation accuracies of the selected features;

รูปที่ 3.2 อัลกอริทึม BPSO สำหรับการเลือกคุณลักษณะ

2. วิธีการเลือกคุณลักษณะด้วย BCS

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีเลือกคุณลักษณะของหลายชุดข้อมูลเพื่อลดมิติด้วยวิธี BCS เริ่มต้นด้วยการ กำหนดค่าเริ่มต้นของจำนวนรอบการค้นหาคำตอบ จำนวนของรังไข่หรือจำนวนวิธีการค้นหาคำตอบ จำนวนประชากรไขในรังไข่หรือจำนวนคำตอบเท่ากับจำนวนคุณลักษณะทั้งหมด กำหนดค่าพารามิเตอร์ความน่าจะเป็นในการละทิ้งไข่ P_d เท่ากับ 0.25 กำหนดค่าพารามิเตอร์ของการกระจาย Levy (λ) และกำหนดค่าน้ำหนักของ Step size (α) ตามสมการ 2.47 อัลกอริทึมของ BCS แสดงดังรูปที่ 3.3

อัลกอริทึม BCS

Input: N : the number of nests size, T : maximum iteration, D : Dimensionality of search space (total number of features), the parameters of Levy fight λ and α .

Output: *Global best*, training and validation accuracies of the selected features

Begin

```

1  divide dataset into a training set and a validation set
2  for  $i=1$  to  $N$  do
3      randomly initialize each dimension  $\{0,1\}$ ;
4      evaluate fitness (3.2) of new solution  $F_i$ ;
5      find the best nest to be Global best
end
6  for  $t=1$  to  $T$  do
7      for  $i=1$  to  $N$  do
8          generate new solution by  $x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1+e^{-x_i(t)}} \geq \text{rand} \\ 0, & \text{otherwise} \end{cases}$ ;
9          evaluate fitness (3.2) of new solution  $F_j$ ;
10         if  $F_j > F_i$  then
11             replace the solution  $i$  by the new solution  $j$ ;
12         end
13         if  $F_j > \text{Global best}$  then
14             replace the global solution by the new solution  $j$ ;
15         end
16         select the worst nest according to  $P_d=0.25$  and replace them for new solution;
17     end
18     return the Global best, training and validation accuracies of the selected features;
19 end

```

รูปที่ 3.3 อัลกอริทึม BCS สำหรับการเลือกคุณลักษณะ

3. วิธีการเลือกคุณลักษณะด้วย EBCS

วิธีเลือกคุณลักษณะด้วยวิธี EBCS มีขั้นตอนคล้ายกับวิธี BCS แต่จะมีการเพิ่มขึ้นขั้นตอนการสร้างวิธีการค้นหาคำตอบด้วยการกลายพันธุ์เข้ามาด้วย วิธีการค้นหาคำตอบจะถูกสุ่มเลือกจำนวน S_k วิธีการเพื่อสร้างวิธีการค้นหาคำตอบด้วยวิธีการกลายพันธุ์ โดยค่าคำตอบของวิธีค้นหาคำตอบที่ถูกสุ่มเลือกนั้นจะถูกสุ่มเลือกคำตอบจำนวน N_k และค่าของคำตอบที่ถูกสุ่มนั้นถูกสลับค่าของคำตอบให้เป็นค่าตรงข้าม เช่น หากค่าไบนารีเป็น 1 ถูกเปลี่ยนเป็น 0 หรือหากค่าไบนารีเป็น 0 ถูกเปลี่ยนเป็น 1 หลักจากนั้นวิธีการ

ค้นหาคำตอบด้วยการกลายพันธุ์ S_k ถูกประเมินค่าความเหมาะสมด้วยสมการ 3.2 จะได้ฟังก์ชันความเหมาะสม F_{mut} และถูกเปรียบเทียบกับฟังก์ชันความเหมาะสม F_i ถ้าหาก F_{mut} มากกว่า F_i วิธีการค้นหาคำตอบที่ดีที่สุดจะถูกแทนด้วยวิธีการค้นหาตอบที่กลายพันธุ์แล้ว $S_k(t)$ โดยรายละเอียดของอัลกอริทึมของ BCS แสดงได้ดังรูปที่ 3.4

อัลกอริทึม EBCS

Input : N : the number of nests size, T : maximum iteration, D : Dimensionality of search space (total number of features), the parameters of Levy fight λ and α , the number of mutations N_{mut} , the number of mutated solution S_{mut}

Output: *Global best*, training and validation accuracies of the selected features

Begin

```

1  divide dataset into a training set and a validation set
2  for i=1 to N do
3      randomly initialize each dimension {0,1};
4      evaluate fitness (3.2) of new solution  $F_i$ ;
5      find the best nest to be Global best
6  end
7  for  $t=1$  to  $T$  do
8      Create Line 7-13 similar to line 7-13 of BCS algorithm as shown in Figure 3.3
9      randomly select  $k$  nest to be the mutation nests  $S_{mut}$ 
10     for  $k=1$  to  $S_{mut}$ 
11         randomly select  $N_{mut}$  dimensions ( $p=1, \dots, N_{mut}$ ) do
12             if value of  $p = 1$  then
13                 convert to be 0
14             if else value of  $p = 0$  then
15                 convert to be 1
16             If else
17                 end
18             end
19         end
20         evaluate fitness (3.2) of the mutation solution  $F_{mut}$ 
21         if  $F_{mut} > F_i$  then
22             replace the solution  $i$  by the mutation solution  $S_{mut}$ ;
23         end
24         if  $F_{mut} > Global\ best$  then
25             replace the global solution by the mutation solution  $S_{mut}$  ;
26         end
27         select the worst nest according to  $P_\alpha = 0.25$  and replace them for new solution;
28     end
29     return the Global best, training and validation accuracies of the selected features;
30 end

```

รูปที่ 3.4 อัลกอริทึม EBCS สำหรับการเลือกคุณลักษณะ

4. วิธีการเลือกคุณลักษณะด้วย BGWO

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีเลือกคุณลักษณะของหลายชุดข้อมูลเพื่อลดมิติด้วยวิธี BGWO เริ่มต้นด้วยการกำหนดค่าเริ่มต้นของจำนวนรอบการค้นหาคำตอบ จำนวนประชากรหมาป่าในฝูงหรือวิธีการค้นหาคำตอบ จำนวนค่าของคำตอบของวิธีการค้นหาเท่ากับจำนวนคุณลักษณะทั้งหมด และกำหนดค่า m และ n พารามิเตอร์ของฟังก์ชัน sigmoid ตามสมการ 2.55 เซตคุณลักษณะที่ได้จากขั้นตอน 3.1.1 ถูกนำมาใช้เลือกเซตคุณลักษณะที่เหมาะสมด้วยอัลกอริทึมการเลือกคุณลักษณะด้วย BGWO แสดงดังรูปที่ 3.5

อัลกอริทึม BGWO

Input : N : the population of gray wolves in the pack, T : maximum iteration, D : Dimensionality of search space (total number of features), the parameters of m and n of sigmoid function.

Output: the optimal wolf position, training and validation accuracies of the selected features

Begin

- 1 divide dataset into a training set and a validation set;
 - 2 randomly initialize the population of wolves positions at $[0,1]$;
 - 3 evaluate fitness of each gray wolf using Equation 3.2;
 - 4 find $Alpha$ =the gray wolf with the first maximum fitness, $Beta$ =the gray wolf with the second maximum fitness, $Delta$ =the gray wolf with the third maximum fitness
 - 5 for $t=1$ to T do
 - 6 for $n=1$ to N do
 - 7 calculate x_1, x_2, x_3 as Equation 2.52-2.54;
 - 8 update the position of the current wolf by $x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1+e^{-m\left(\frac{x_1+x_2+x_3}{3}\right)-n}} \geq rand \\ 0, & \text{otherwise} \end{cases}$;
 - 9 end
 - 9 evaluate fitness of all grey wolves using Equation 3.2
 - 10 update $Alpha, Beta, Delta$ solutions
 - 11 end
 - 11 return $Alpha$ solution and training and validation accuracies of the selected features;
-

รูปที่ 3.5 อัลกอริทึม BGWO สำหรับการเลือกคุณลักษณะ

5. วิธีการเลือกคุณลักษณะด้วย DisBGWO

เพื่อเพิ่มประสิทธิภาพของการค้นหาคำตอบของอัลกอริทึม BGWO เรานำเสนออัลกอริทึมที่เรียกว่า DisBGWO ซึ่งเป็นอัลกอริทึมที่นำแนวคิดของกระบวนการ Mutation มาประยุกต์ใช้และเพิ่มในขั้นตอนของ BGWO แบบเดิม ซึ่งอัลกอริทึมที่นำเสนอนี้สามารถค้นหาคำตอบโดยพิจารณาทั้งข้อมูลปริภูมิเฉพาะที่และปริภูมิครอบคลุม นอกจากนี้เราได้สร้างวิธีการค้นหาคำตอบของ Mutation ด้วยการวัดความแตกต่างระหว่างสองเวกเตอร์ด้วยค่าสัมประสิทธิ์ Jaccard

อัลกอริทึมการเลือกคุณลักษณะที่นำเสนอเรียกว่า DisBGWO เริ่มต้นด้วยกำหนดค่าเริ่มต้นของจำนวนรอบการค้นหาคำตอบ จำนวนประชากรหมาป่าในฝูงหรือวิธีการค้นหาคำตอบ จำนวนคำตอบของวิธีการค้นหาเท่ากับจำนวนคุณลักษณะทั้งหมด กำหนดค่า m และ n พารามิเตอร์ของฟังก์ชัน sigmoid

ตามสมการ 2.55 ขั้นตอนการสร้างวิธีการค้นหาคำตอบด้วยการกลายพันธุ์ของฝูงหมาป่าใหม่ (Mutated Solution) $S_m(t)$ ด้วยหลักการของ differential evolution เริ่มจากนำวิธีการค้นหาคำตอบที่มีค่ามากที่สุดสามอันดับแรก ประกอบด้วย $S_1(t)$, $S_2(t)$ และ $S_3(t)$ ซึ่งเราสามารถหาสร้างวิธีการค้นหาคำตอบการกลายพันธุ์ $S_m(t)$ ด้วยสมการ

$$S_m(t) = S_1(t) + \phi(S_2(t) - S_3(t)) \quad (3.3)$$

โดยที่ ϕ ค่าคงที่ แปรจากสมการ 3.3 เพื่อให้เป็นไปตามหลักการของการหาค่าสัมประสิทธิ์ของ Jaccard ในหัวข้อที่ 2.9 เราสามารถเขียนสมการใหม่ได้ด้วย

$$S_m(t) - S_1(t) = \phi(S_2(t) - S_3(t)) \quad (3.4)$$

เมื่อเครื่องหมาย “-” ถูกแทนด้วย Dis ซึ่งหมายถึง ความต่างระหว่างสองเวกเตอร์ สามารถเขียนได้ดังสมการ

$$Dis(S_m(t), S_1(t)) = \phi \times Dis(S_2(t), S_3(t)) \quad (3.5)$$

จากนั้นเราใช้เครื่องหมาย “ \sim ” แทนที่ “=” เพราะความต่างระหว่าง $S_m(t)$ และ $S_1(t)$ ควรจะมีค่าใกล้เคียงกับความต่างระหว่าง $S_2(t)$ และ $S_3(t)$ มากที่สุด ดังนั้น $S_m(t)$ สามารถหาได้จากสมการ

$$\min\{Dis(S_m(t), S_1(t)) - \phi \times Dis(S_2(t), S_3(t))\} \quad (3.6)$$

การแก้สมการหาเวกเตอร์ของ $S_m(t)$ ด้วยเงื่อนไขของสมการทางคณิตศาสตร์ จากนั้นวิธีการค้นหาคำตอบด้วยการกลายพันธุ์ที่ได้ (S_m) ถูกนำไปคำนวณค่าฟังก์ชันความเหมาะสมได้ F_m และถูกนำไปเปรียบเทียบกับค่าฟังก์ชันความเหมาะสมกับ F_i ถ้าหากค่าฟังก์ชันความเหมาะสม F_m มากกว่าค่าฟังก์ชันความเหมาะสม F_i วิธีการค้นหาคำตอบที่ดีที่สุดหรือวิธีการค้นหาของหัวหน้าหมาป่า $x_{alpha}(t)$ จะถูกแทนที่ด้วยวิธีการค้นหาคำตอบ $S_m(t)$ รายละเอียดอัลกอริทึม DisBGWO แสดงดังรูปที่ 3.6

อัลกอริทึม DisBGWO

Input : N : the population of gray wolves in the pack, T : maximum iteration, D : Dimensionality of search space (total number of features), parameters of m and n of sigmoid function.

Output: the optimal wolf position, training and validation accuracies of the selected features

Begin

- 1 divide dataset into a training set and a validation set;
- 2 randomly initialize the population of wolves positions at $[0,1]$;
- 3 evaluate fitness of each gray wolf using Equation 3.2;

```

4   find  $Alpha$ =the gray wolf with the first maximum fitness,  $Beta$ =the gray wolf with the
    second maximum fitness,  $Delta$  =the gray wolf with the third maximum fitness
5   for  $t=1$  to  $T$  do
6     for  $n=1$  to  $N$  do
7       calculate  $x_1, x_2, x_3$  as Equation 2.52-2.54;
8       update the position of the current wolf by  $x_i(t+1) = \begin{cases} 1, & \text{if } \frac{1}{1 + e^{-m\left(\frac{x_1+x_2+x_3}{3}\right)^{-n}}} \geq rand \\ 0, & \text{otherwise} \end{cases}$ ;
9     end
10    evaluate fitness of all grey wolves using Equation 3.2 ;
11    find three best solution  $S_1, S_2, S_3$  to create the mutation solution using dissimilarity
    Jaccard coefficient as  $S_m(t) = S_1(t) + \phi(S_2(t) - S_3(t))$ ;
12    evaluate fitness of the mutation solution  $F_m$ ;
13    if  $F_m > F_{Alpha}$  then
14      Replace the alpha solution  $Alpha$  by the mutation solution  $S_m$ ;
15    end
16    update  $Alpha, Beta, Delta$  solutions
17  end
18  return  $Alpha$  solution and training and validation accuracies of the selected features;

```

รูปที่ 3.6 อัลกอริทึม DisBGWO สำหรับการเลือกคุณลักษณะ

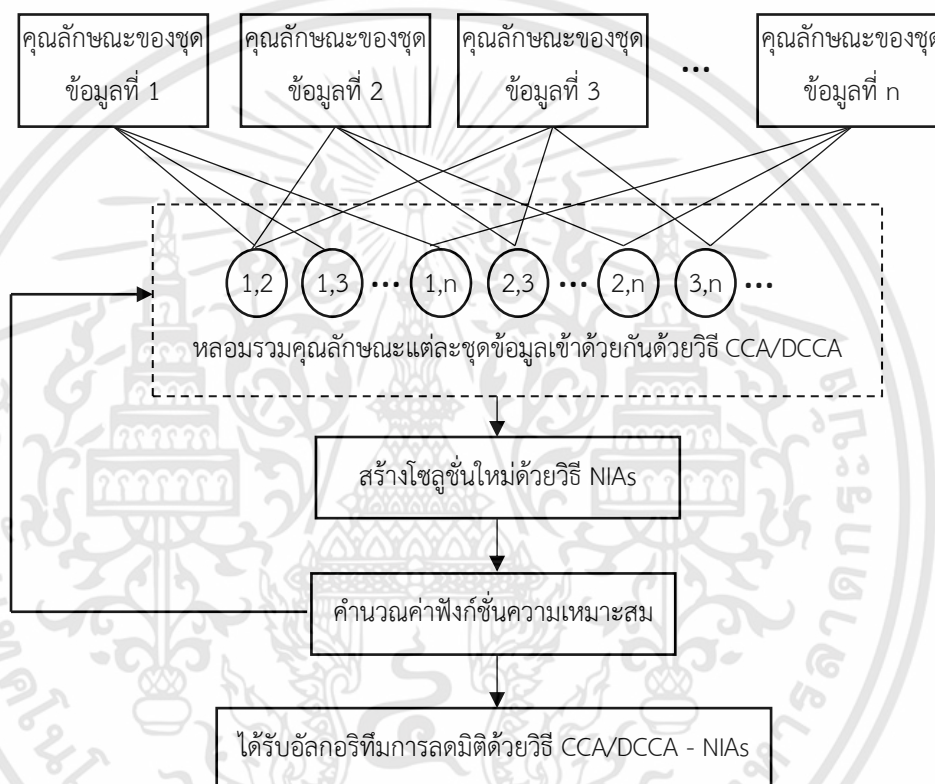
3.1.2.3 วิธีการลดมิติแบบผสมผสานด้วยอัลกอริทึมที่นำเสนอ

เพื่อเพิ่มประสิทธิภาพของการวิเคราะห์การนอนอัตโนมัติ งานวิจัยนี้ได้มุ่งเน้นการเพิ่มประสิทธิภาพของการจำแนกข้อมูลด้วยขั้นตอนการลดมิติข้อมูล เราได้นำเสนออัลกอริทึมใหม่ซึ่งเป็นวิธีการลดมิติแบบผสมผสานตามรูปแบบที่ 3 และรูปแบบที่ 4 ตามคำอธิบายในหัวข้อ 2.7 เรียกว่าวิธีการลดมิติแบบผสมผสาน CCA/DCCA-NIAs และ NIAs-CCA/DCCA โดยกระบวนการลดมิติแบบผสมผสาน CCA/DCCA-NIAs เป็นอัลกอริทึมที่เกิดจากการผสมผสานของวิธีการลดมิติสองวิธี โดยเริ่มต้นด้วยวิธีการหลอมรวมคุณลักษณะด้วยวิธีการแปลง CCA/DCCA ซึ่งพิจารณาความสัมพันธ์ระหว่างเซตข้อมูลหลายเซต จากนั้นตามด้วยวิธีการเลือกคุณลักษณะที่เหมาะสมด้วยวิธีการค้นหาคำตอบแบบสุ่มด้วยอัลกอริทึม NIAs แสดงดังรูป 3.7 และกระบวนการลดมิติแบบผสมผสาน NIAs-CCA/DCCA แสดงดังรูปที่ 3.9 ซึ่งจะแตกต่างจากวิธี CCA/DCCA-NIAs เพียงเล็กน้อย คือ การสลับลำดับของกระบวนการ โดยจะเริ่มจากการเลือกเซตคุณลักษณะที่เหมาะสมของแต่ละเซตคุณลักษณะด้วยอัลกอริทึมการค้นหาคำตอบ NIAs หลังจากนั้นแต่ละเซตคุณลักษณะที่ถูกเลือกจะถูกหลอมรวมกันด้วยวิธีการแปลงคุณลักษณะ CCA/DCCA โดยแต่ละอัลกอริทึมที่นำเสนอแสดงรายละเอียดดังนี้

1. วิธีการลดมิติแบบผสมผสาน CCA/DCCA-NIAs

จากรูปที่ 3.7 แสดงกระบวนการลดมิติแบบผสมผสาน CCA/DCCA-NIAs เริ่มจากเซตคุณลักษณะที่มาจากแหล่งกำเนิดสัญญาณต่างๆ ถูกหลอมรวมคุณลักษณะเข้าด้วยกันด้วยวิธี CCA หรือ DCCA

จากนั้นจะถูกเลือกคุณลักษณะด้วยวิธีการค้นหาคำตอบ NIAs ต่างๆ ได้แก่ BPSO, BCS, EBCS, BGWO และ DisBGWO หลังจากได้เซตคุณลักษณะแล้วจะเข้าสู่กระบวนการประเมินความเหมาะสมด้วยสมการ 3.2 ซึ่งเน้นค่าความถูกต้องแม่นยำของการจำแนกข้อมูลและจำนวนของคุณลักษณะที่ถูกเลือก อัลกอริทึมการลดมิติแบบผสมผสาน CCA/DCCA-NIAs ที่นำเสนอ ประกอบด้วย 10 อัลกอริทึม ดังนี้ CCA-BPSO, CCA-BCS, CCA-EBCS, CCA-BGWO, CCA-DisBGWO, DCCA-BPSO, DCCA-BCS, DCCA-EBCS, DCCA-BGWO และ DCCA-DisBGWO



รูปที่ 3.7 กระบวนการลดมิติแบบผสมผสานของที่นำเสนอ CCA/DCCA-NIAs

อัลกอริทึมการลดมิติแบบผสมผสาน CCA/DCCA-NIAs เริ่มต้นด้วยคุณลักษณะที่ได้จากสัญญาณต่างๆ ในขั้นตอนที่ 3.1.1 ถูกนำมาลดมิติข้อมูล โดยเริ่มต้นจากกำหนดค่าเริ่มต้นต่างๆ ของอัลกอริทึม NIAs ได้แก่ BPSO, BCS, EBCS, BGWO และ DisBGWO ตามอัลกอริทึมหัวข้อ 3.1.2.2 ข้อ 2-6 ตามลำดับ จากนั้นแบ่งเซตคุณลักษณะแบ่งออกเป็น 2 เซต คือ เซตคุณลักษณะสำหรับสร้างโมเดล และเซตคุณลักษณะสำหรับการสอบทวน และสุ่มค่าของคำตอบของวิธีการค้นหาเริ่มต้น แต่ละรอบของการคำนวณแต่ละเซตของคุณลักษณะถูกห่อรวมเข้าด้วยกัน โดยดำเนินการทีละคู่ด้วยวิธี CCA หรือ DCCA ตามหัวข้อ 3.1.2.1 รวมทั้งหมด $K(K-1)/(2 \times I)$ คู่ โดยที่ K คือ จำนวนของเซตของคุณลักษณะทั้งหมด และทุกคู่ของเซตของคุณลักษณะถูกห่อรวมกัน จากนั้นเข้าสู่กระบวนการเลือกคุณลักษณะด้วยวิธีการ

สุ่ม NIAs ตามอัลกอริทึมต่างๆ (BPSO, BCS, EBCS, BGWO, และ DisBGWO) ตามรูปที่ 3.3-3.6 และเซตคุณลักษณะถูกประเมินฟังก์ชันความเหมาะสมตามสมการ 3.2 จนครบรอบการคำนวณ ผลลัพธ์ที่ได้คือเซตของคุณลักษณะที่มีค่าฟังก์ชันความเหมาะสมที่ดีที่สุด และค่าฟังก์ชันความเหมาะสมในแต่ละรอบของการคำนวณ รายละเอียดอัลกอริทึม CCA/DCCA-NIAs ที่นำเสนอแสดงดังรูปที่ 3.8

อัลกอริทึม CCA/DCCA-NIAs

Input : T : maximum iteration, D : Dimensionality of search space (total number of features)

Output: The optimal position, training and validation accuracies of the selected features

Begin

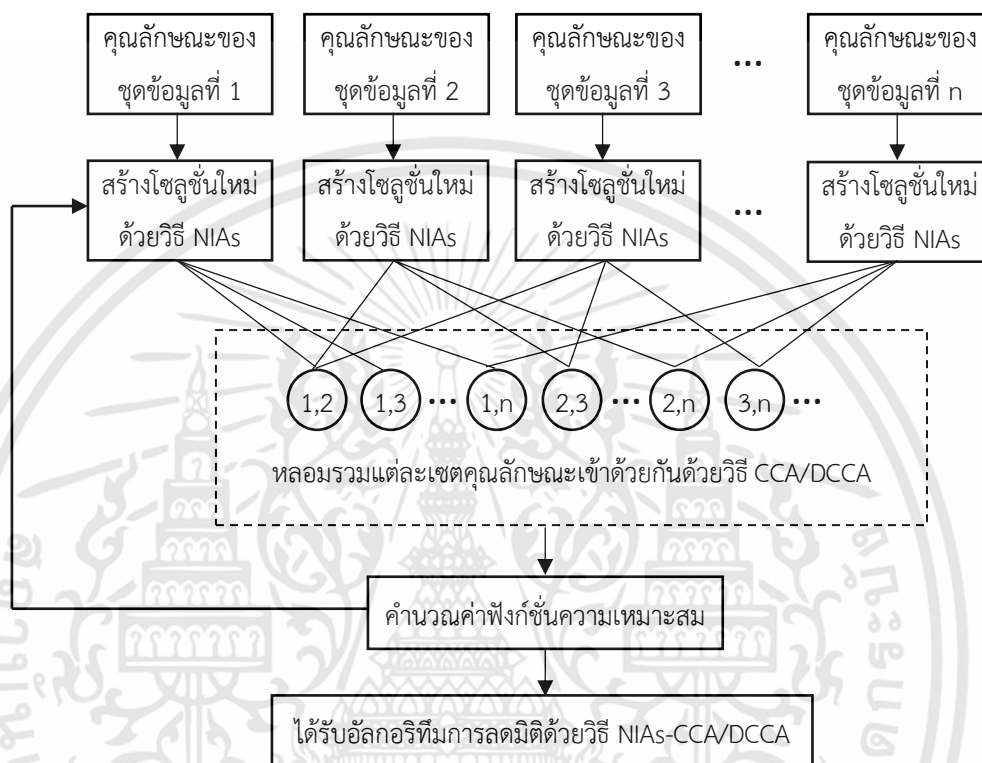
- 1 divide dataset into a training set and a validation set;
 - 2 randomly initialize the position each dimension $\{0,1\}$;
 - 3 for $t=1$ to T do
 - 4 match datasets by the combination techniques $K(K-1)/2$, K = the total of different datasets;
 - 5 for $k=1$ to $K(K-1)/2$ each pairwise i.e. (X_1, X_2)
 - 6 // **For CCA technique.**
calculate eigenvalues and eigenvectors of X_1 and X_2 by Equation 2.32;
 - 6 // **For DCCA technique.**
calculate eigenvalues and eigenvectors of X_1 and X_2 by Equation 2.44;
 - 7 obtain eigenvalues as where $\Omega_{X1}=[w_{X1}^1, \dots, w_{X1}^d]$ and $\Omega_{X2}=[w_{X2}^1, \dots, w_{X2}^d]$, d is reduced dimension;
 - end
 - 8 fuse all feature sets by the parallel combination; $Z_k = \frac{1}{K-1} \sum_{m=1}^K \Omega_{X_m}^T X_m$, where Z_k is the feature fusion of multi-datasets;
 - 9 // **For CCA/DCCA-BPSO algorithm**
select optimal feature using Line 4-13 of BPSO algorithm as shown in Fig.3.2;
 - // **For CCA/DCCA-BCS algorithm**
select the feature fusion using Line 7-15 of BCS algorithm as shown in Fig. 3.3;
 - // **For CCA/DCCA-EBCS algorithm**
select the feature fusion using Line 4-5 and 7-27 of EBCS algorithm as shown in Fig. 3.4;
 - // **For CCA/DCCA-BGWO algorithm**
select the feature fusion using Line 3-4 and 6-11 of BGWO algorithm as shown in Fig. 3.5;
 - // **For CCA/DCCA-DisBGWO algorithm**
select the feature fusion using Line 3-4 and 6-15 of DisBGWO algorithm as shown in Fig.3.6;
 - 10 evaluate fitness function of each solution
 - 11 return the optimal position and training and validation accuracies of the selected features;
-

รูปที่ 3.8 อัลกอริทึม CCA/DCCA-NIAs

2. วิธีการลดมิติแบบผสมผสาน NIAs-CCA/DCCA

การกระบวนกรลดมิติแบบผสมผสานของอัลกอริทึมที่นำเสนอ NIAs-CCA/DCCA แสดงดังรูป 3.9 เริ่มจากเซตคุณลักษณะที่มาจากแหล่งกำเนิดสัญญาณต่างๆ ถูกเลือกคุณลักษณะด้วยแต่ละวิธี NIAs และเซตของคุณลักษณะที่ถูกเลือกของแต่ละแหล่งกำเนิดต่างๆ ถูกหลอมรวมคุณลักษณะเข้าด้วยกันด้วยวิธีการแปลง CCA หรือ DCCA และคุณลักษณะที่หลอมรวมแล้วจะเข้าสู่กระบวนการประเมินฟังก์ชัน

ความเหมาะสมเพื่อหาเซตคุณลักษณะที่เหมาะสมที่สุด และฟังก์ชันการแปลงของ CCA หรือ DCCA ของเซตคุณลักษณะที่เหมาะสมที่สุด อัลกอริทึมการลดมิติแบบผสมผสาน NIAs-CCA/DCCA ที่นำเสนอประกอบด้วย 10 อัลกอริทึม ดังนี้ BPSO-CCA, BCS-CCA, EBCS-CCA, BGWO-CCA, DisBGWO-CCA, BPSO-DCCA, BCS -DCCA, EBCS-DCCA, BGWO-DCCA และ DisBGWO-DCCA



รูปที่ 3.9 กระบวนการลดมิติแบบผสมผสานของที่นำเสนอ NIAs-CCA/DCCA

อัลกอริทึมการลดมิติแบบผสมผสาน NIAs-CCA/DCCA เริ่มต้นด้วยคุณลักษณะที่ได้จากสัญญาณต่างๆ ในขั้นตอนที่ 3.1.1 ถูกนำมาลดมิติข้อมูล โดยเริ่มต้นจากกำหนดค่าเริ่มต้นต่างๆ ของอัลกอริทึม NIAs ได้แก่ BPSO, BCS, EBCS, BGWO และ DisBGWO ตามอัลกอริทึมหัวข้อ 3.1.2.2 ข้อ 2-6 ตามลำดับ จากนั้นแบ่งแต่ละเซตคุณลักษณะแบ่งออกเป็น 2 เซต คือ เซตคุณลักษณะสำหรับสร้างโมเดล และเซตคุณลักษณะสำหรับการสอบทาน และสุ่มค่าเพื่อหาค่าตอบเริ่มต้น แต่ละรอบของการคำนวณคุณลักษณะของแต่ละแหล่งที่มาถูกเลือกเซตคุณลักษณะที่เหมาะสมด้วยแต่ละอัลกอริทึมการค้นหาค่าตอบแบบสุ่มด้วย NIAs ได้แก่ BPSO, BCS, EBCS, BGWO และ DisBGWO ดังรูปที่ 3.3-3.6 เมื่อได้เซตคุณลักษณะที่เหมาะสมมาแล้วแต่ละเซตคุณลักษณะเหล่านั้นถูกลอมรวมคุณลักษณะ โดยดำเนินการที่ละคู่ด้วยวิธี CCA หรือ DCCA ตามหัวข้อ 3.1.2.1 รวมทั้งหมด $K(K-1)/(2 \times I)$ คู่ จากนั้นจะถูกประเมินฟังก์ชันความเหมาะสมตามสมการ 3.2 จนครบรอบการคำนวณ ผลลัพธ์ที่ได้ คือ เซตคุณลักษณะที่มีค่าฟังก์ชันความ

เหมาะสมที่สุด และค่าฟังก์ชันความเหมาะสมในแต่ละรอบของการคำนวณ อัลกอริทึม CCA/DCCA-NIAs ที่นำเสนอแสดงดังรูปที่ 3.10

อัลกอริทึม NIAs-CCA/DCCA

Input : T : maximum iteration, D : Dimensionality of search space (total number of features)

Output: The optimal position, training and validation accuracies of the selected features

Begin

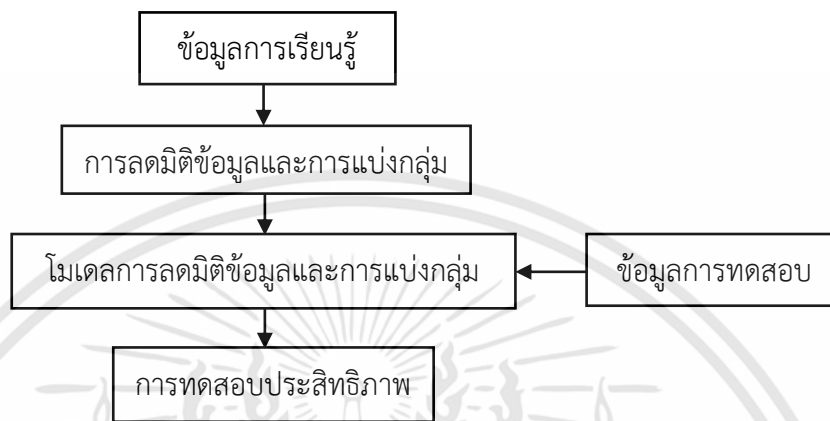
- 1 divide dataset into a training set and a validation set;
 - 2 randomly initialize the position each dimension $\{0,1\}$;
 - 3 for $t=1$ to T do
 - 4 // **For CCA/DCCA-BPSO algorithm**
select optimal feature using Line 4-12 of BPSO algorithm as shown in Fig.3.2;
 - 4 // **For CCA/DCCA-BCS algorithm**
select the feature fusion using Line 7-14 of BCS algorithm as shown in Fig. 3.3;
 - 4 // **For CCA/DCCA-EBCS algorithm**
select the feature fusion using Line 4-5 and 7-24 of EBCS algorithm as shown in Fig. 3.4;
 - 4 // **For CCA/DCCA-BGWO algorithm**
select the feature fusion using Line 3-4 and 6-10 of BGWO algorithm as shown in Fig. 3.5;
 - 4 // **For CCA/DCCA-DisBGWO algorithm**
select the feature fusion using Line 3-4 and 6-14 of DisBGWO algorithm as shown in Fig. 3.6;
 - 5 match selected feature sets by the combination techniques $K(K-1)/2$, K = the total of Different datasets;
 - 6 for $k=1$ to $K(K-1)/2$ each pairwise i.e. (X_1, X_2)
 - 7 // **For CCA technique**
calculate eigenvalues and eigenvectors of X_1 and X_2 by Equation 2.32;
 - 7 // **For DCCA technique**
calculate eigenvalues and eigenvectors of X_1 and X_2 by Equation 2.44;
 - 8 obtain eigenvalues as where $\Omega_{X1}=[w_{X1}^1, \dots, w_{X1}^d]$ and $\Omega_{X2}=[w_{X2}^1, \dots, w_{X2}^d]$, d is reduced dimension;
end
 - 9 fuse all feature sets as parallel combination; $Z_K = \frac{1}{K-1} \sum_{k=1}^K \Omega_{X_m}^T X_m$, where Z_K is the feature fusion of multi-datasets;
 - 10 evaluate fitness;
end
 - 11 return the optimal position and training and validation accuracies of the selected features;
-

รูปที่ 3.10 อัลกอริทึม NIAs-CCA/DCCA สำหรับการเลือกคุณลักษณะ

3.3 ขั้นตอนการประมวลผล

กระบวนการประมวลผลสำหรับงานวิจัยนี้ประกอบด้วยกระบวนการแบ่งกลุ่มข้อมูลดังรูปที่ 3.11 ข้อมูลทั้งหมดถูกแบ่งออกเป็น 2 ส่วน คือข้อมูลสำหรับการเรียนรู้เพื่อสร้างโมเดล และข้อมูลสำหรับการทดสอบเพื่อประเมินประสิทธิภาพของของโมเดล ข้อมูลสำหรับการเรียนรู้ถูกนำไปใช้เป็นข้อมูลสำหรับการดำเนินการในหัวข้อที่ 3.2 และเข้าสู่กระบวนการแบ่งกลุ่มข้อมูลด้วยอัลกอริทึมการแบ่งกลุ่มข้อมูล

ประกอบด้วย kNN, DT, SVM และ RF ตามหัวข้อที่ 2.10 โดยผลลัพธ์ที่ได้ประกอบด้วย โมเดลการลดมิติข้อมูล ข้อมูลสำหรับการทดสอบเป็นข้อมูลที่ไม่เคยถูกนำมาใช้มาก่อนถูกนำไปใช้กับโมเดลที่ได้จากการดำเนินการในหัวข้อ 3.2 เพื่อทดสอบประสิทธิภาพของโมเดล



รูปที่ 3.11 ขั้นตอนการประมวลผล

3.3.1 การประเมินประสิทธิภาพของอัลกอริทึม

3.3.1.1 ค่าเฉลี่ยความถูกต้องแม่นยำ (Average of Accuracy) คืออัตราส่วนของจำนวนทั้งหมดของคำตอบที่โมเดลตอบถูกและจำนวนคำตอบทั้งหมดสามารถอธิบายได้ตามสมการดังต่อไปนี้

$$ArgAcc = \frac{1}{K} \sum_{k=1}^K \left(\frac{TP + TN}{TP + FP + FN + TN} \times 100\% \right) \quad (3.7)$$

โดยที่ $ArgAcc$ คือ ค่าเฉลี่ยความถูกต้องแม่นยำของอัลกอริทึมทั้งหมด K รอบการประเมินประสิทธิภาพ TP คือ จำนวนที่ตอบถูกในกลุ่มที่สนใจ TN คือ จำนวนที่ตอบถูกในกลุ่มอื่นๆ FP คือ จำนวนที่ตอบผิดในกลุ่มที่สนใจ และ FN คือ จำนวนที่ตอบผิดในกลุ่มอื่นๆ

3.3.1.2 ค่าเฉลี่ยความไว (Average of Sensitivity) คืออัตราส่วนของจำนวนคำตอบที่โมเดลตอบถูกในกลุ่มที่สนใจ และจำนวนคำตอบของกลุ่มที่สนใจทั้งหมด สามารถอธิบายได้ตามสมการดังต่อไปนี้

$$ArgSent = \frac{1}{K} \sum_{k=1}^K \left(\frac{TP}{TP + FN} \times 100\% \right) \quad (3.8)$$

โดยที่ $ArgSent$ คือ ค่าเฉลี่ยความไวของอัลกอริทึมของอัลกอริทึมทั้งหมด K รอบการประเมินประสิทธิภาพ

3.3.1.3 ค่าเฉลี่ยความจำเพาะ (Average of Specificity) คืออัตราส่วนของจำนวนคำตอบที่

โมเดลตอบถูกในกลุ่มอื่นๆ และจำนวนคำตอบของกลุ่มอื่นๆทั้งหมด สามารถอธิบายได้ตามสมการดังต่อไปนี้

$$ArgSpec = \frac{1}{K} \sum_{k=1}^K \left(\frac{TN}{TN + FP} \times 100\% \right) \quad (3.9)$$

โดยที่ $ArgSpec$ คือ ค่าเฉลี่ยความจำเพาะของอัลกอริทึมของอัลกอริทึมทั้งหมด K รอบการประเมินประสิทธิภาพ

3.3.1.4 ค่าความแปรปรวนของผลการประเมิน คือ ค่าความแปรปรวนของการประเมินประสิทธิภาพต่างๆ ประกอบด้วย ค่าความถูกต้องแม่นยำ ค่าความไว และค่าความจำเพาะ ทั้งหมด K รอบ ซึ่งจะเป็นเครื่องที่ใช้เพื่อบ่งบอกถึงความเสถียรภาพของอัลกอริทึม สามารถหาได้คำนวณได้จากสมการ

$$Std = \sqrt{\frac{1}{K} \sum_{k=1}^K (p_k - mean(p))^2} \quad (3.10)$$

โดยที่ Std คือค่าความแปรปรวนของการประเมินทั้งหมด K รอบ p_k ค่าผลของการประเมินแต่ละรอบที่ k และ $mean(p)$ คือค่าเฉลี่ยของผลการประเมินทั้งหมด

บทที่ 4

วิธีการทดลอง ผลการทดลอง และวิเคราะห์ผลการทดลอง

การทดลองการวิเคราะห์การนอนอัตโนมัติของงานวิจัยนี้เราได้แบ่งออกเป็น 2 การทดลอง ได้แก่ การทดลองการจำแนกระยะการนอนแบบอัตโนมัติ และการทดลองการจำแนกภาวะหยุดหายใจขณะนอนหลับอัตโนมัติ โดยมีรายละเอียดดังนี้

4.1 การทดลองการจำแนกระยะการนอนแบบอัตโนมัติ

กระบวนการทดลองการจำแนกระยะการนอนแบบอัตโนมัติ ประกอบด้วย กระบวนการก่อนการประมวลผล ได้แก่ ขั้นตอนการเตรียมข้อมูล ขั้นตอนการประมวลผลสัญญาณ ขั้นตอนการสกัดคุณลักษณะ และขั้นตอนการลดมิติข้อมูล และกระบวนการประมวลผลสำหรับการจำแนกระยะการนอน ซึ่งเป็นไปตามขั้นตอนการออกแบบในบทที่ 3 การทดลองของงานวิจัยนี้ได้มุ่งเน้นเพื่อการสร้าง ออกแบบ ทดลองและเปรียบเทียบอัลกอริทึมต่างๆ เพื่อให้ได้อัลกอริทึมที่เหมาะสมที่สุดที่สามารถเพิ่มประสิทธิภาพสำหรับวิเคราะห์การนอนแบบอัตโนมัติ โดยเฉพาะอย่างยิ่งมุ่งเน้นการปรับปรุงประสิทธิภาพด้านความถูกต้องแม่นยำของการจำแนกระยะการนอนและลดเวลาในการประมวลผล

4.1.1 ขั้นตอนการเตรียมข้อมูล

งานวิจัยนี้ได้วิเคราะห์ข้อมูลการนอนด้วยการวัดสัญญาณ PSG ประกอบด้วย สัญญาณ EEG จำนวน 2 ช่องสัญญาณ ณ ตำแหน่ง C3-A1 และ C4-A2 ตามมาตรฐานตำแหน่งอิเล็กโทรด 10-20, สัญญาณ EOG จำนวน 2 ช่องสัญญาณที่ได้จากตำแหน่ง LOC และ ROC และสัญญาณ ECG จากตำแหน่ง Lead II ด้วยความถี่สุ่ม 125 เฮิร์ตซ์ ซึ่งสัญญาณ PSG ที่ใช้สำหรับการทดลองนี้มาจากฐานข้อมูล Sleep Heart Health Study (SHHS) จำนวน 100 ชุดข้อมูล [12] แต่ละระยะการนอน และถูกกำหนดจากผู้เชี่ยวชาญตามมาตรฐาน AASM ขั้นตอนการเตรียมสัญญาณเริ่มต้นด้วยแต่ละช่องสัญญาณของสัญญาณ EEG, สัญญาณ EOG และสัญญาณ ECG ถูกแบ่งออกเป็นช่วงๆ ช่วงละ 30 วินาที ซึ่งแบ่งตามการกำหนดช่วงเวลา (Epoch) ระยะการนอนตามมาตรฐาน AASM จำนวนตัวอย่างของแต่ละระยะการนอนของฐานข้อมูล SHHS สามารถแสดงได้ดังตารางที่ 4.1 การทดลองทั้งหมดเราได้ดำเนินการทดลองโดยใช้โปรแกรม MATLAB และใช้บางไลบรารีของโปรแกรม และประมวลผลด้วยเครื่องคอมพิวเตอร์ส่วนบุคคล (Intel Core i7-9700K CPU, 4.7 GHz, with 7.89 Gb. บน Windows 8.1)

ตารางที่ 4.1 จำนวนตัวอย่างของระยะ Wake,ระยะ NREM1-3 และระยะ REM จากฐานข้อมูล SHHS

แหล่งข้อมูล	ระยะ Wake	ระยะ REM	ระยะ NREM1	ระยะ NREM2	ระยะ NREM3
SHHS	18,061	16,231	13,593	14,369	20,498

ก่อนเข้าสู่กระบวนการสกัดคุณลักษณะ แต่ละสัญญาณจะถูกประมวลผลสัญญาณก่อนเพื่อกรองสัญญาณรบกวน สัญญาณ EEG ผ่านกระบวนการกรองความถี่ด้วย Butterworth filter ที่ย่านความถี่ช่วง 0.3-30 เฮิรตซ์ สัญญาณ EOG ผ่านกระบวนการกรองความถี่ย่านความถี่ช่วง 0.3-10 เฮิรตซ์ และสัญญาณ ECG ผ่านกระบวนการกรองความถี่ Band pass filter ที่ย่านความถี่ช่วง 5-50 เฮิรตซ์ จากนั้นแต่ละสัญญาณถูกสกัดคุณลักษณะตามรายละเอียดหัวข้อที่ 3.1.1 จำนวนคุณลักษณะรวมทั้งสิ้น 364 คุณลักษณะ ได้แก่ 2 ช่องสัญญาณของ EEG จำนวน $81 \times 2 = 162$ คุณลักษณะ, 2 ช่องสัญญาณของ EOG จำนวน $73 \times 2 = 146$ คุณลักษณะ และ 1 ช่องสัญญาณของ ECG ซึ่งเป็นคุณลักษณะที่ได้จากการวิเคราะห์ HRV จำนวน $56 \times 1 = 56$ คุณลักษณะ

4.1.2 การทดลองและผลการทดลองการทดสอบประสิทธิภาพของวิธีการลดมิติข้อมูลแบบเดี่ยวสำหรับการจำแนกระยะการนอนด้วยหลายชุดสัญญาณ

ขั้นตอนการลดมิติข้อมูลแบบเดี่ยวคือ การลดมิติข้อมูลด้วยวิธีการแปลงคุณลักษณะหรือ การลดมิติข้อมูลด้วยวิธีการเลือกคุณลักษณะเพียงอย่างใดอย่างหนึ่ง ดังนั้นขั้นตอนการลดมิติของการทดลองในงานวิจัยนี้ได้นำวิธีการแปลงคุณลักษณะด้วยวิธี PCA, LDA, CCA และ DCCA มาทดสอบประสิทธิภาพและนำมาประยุกต์ใช้สำหรับการจำแนกระยะการนอน และได้นำวิธีการเลือกคุณลักษณะด้วยวิธี NIAs ได้แก่ BPSO, BCS, EBCS และ BGWO มาทดสอบประสิทธิภาพและนำมาเปรียบเทียบกับประสิทธิภาพของวิธีการเลือกคุณลักษณะที่งานวิจัยนี้ได้นำเสนอ DisBGWO สำหรับประยุกต์ใช้ในการจำแนกระยะการนอน

4.1.2.1 การทดลองและผลการทดลองของการทดสอบประสิทธิภาพของวิธีการลดมิติข้อมูลด้วยการแปลงคุณลักษณะ

เซตของคุณลักษณะที่ได้จากการทดลองในหัวข้อที่ 4.1.1 ถูกนำมาใช้ในการทดลอง เซตของคุณลักษณะทั้งหมดจะถูกสุ่มแบ่งเป็น 2 เซตข้อมูล คือ เซตข้อมูลสำหรับการเรียนรู้ และเซตของข้อมูลสำหรับการทดสอบ ขั้นตอนการทดลองประกอบด้วย 2 ขั้นตอนหลัก คือ ขั้นตอนการเรียนรู้ เริ่มต้นด้วยแต่ละเซตของคุณลักษณะถูกหลอมรวมด้วยวิธีการแปลง PCA, LDA, CCA และ DCCA ตามหัวข้อ 3.1.2.1 ข้อ 1-4 ตามลำดับ จากนั้นเข้าสู่กระบวนการจำแนกข้อมูล สำหรับขั้นตอนการทดสอบนั้น เซตข้อมูลสำหรับการทดสอบจะถูกใช้เพื่อทดสอบประสิทธิภาพและเปรียบเทียบวิธีต่างๆ สิ่งที่พิจารณาความ

เหมาะสมของวิธีการเหล่านั้น คือ ค่าความประสิทธิภาพความถูกต้องแม่นยำของการจำแนกกระยะการนอน และจำนวนของคุณลักษณะที่ลดลง

1. การทดลองที่ 1 การทดลองเปรียบเทียบประสิทธิภาพของวิธีการลดมิติด้วยการแปลงคุณลักษณะสำหรับการจำแนกกระยะการนอนโดยใช้สัญญาณ EEG EOG และ ECG

จุดประสงค์การทดลอง เพื่อเปรียบเทียบวิธีการลดมิติด้วยการแปลงคุณลักษณะวิธี PCA, LDA, CCA แบบอนุกรมและแบบขนาน และ DCCA แบบอนุกรมและแบบขนาน และการทดลองใช้อัลกอริทึมการจำแนกข้อมูลต่างๆ ได้แก่ KNN, SVM, Decision tree และ Random forest เพื่อเปรียบเทียบและเลือกอัลกอริทึมการจำแนกข้อมูลที่เหมาะสมที่สุดสำหรับเป็นเงื่อนไขในการทดลอง

เงื่อนไขการทดลอง

1. เซตของคุณลักษณะทั้งหมดถูกสุ่มแบ่งด้วยวิธี Cross validation ที่ $k=5$ เซตคุณลักษณะแบ่งเป็น 2 ส่วน คือ เซตข้อมูลสำหรับการเรียนรู้ และเซตของข้อมูลสำหรับการทดสอบ
2. พารามิเตอร์สำหรับการกำหนดอัลกอริทึมการจำแนกข้อมูลของแต่ละอัลกอริทึมแสดงได้ดังตารางที่ 4.2

ตารางที่ 4.2 พารามิเตอร์ที่สำคัญสำหรับอัลกอริทึมการจำแนกข้อมูลต่างๆ

อัลกอริทึมการจำแนกข้อมูล	พารามิเตอร์ต่างๆ
kNN	จำนวนของข้อมูลที่ถูกนำมาพิจารณา $k=5$
Decision tree	อัลกอริทึมการตัดสินใจด้วยวิธี ID3
SVM	ฟังก์ชัน kernel คือ Radial Based Function
Random forest	จำนวนต้นไม้เท่ากับ $n=15$

3. ผลการทดสอบประสิทธิภาพของแต่ละวิธีแสดงในรูปแบบของ ค่าเฉลี่ยของความแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าความจำเพาะ ค่าความแปรปรวน และเวลาในการประมวลผล

2. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการเรียนรู้การทดลองที่ 1

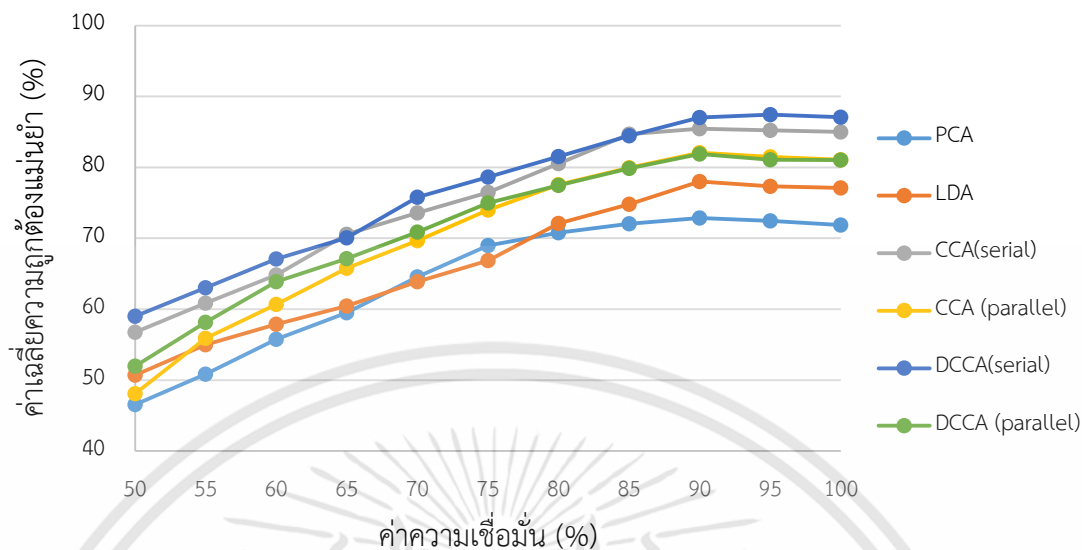
จากวิธีการลดมิติด้วยการแปลงคุณลักษณะ PCA, LDA, CCA และ DCCA ตามหัวข้อที่ 3.1.2.1 จำเป็นต้องกำหนดเงื่อนไขของการลดมิติ ดังนั้นการทดลองนี้จึงได้ทดลองเพื่อหาเงื่อนไขที่เหมาะสมสำหรับการลดมิติข้อมูลด้วยพื้นฐานร้อยละความเชื่อมั่นของคุณลักษณะที่ถูกนำไปใช้ในการประมวลผล จากผลการทดลองตารางที่ 4.3 – 4.7 แสดงค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยวิธี PCA, LDA, CCA แบบอนุกรมและแบบขนาน และ DCCA แบบอนุกรมและแบบขนาน ของการจำแนก

ข้อมูลด้วยอัลกอริทึมต่างๆ ได้แก่ kNN, DT, SVM และ RF นอกจากนี้ผลการทดลองรูปที่ 4.1 - 4.4 แสดงความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ และค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอน

ตารางที่ 4.3 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนกกระยะการนอนด้วยอัลกอริทึม kNN

ความเชื่อมั่น (%)	ค่าเฉลี่ยค่าความถูกต้องแม่นยำ(%±SD)					
	PCA	LDA	CCA (serial)	CCA (parallel)	DCCA (serial)	DCCA (parallel)
50	46.56±0.41	50.70±0.67	56.74±0.52	48.08±0.75	58.98±0.72	51.98±0.67
55	50.86±0.71	54.98±0.77	60.87±0.67	55.87±0.54	63.00±0.78	58.13±0.57
60	55.76±0.42	57.86±0.87	64.84±0.97	60.65±0.43	67.07±0.66	63.89±0.73
65	59.48±0.91	60.44±0.92	70.54±1.21	65.75±0.83	70.07±0.64	67.13±0.66
70	64.54±0.67	63.87±0.65	73.58±0.42	69.65±0.41	75.78±0.35	70.87±0.69
75	68.98±0.71	66.87±0.68	76.45±0.24	73.98±0.72	78.64±0.73	74.99±0.71
80	70.78±0.64	72.07±0.49	80.54±0.31	77.56±0.38	81.54±0.31	77.47±0.68
85	72.06±1.11	74.79±0.68	84.67±0.45	79.96±0.44	84.43±0.52	79.86±0.87
90	72.86±0.97	77.98±0.62	85.45±0.63	82.07±0.47	87.04±0.55	81.88±0.89
95	72.44±0.69	77.31±0.81	85.23±0.61	81.47±0.65	87.45±0.33	81.08±0.56
100	71.86±0.88	77.08±0.72	85.00±0.78	81.05±0.74	87.08±0.64	81.03±0.48

จากนั้นนำผลต่างๆที่ได้จากตารางที่ 4.3 มาพล็อตเป็นกราฟความสัมพันธ์แสดงดังรูปที่ 4.1

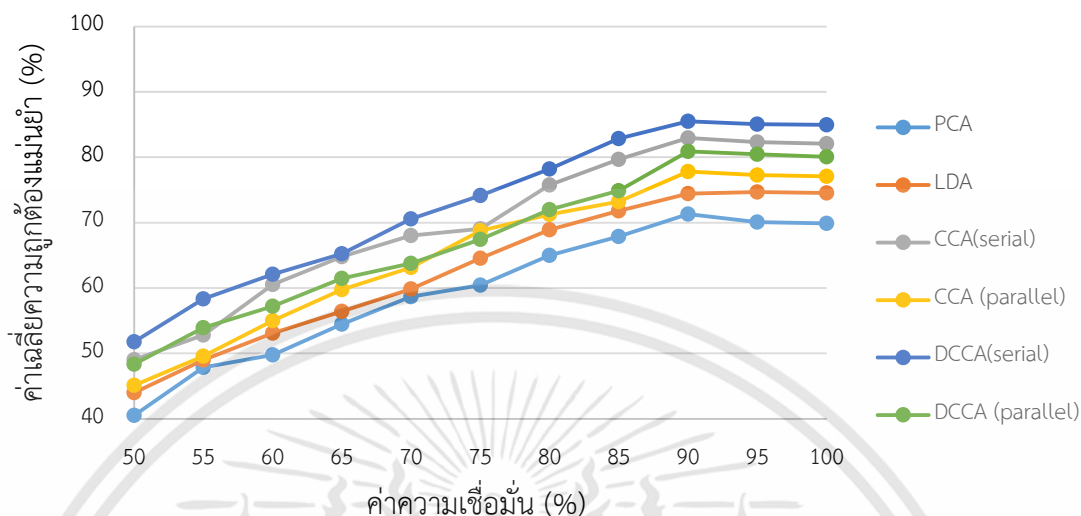


รูปที่ 4.1 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะ และค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม kNN

ตารางที่ 4.4 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนกกระยะการนอนด้วยอัลกอริทึม DT

ค่าความ เชื่อมั่น (%)	ค่าเฉลี่ยค่าความถูกต้องแม่นยำ(%±SD)					
	PCA	LDA	CCA (serial)	CCA (parallel)	DCCA (serial)	DCCA (parallel)
50	40.52±1.22	43.98±0.69	49.00±0.98	45.12±0.87	51.78±0.55	48.32±0.68
55	47.86±0.63	49.02±0.68	52.78±0.97	49.56±0.72	58.35±0.62	53.93±0.71
60	49.76±0.41	53.11±0.74	60.54±0.76	54.98±0.57	62.11±0.76	57.21±0.94
65	54.48±0.55	56.42±1.27	64.77±0.85	59.75±0.61	65.24±0.69	61.44±0.54
70	58.65±0.36	59.85±0.54	68.01±1.23	63.12±0.76	70.56±0.64	63.76±0.54
75	60.44±1.01	64.57±0.65	69.05±0.84	68.75±1.27	74.15±0.76	67.43±0.96
80	65.00±0.97	68.88±0.75	75.78±0.82	71.25±1.03	78.22±0.64	71.97±0.81
85	67.87±0.66	71.77±0.69	79.67±1.08	73.22±0.56	82.87±0.58	74.89±0.76
90	71.28±0.64	74.43±0.78	82.97±0.46	77.82±0.45	85.48±0.58	80.91±0.57
95	70.06±0.99	74.69±0.75	82.33±0.81	77.27±0.65	85.08±0.33	80.45±0.61
100	69.87±0.65	74.52±0.92	82.08±1.19	77.07±0.72	84.97±0.81	80.06±0.45

จากนั้นนำผลต่างๆที่ได้จากตารางที่ 4.4 มาพล็อตเป็นกราฟความสัมพันธ์แสดงดังรูปที่ 4.2

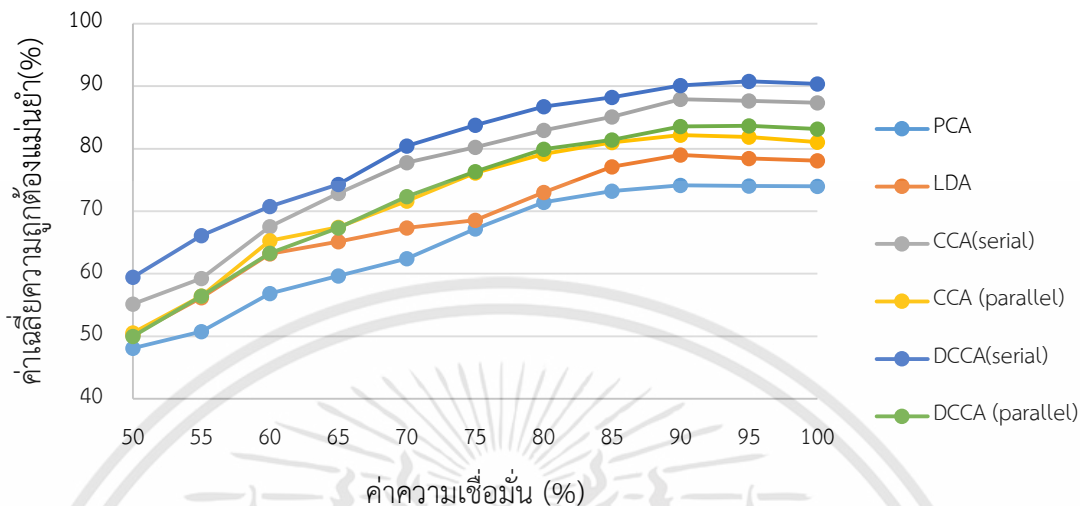


รูปที่ 4.2 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะ และค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม DT

ตารางที่ 4.5 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนกกระยะการนอนด้วยอัลกอริทึม SVM

ค่าความ เชื่อมั่น (%)	ค่าเฉลี่ยค่าความถูกต้องแม่นยำ(%±SD)					
	PCA	LDA	CCA (serial)	CCA (parallel)	DCCA (serial)	DCCA (parallel)
50	48.11±0.67	50.23±0.17	55.17±0.45	50.55±0.34	59.43±0.24	50.01±0.56
55	50.76±0.44	56.20±0.31	59.22±0.81	56.44±0.37	66.10±0.85	56.43±0.18
60	56.85±0.76	63.20±0.74	67.54±1.34	65.31±0.28	70.76±0.40	63.31±0.52
65	59.65±0.43	65.15±0.83	72.86±0.62	67.44±0.31	74.32±1.15	67.32±0.75
70	62.42±1.21	67.33±0.61	77.75±1.02	71.65±0.49	80.42±0.44	72.34±0.75
75	67.20±0.53	68.56±0.83	80.21±0.45	76.12±0.83	83.75±0.64	76.32±0.58
80	71.44±1.12	73.01±0.45	82.96±0.87	79.15±0.33	86.74±0.65	79.90±0.43
85	73.23±0.81	77.11±0.66	85.11±0.61	80.97±0.56	88.21±0.86	81.38±0.65
90	74.12±0.74	78.98±0.56	87.88±1.18	82.15±0.58	90.13±0.66	83.54±0.65
95	74.03±0.54	78.43±0.54	87.66±0.72	81.87±0.62	90.76±0.93	83.67±0.76
100	74.00±0.29	78.10±0.27	87.31±0.77	81.06±0.89	90.36±0.78	83.15±0.51

จากนั้นนำผลต่างๆที่ได้จากตารางที่ 4.5 มาพล็อตเป็นกราฟความสัมพันธ์แสดงดังรูปที่ 4.3

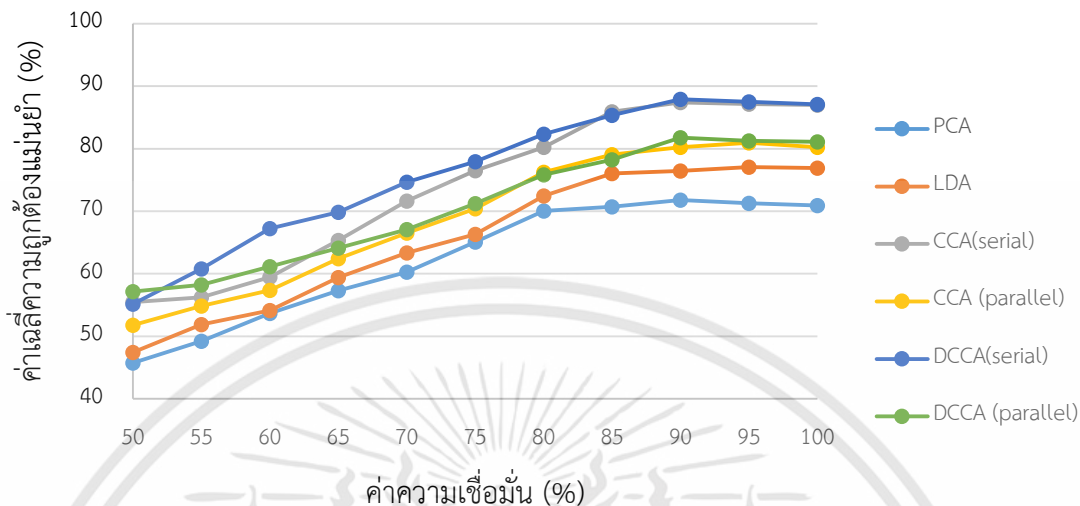


รูปที่ 4.3 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะ และค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม SVM

ตารางที่ 4.6 ค่าเฉลี่ยความถูกต้องแม่นยำที่ได้จากวิธีการลดมิติด้วยการแปลงคุณลักษณะต่างๆ ของการจำแนกกระยะการนอนด้วยอัลกอริทึม RF

ค่าความ เชื่อมั่น (%)	ค่าเฉลี่ยค่าความถูกต้องแม่นยำ(%±SD)					
	PCA	LDA	CCA (serial)	CCA (parallel)	DCCA (serial)	DCCA (parallel)
50	45.75±0.53	47.45±0.75	55.44±0.73	51.76±0.76	55.13±0.62	57.14±0.68
55	49.22±0.67	51.87±0.83	56.25±0.73	54.85±0.72	60.79±0.75	58.21±0.57
60	53.64±0.94	54.11±0.87	59.44±0.57	57.38±0.71	67.22±0.78	61.13±0.57
65	57.32±1.13	59.42±0.96	65.35±0.43	62.44±0.53	69.80±0.42	64.13±0.65
70	60.26±0.54	63.33±0.64	71.64±0.65	66.54±0.51	74.65±0.15	67.07±0.43
75	65.08±0.76	66.31±0.73	76.52±1.15	70.41±0.55	77.93±0.43	71.22±0.81
80	70.06±0.57	72.44±0.75	80.24±0.63	76.23±0.52	82.34±0.82	75.84±0.53
85	70.73±1.06	76.03±0.63	85.91±0.33	79.07±0.86	85.33±0.28	78.24±0.75
90	71.77±0.77	76.46±0.62	87.41±0.94	80.24±1.04	87.92±0.49	81.78±0.67
95	71.29±0.51	77.07±0.64	87.15±0.65	80.96±0.46	87.47±0.47	81.26±0.63
100	70.92±0.84	76.91±0.38	87.11±0.78	80.22±0.88	87.09±0.27	81.12±0.68

จากนั้นนำผลต่างๆที่ได้จากตารางที่ 4.6 มาพล็อตเป็นกราฟความสัมพันธ์แสดงดังรูปที่ 4.4



รูปที่ 4.4 ความสัมพันธ์ระหว่างค่าความเชื่อมั่นต่างๆ ของแต่ละวิธีการลดมิติด้วยการแปลงคุณลักษณะ และค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม RF

การเปรียบเทียบของผลการทดลองจากตารางที่ 4.3 และรูปที่ 4.1 จะเห็นได้ว่าวิธีการลดมิติด้วยวิธี PCA, LDA, CCA แบบอนุกรมและแบบขนาน และ DCCA แบบขนาน แสดงค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม kNN สูงสุดบนพื้นฐานความเชื่อมั่นร้อยละ 90 ขณะที่บนพื้นฐานความเชื่อมั่นร้อยละ 95 และ 100 มีค่าลดต่ำลง ดังนั้นจะเห็นได้ว่าที่ความเชื่อมั่นร้อยละ 100 หรือจำนวนคุณลักษณะทั้งหมดไม่ได้ให้ผลความถูกต้องแม่นยำในการการจำแนกสูงสุดเสมอไป และเป็นตัวอย่างของการเกิดปรากฏการณ์ the curse of dimension ซึ่งจึงจำเป็นอย่างยิ่งที่จะต้องทดลองหาความเชื่อมั่นที่เหมาะสม

จากผลการทดลองตารางที่ 4.4 และรูปที่ 4.2 แสดงให้เห็นว่าการลดมิติด้วยวิธี PCA, CCA แบบอนุกรมและแบบขนาน และ DCCA แบบอนุกรมและแบบขนาน แสดงค่าเฉลี่ยของค่าความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม DT สูงสุดบนพื้นฐานความเชื่อมั่นร้อยละ 90 ขณะที่บนพื้นฐานความเชื่อมั่นที่ร้อยละ 95 และ 100 มีค่าลดต่ำลง การลดมิติด้วยวิธี LDA แสดงค่าความถูกต้องแม่นยำสูงสุดบนพื้นฐานความเชื่อมั่นร้อยละ 95 และลดลงที่ค่าความเชื่อมั่นที่ 100

จากผลการทดลองตารางที่ 4.5 และรูปที่ 4.3 วิธีการลดมิติด้วยวิธี PCA, LDA และ CCA แบบอนุกรมและแบบขนาน แสดงค่าเฉลี่ยของค่าความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึม SVM สูงสุดบนพื้นฐานความเชื่อมั่นร้อยละ 90 ขณะที่ความเชื่อมั่นร้อยละ 95 และ 100 มีค่าลดต่ำลง สำหรับการลดมิติด้วยวิธี DCCA แบบอนุกรมและแบบขนานมีค่าความถูกต้องแม่นยำสูงสุดบนพื้นฐานความเชื่อมั่นร้อยละ 95 และลดลงที่ความเชื่อมั่นที่ 100

จากผลการทดลองตารางที่ 4.6 และรูป 4.4 วิธีการลดมิติด้วยวิธี PCA, CCA แบบอนุกรม และ DCCA แบบอนุกรมและแบบขนาน แสดงค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วย อัลกอริทึม RF สูงสุดบนพื้นฐานความเชื่อมั่นร้อยละ 90 ขณะที่ความเชื่อมั่นร้อยละ 95 และ 100 มีค่าลดต่ำลง สำหรับวิธีการลดมิติ LDA และ CCA แบบขนาน แสดงค่าความถูกต้องแม่นยำสูงสุดบนพื้นฐานค่าความเชื่อมั่นร้อยละ 95 และลดลงเมื่อค่าความเชื่อมั่นที่ 100

จากการทดลองการทั้งหมดจะเห็นได้ว่าค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนด้วยอัลกอริทึมการจำแนกข้อมูลต่างๆ ส่วนใหญ่แสดงค่าสูงสุดบนพื้นฐานความเชื่อมั่นที่ร้อยละ 90 ดังนั้นเราจึงกำหนดเงื่อนไขการลดมิติของวิธี PCA, LDA, CCA และ DCCA ด้วยพื้นฐานความเชื่อมั่นร้อยละ 90 นอกจากนี้ผลการทดลองแสดงให้เห็นว่าการใช้อัลกอริทึมสำหรับการจำแนกกระยะการนอนที่ต่างกัน มีผลกระทบต่อค่าความถูกต้องแม่นยำ จากผลการทดลองแสดงให้เห็นว่าค่าความถูกต้องแม่นยำการจำแนกกระยะการนอนด้วยอัลกอริทึม SVM แสดงค่าความถูกต้องแม่นยำสูงสุดเมื่อเปรียบเทียบกับอัลกอริทึมอื่นๆ

3. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการทดสอบของการทดลองที่ 1

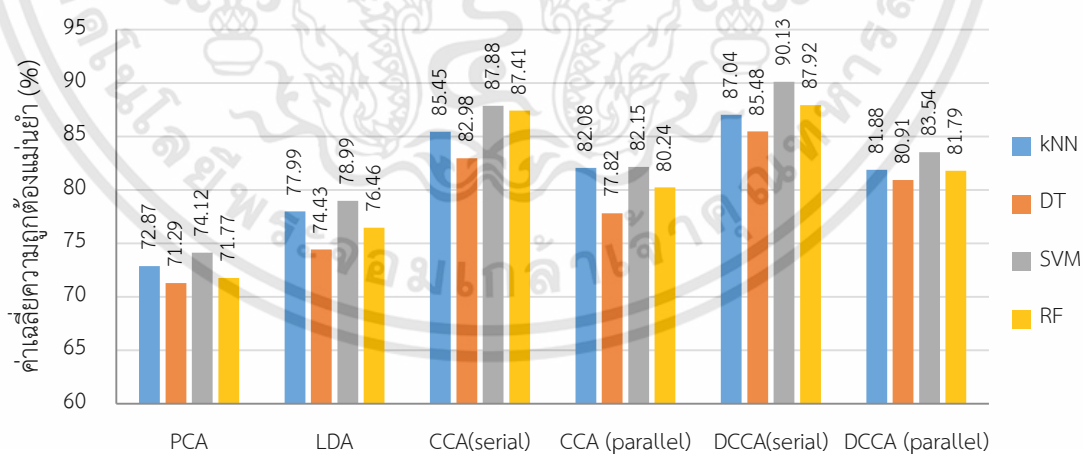
จากผลการทดลองข้างต้นเราได้กำหนดเงื่อนไขการลดมิติของแต่ละวิธีด้วยพื้นฐานความเชื่อมั่นที่ร้อยละ 90 ดังนั้นเราจึงนำเงื่อนไขดังกล่าวมาใช้เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมต่างๆ แสดงดังรูปที่ 4.7 ซึ่งแสดงด้วยค่าเฉลี่ยความถูกต้องแม่นยำ (ArgAcc.) ค่าเฉลี่ยของค่าความไว (ArgSent.) ค่าเฉลี่ยของค่าจำเพาะ (AccSpec.) และค่าเวลาในการประมวลผล

ตารางที่ 4.7 ประสิทธิภาพการจำแนกกระยะการนอนด้วยวิธีการแปลงคุณลักษณะ PCA, LDA, CCA, DCCA และอัลกอริทึมการจำแนกข้อมูล kNN, DT, SVM, และ RF

อัลกอริทึมการจำแนกข้อมูล	วิธีการหลอมรวมคุณลักษณะ	ArgAcc. (%±SD)	ArgSent. (%±SD)	ArgSpec. (%±SD)	Time (sec.)
kNN	PCA	72.8683±0.97	70.4273±0.75	74.7556±0.89	27.5695±0.55
	LDA	77.9856±0.62	74.3237±0.56	80.0643±1.24	28.6604±0.43
	CCA(serial)	85.4532±0.63	81.9643±0.42	87.7553±0.59	33.3702±0.55
	CCA(parallel)	82.0761±0.47	78.0753±1.11	84.6762±0.62	33.2128±0.42
	DCCA(serial)	87.0404±0.55	83.6433±0.71	90.4533±0.76	34.3401±0.68
	DCCA(parallel)	81.8763±0.89	77.6563±0.55	84.9576±0.74	34.8960±0.59
DT	PCA	71.2872±0.64	66.9327±0.52	73.7641±0.69	23.7382±0.43
	LDA	74.4324±0.78	71.9432±0.41	76.1134±0.74	24.5070±0.74
	CCA(serial)	82.9756±0.56	79.8743±0.63	84.4322±0.45	26.4394±0.78

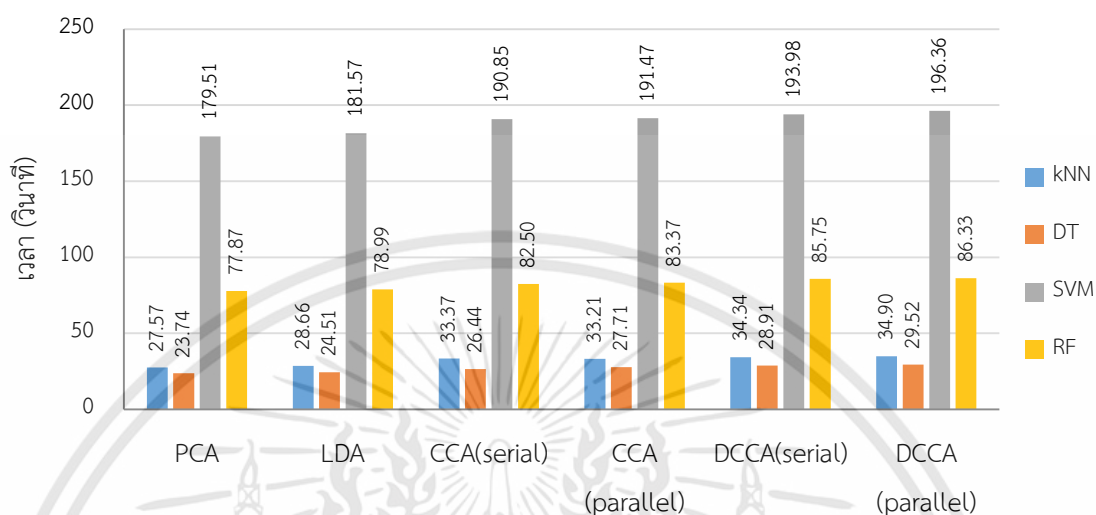
	CCA(parallel)	77.8236±0.45	74.0785±0.39	81.9311±0.48	27.7083±0.69
	DCCA(serial)	85.4835±0.58	83.0327±0.61	86.0843±0.43	28.9108±0.72
	DCCA(parallel)	80.9112±0.57	76.0313±0.66	82.1892±0.49	29.5230±0.47
SVM	PCA	74.1234±0.74	71.4371±0.61	78.0431±1.17	179.5070±0.79
	LDA	78.9872±0.56	76.9822±0.77	80.3430±0.91	181.5711±0.68
	CCA(serial)	87.8843±1.18	85.9234±0.73	88.9643±0.69	190.8524±0.77
	CCA(parallel)	82.1543±0.58	79.0374±0.68	85.6679±0.87	191.4713±0.93
	DCCA(serial)	90.1258±0.66	86.9831±0.37	92.0341±1.21	193.9791±0.82
	DCCA(parallel)	83.5428±0.65	79.8755±0.46	85.8753±0.71	196.3553±0.92
RF	PCA	71.7708±0.77	68.2311±0.51	73.8751±0.83	77.8702±0.44
	LDA	76.4576±0.62	73.4848±0.94	78.6053±0.78	78.9851±0.91
	CCA(serial)	87.4145±0.94	84.0861±0.71	90.0014±0.69	82.4985±0.66
	CCA(parallel)	80.2431±1.04	75.7655±0.98	84.0642±1.12	83.3657±0.53
	DCCA(serial)	87.9216±0.49	84.0875±1.12	88.8601±0.78	85.7495±0.69
	DCCA(parallel)	81.7864±0.67	77.4212±1.28	84.5591±0.95	86.3336±0.83

จากนั้นนำผลของค่าเฉลี่ยของค่าความถูกต้องแม่นยำของแต่ละวิธีที่ได้จากตารางที่ 4.7 มาพล็อตเป็นกราฟความสัมพันธ์แสดงดังรูปที่ 4.5



รูปที่ 4.5 ค่าเฉลี่ยความถูกต้องแม่นยำของจำแนกกระยะการนอนที่ได้จากการลดมิติด้วยวิธีการแปลงคุณลักษณะต่างๆ

จากนั้นนำเวลาการประมวลผลของเซตข้อมูลสำหรับการทดสอบของแต่ละวิธีที่ได้จากตารางที่ 4.7 มาพล็อตเป็นกราฟความสัมพันธ์แสดงดังรูปที่ 4.6



รูปที่ 4.6 เวลาเฉลี่ยในการประมวลผลของจำแนกกระยะการนอนที่ได้จากการลดมิติด้วยวิธีการแปลงคุณลักษณะต่างๆ

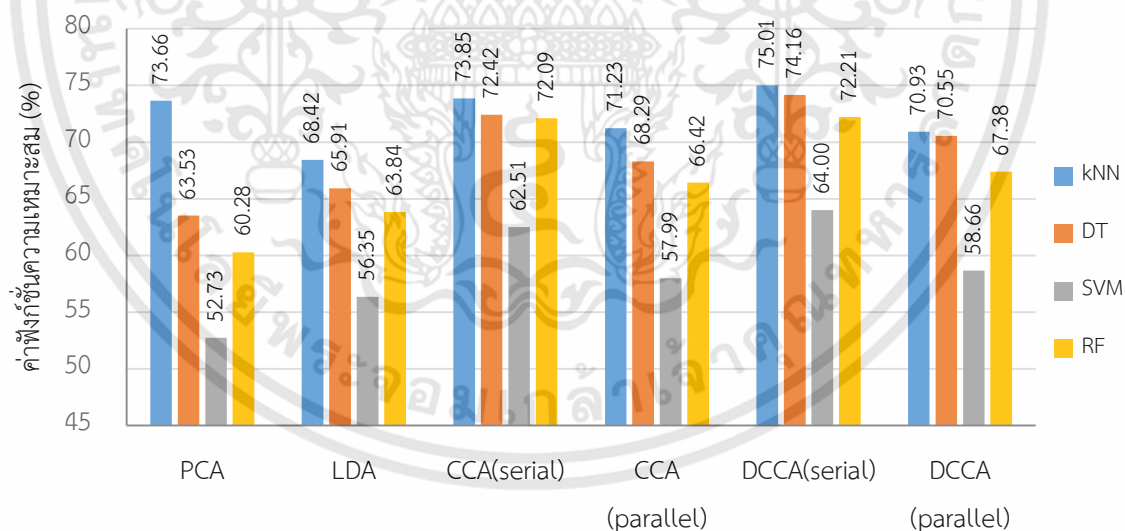
จากตารางผลการทดลองที่ 4.7 เมื่อทำการเปรียบเทียบระหว่างค่าเฉลี่ยความถูกต้องแม่นยำของแต่ละวิธีการลดมิติข้อมูลด้วยการแปลงคุณลักษณะ จะเห็นว่าค่าเฉลี่ยของความแม่นยำของการจำแนกกระยะการนอนด้วยวิธี DCCA และ CCA แสดงค่าสูงกว่าวิธี PCA และ LDA เนื่องจากข้อดีของวิธีการหาฟังก์ชันการแปลงของ DCCA และ CCA จะพิจารณาจากความสัมพันธ์ของคุณลักษณะทั้งสองด้าน คือ ความสัมพันธ์ระหว่างคุณลักษณะด้วยกันเองและความสัมพันธ์กับข้อมูลเอาต์พุต ในขณะที่การหาฟังก์ชันการแปลงของวิธี PCA และ LDA พิจารณาจากความสัมพันธ์เพียงด้านเดียว คือ ความสัมพันธ์ระหว่างคุณลักษณะของเซตข้อมูลการเรียนรู้และเอาต์พุตเท่านั้น เมื่อนำมาใช้กับเซตข้อมูลสำหรับการทดสอบซึ่งเป็นเซตข้อมูลที่ไม่เคยถูกนำมาเรียนรู้มาก่อนที่ได้มาใช้กับฟังก์ชันการแปลง PCA และ LDA จึงทำให้การจำแนกกระยะการนอนมีประสิทธิภาพลดลง เมื่อเปรียบเทียบผลการทดลองระหว่างวิธี DCCA และ CCA จะเห็นได้ว่าวิธี DCCA แสดงค่าความถูกต้องแม่นยำสูงสุด ทั้งนี้เนื่องจากวิธี DCCA ไม่เพียงแต่หาฟังก์ชันการแปลงด้วยความสัมพันธ์ระหว่างคุณลักษณะด้วยเองกันเท่านั้น แต่ยังใช้ข้อมูลของกลุ่มหรือ Class information มาร่วมในการพิจารณาด้วยซึ่งช่วยเพิ่มประสิทธิภาพการจำแนกข้อมูลได้มากขึ้น จึงทำให้มีประสิทธิภาพการจำแนกกระยะการนอนสูงกว่าวิธี CCA ขณะเดียวกันเมื่อเปรียบเทียบระหว่างวิธีการหลอมรวมคุณลักษณะแบบขนานและแบบอนุกรมของวิธี CCA และ DCCA จะเห็นได้ว่า การหลอมรวมแบบอนุกรมแสดงค่าความถูกต้องแม่นยำสูงกว่าการหลอมรวมแบบขนาน ทั้งนี้เนื่องจากเงื่อนไขการหาไอเกนเวกเตอร์และค่าไอเกนของแต่ละคู่ที่มีขนาดมิติไม่เท่ากัน ทำให้ผลลัพธ์ของไอเกนเวกเตอร์และค่าไอเกนที่

ได้จะมีจำนวนเท่ากับมิติที่น้อยที่สุด ดังนั้นเมื่อนำข้อมูลมาหลอมรวมกันแบบขนานจึงต้องมีการจัดเรียงให้เมทริกซ์มีขนาดเท่ากัน ดังนั้นค่าไอเกนที่มีค่าน้อยซึ่งอาจมีผลต่อการประมวลผลถูกตัดออกไป จึงเป็นสาเหตุทำให้เกิดการสูญหายของข้อมูล และทำให้ประสิทธิภาพการจำแนกกระยะการนอนที่มีการหลอมรวมแบบขนานแสดงค่าความถูกต้องแม่นยำลดลง ดังนั้นจากการทดลองนี้แสดงให้เห็นว่า วิธีการแปลงคุณลักษณะสำหรับการลดมิติของชุดข้อมูลหลายชุดคือ วิธีการแปลงคุณลักษณะ DCCA ที่หลอมรวมแต่ละเซตข้อมูลแบบอนุกรม

นอกจากนี้เราได้ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึมต่างๆ ในขั้นตอนการทดสอบด้วยการคำนวณฟังก์ชันความเหมาะสมของขั้นตอนการจำแนก ($Fitness_{classifier}$) ที่พิจารณาจากค่าความถูกต้องแม่นยำของการจำแนกข้อมูลและเวลาที่ใช้ในการประมวลผล ด้วยสมการ

$$Fitness_{classifier} = \left(\alpha \times \left(\frac{Acc}{100} \right) + (1 - \alpha) \left(\frac{Time_{max} - Time}{Time_{max}} \right) \right) \times 100 \quad (4.1)$$

โดยที่ α คือ ค่าน้ำหนักความสำคัญของจุดประสงค์ที่ต้องการให้ความสำคัญ ในการทดลองนี้เราได้เน้นความสำคัญของความถูกต้องแม่นยำมากกว่าเวลาที่ใช้ในการประมวลผล จึงกำหนดให้ค่า α เท่ากับ 0.8 Acc คือ ร้อยละความถูกต้องแม่นยำ $Time_{max}$ คือเวลาในการประมวลผลที่สูงที่สุด $Time$ คือเวลาในการประมวลผล ผลการทดลองค่าฟังก์ชันความเหมาะสมของขั้นตอนการทดสอบแสดงดังรูปที่ 4.7



รูปที่ 4.7 ค่าฟังก์ชันความเหมาะสมของแต่ละการลดมิติข้อมูลด้วยวิธีการแปลงข้อมูลต่างๆและการจำแนกข้อมูลด้วยอัลกอริทึม kNN, DT, SVM และ RF

จากผลการทดลองรูปที่ 4.7 เมื่อเปรียบเทียบผลการทดลองจะเห็นว่าค่าฟังก์ชันความเหมาะสมของการลดมิติด้วยการแปลงคุณลักษณะ DCCA ที่มีการหลอมรวมแบบอนุกรมและการจำแนกข้อมูลด้วย

อัลกอริทึม kNN แสดงค่าสูงสุดเท่ากับร้อยละ 75.01 ดังนั้นจากผลการทดลองนี้เราจึงได้กำหนดให้การจำแนกข้อมูลด้วยอัลกอริทึม kNN เป็นวิธีที่เหมาะสมสำหรับนำไปใช้ในการเพิ่มประสิทธิภาพการจำแนกกระยะการนอน

4.1.2.2 การทดลองและผลการทดลองของประสิทธิภาพของวิธีการลัดมิติข้อมูลด้วยการเลือกคุณลักษณะ

การทดลองนี้มีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพระหว่างวิธีการเลือกคุณลักษณะด้วยการค้นหาคำตอบ NIAs ที่ใช้ในงานวิจัยทั่วไป ได้แก่ BPSO, BCS, EBCS, BGWO และวิธีการเลือกคุณลักษณะที่งานวิจัยนี้ได้นำเสนอ DisBGWO ซึ่งเป็นวิธีการคุณลักษณะแบบใหม่โดยนำค่าประสิทธิภาพความแตกต่างของแฉีกคาร์ตมาใช้ในการปรับปรุงอัลกอริทึม BGWO แบบเดิมด้วยแนวความคิดการกลายพันธุ์โดยเซตของคุณลักษณะที่ได้จากการทดลองหัวข้อที่ 4.1.1 ถูกนำมาใช้ในการทดลองนี้ซึ่งเซตของคุณลักษณะทั้งหมดถูกสุ่มแบ่งเป็น 2 เซตข้อมูลคือ เซตข้อมูลสำหรับการเรียนรู้จำนวนร้อยละ 80 และเซตของข้อมูลสำหรับการทดสอบจำนวนร้อยละ 20 ขั้นตอนของการทดลองประกอบด้วย 2 ขั้นตอนหลัก ขั้นตอนแรกเป็นขั้นตอนการเรียนรู้ เริ่มต้นด้วยเซตคุณลักษณะถูกเลือกเซตคุณลักษณะที่เหมาะสมด้วยวิธีการเลือกคุณลักษณะต่างๆ รายละเอียดอัลกอริทึมของวิธีการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO ตามหัวข้อ 3.1.2.2 จากนั้นเซตของคุณลักษณะที่ถูกเลือกจะถูกนำเข้าสู่วิธีการจำแนกข้อมูลด้วยอัลกอริทึม kNN และจะถูกประเมินประสิทธิภาพด้วยฟังก์ชันความเหมาะสมตามสมการ 3.2 โดยกำหนดให้ค่าคงที่ γ เท่ากับ 0.8 ซึ่งหมายถึงงานวิจัยนี้ให้ความสำคัญกับค่าความถูกต้องแม่นยำมากกว่าจำนวนคุณลักษณะ เมื่อดำเนินการคำนวณซ้ำจนกระทั่งครบจำนวนรอบที่กำหนด ผลลัพธ์ที่ได้คือ เซตคุณลักษณะที่เหมาะสมและค่าฟังก์ชันความเหมาะสมของแต่ละรอบการคำนวณ สำหรับขั้นตอนการทดสอบประสิทธิภาพ เซตคุณลักษณะที่ถูกเลือกจากวิธีการเลือกคุณลักษณะต่างๆจะถูกทดสอบด้วยเซตข้อมูลสำหรับการทดสอบ

1. การทดลองที่ 2 การทดลองเปรียบเทียบประสิทธิภาพของวิธีการลัดมิติด้วยการเลือกคุณลักษณะสำหรับการจำแนกระยะการนอนจากสัญญาณ EEG EOG และ ECG

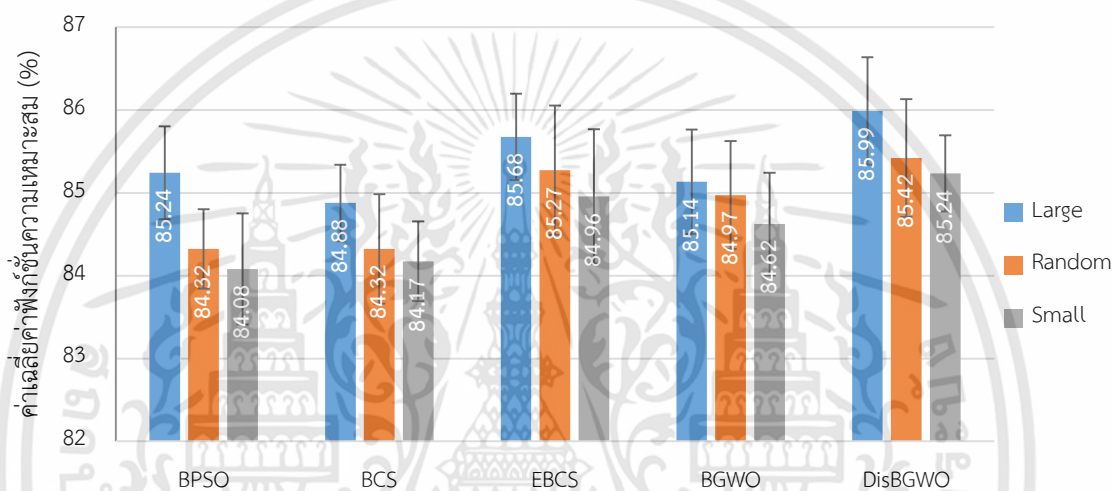
จุดประสงค์การทดลอง เพื่อเปรียบเทียบประสิทธิภาพของการลัดมิติด้วยการเลือกคุณลักษณะ BPSO, BCS, EBCS และ DisBGWO สำหรับการจำแนกระยะการนอน

สำหรับขั้นตอนการเลือกคุณลักษณะด้วยวิธี NIAs การกำหนดค่าเริ่มต้นของวิธีการค้นหาคำตอบเป็นขั้นตอนสำคัญซึ่งมีผลต่อการเลือกเซตคุณลักษณะที่มีประสิทธิภาพ ดังนั้นงานวิจัยนี้จึงได้ตั้งสมมุติฐานว่า การสุ่มค่าคำตอบสำหรับวิธีการค้นหาคำตอบเริ่มต้นที่แตกต่างกันมีผลต่อประสิทธิภาพการจำแนกกระยะการนอน ดังนั้นเราจึงทดลองเพื่อทดสอบสมมุติฐานดังกล่าวด้วยการกำหนดค่าเริ่มต้นของวิธีการค้นหาคำตอบที่แตกต่างกัน 3 รูปแบบ ดังนี้

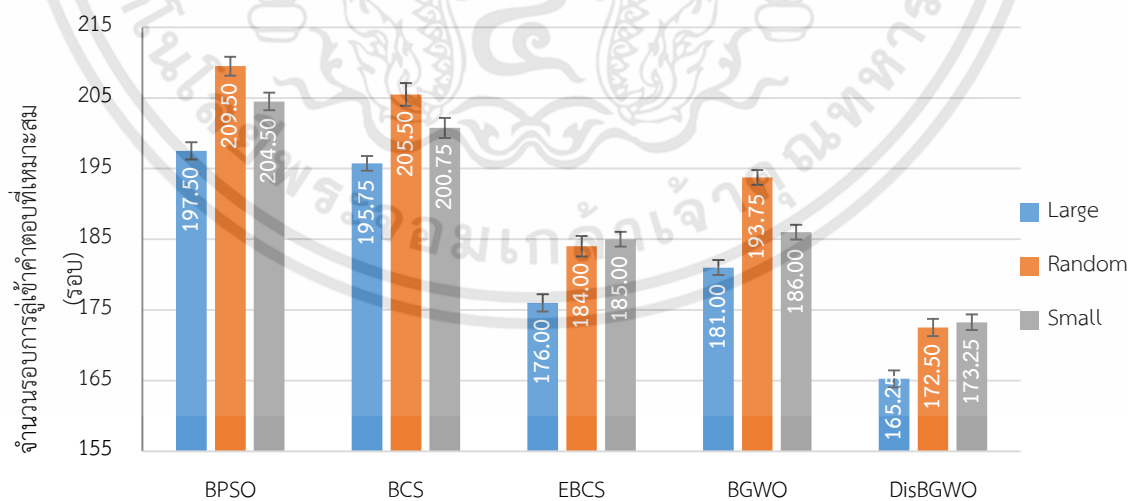
1. การกำหนดค่าเริ่มต้นของวิธีค้นหาคำตอบด้วยการสุ่มเลือกคุณลักษณะที่มีจำนวนน้อย (Small initialization) หรือเริ่มต้นด้วยการสุ่มเลือกของคุณลักษณะเพียงตัวเดียว และจะเพิ่มขึ้นตามอัลกอริทึมการค้นหาคำตอบ NIAs

2. การกำหนดค่าเริ่มต้นของวิธีค้นหาคำตอบด้วยการสุ่มเลือกคุณลักษณะแบบสุ่ม (Random initialization)

3. การกำหนดค่าเริ่มต้นของวิธีค้นหาคำตอบด้วยการสุ่มเลือกคุณลักษณะที่มีจำนวนมาก (Large initialization) หรือเริ่มต้นด้วยการเลือกคุณลักษณะทั้งหมดก่อนและจะลดลงตามอัลกอริทึม NIAs



รูปที่ 4.8 ค่าเฉลี่ยค่าฟังก์ชันความเหมาะสมของการกำหนดค่าเริ่มต้นของแต่ละวิธีค้นหาคำตอบด้วยการสุ่มเลือกแบบต่างๆ ได้แก่ Large, Random และ Small initializations



รูปที่ 4.9 จำนวนรอบการคำนวณการสุ่มเข้าคำตอบที่เหมาะสมของการกำหนดค่าเริ่มต้นของวิธีค้นหาคำตอบด้วยการสุ่มเลือกแบบต่างๆ ได้แก่ Large, Random และ Small initializations

ผลการทดลองรูปที่ 4.8 และ 4.9 แสดงค่าเฉลี่ยฟังก์ชันความเหมาะสม และจำนวนรอบของการลู่เข้าคำตอบที่เหมาะสมของการกำหนดค่าเริ่มต้นของวิธีค้นหาคำตอบด้วยการสุ่มเลือกแบบต่างๆ จะเห็นได้ว่าการสุ่มคุณลักษณะแบบ Large initialization แสดงค่าเฉลี่ยฟังก์ชันความเหมาะสมสูงกว่าและใช้จำนวนรอบการคำนวณในการลู่เข้าคำตอบที่เหมาะสมน้อยกว่าวิธี Random initialization และ Small initialization ดังนั้นจากการทดลองนี้เราจึงได้กำหนดเงื่อนไขการสุ่มคุณลักษณะสำหรับวิธีการค้นหาคำตอบเริ่มต้นด้วยวิธี Large initialization สำหรับการเพิ่มประสิทธิภาพการจำแนกระยะการนอน

เงื่อนไขการทดลอง

1. ข้อมูลสำหรับการเรียนรู้ถูกสุ่มแบ่งด้วย Cross-validation (CV) ที่ $k=5$ ประกอบด้วย 2 เซต คือ เซตของข้อมูลสำหรับการสร้างโมเดล และเซตข้อมูลสำหรับการปรับแต่ง
2. ขั้นตอนการเรียนรู้ เราได้ประเมินประสิทธิภาพของเซตคุณลักษณะที่ถูกเลือก ค่าเฉลี่ยฟังก์ชันความเหมาะสม ค่าเฉลี่ยความถูกต้องแม่นยำ จำนวนคุณลักษณะที่ถูกเลือก จำนวนรอบของการลู่เข้าคำตอบที่เหมาะสม และเวลาในการประมวลผลของขั้นตอนการเรียนรู้
3. การกำหนดพารามิเตอร์ของวิธีเลือกคุณลักษณะของ NIAs มีดังนี้ จำนวนรอบของการค้นหาตอบของวิธี NIAs เท่ากับ 250 รอบ จำนวนคำตอบในแต่ละวิธีค้นหาคำตอบเท่ากับจำนวนคุณลักษณะทั้งหมดเท่ากับ 364 การกำหนดพารามิเตอร์ของวิธีการเลือกคุณลักษณะ NIAs ต่างๆ แสดงรายละเอียดดังตารางที่ 4.8 และกำหนดเงื่อนไขของการสุ่มการค้นหาคำตอบเริ่มต้นด้วยวิธี Large initialization
4. การจำแนกข้อมูลด้วยอัลกอริทึม kNN ที่ $k=5$
5. ผลการทดสอบประสิทธิภาพของแต่ละวิธีแสดงในรูปแบบของ ค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะ ค่าความแปรปรวน เวลาในการประมวลผล และค่าฟังก์ชันความเหมาะสมสำหรับการจำแนกซึ่งผลการทดสอบประสิทธิภาพในขั้นตอนการทดสอบได้มาจากค่าเฉลี่ยของการทดลองจำนวน 10 ครั้ง

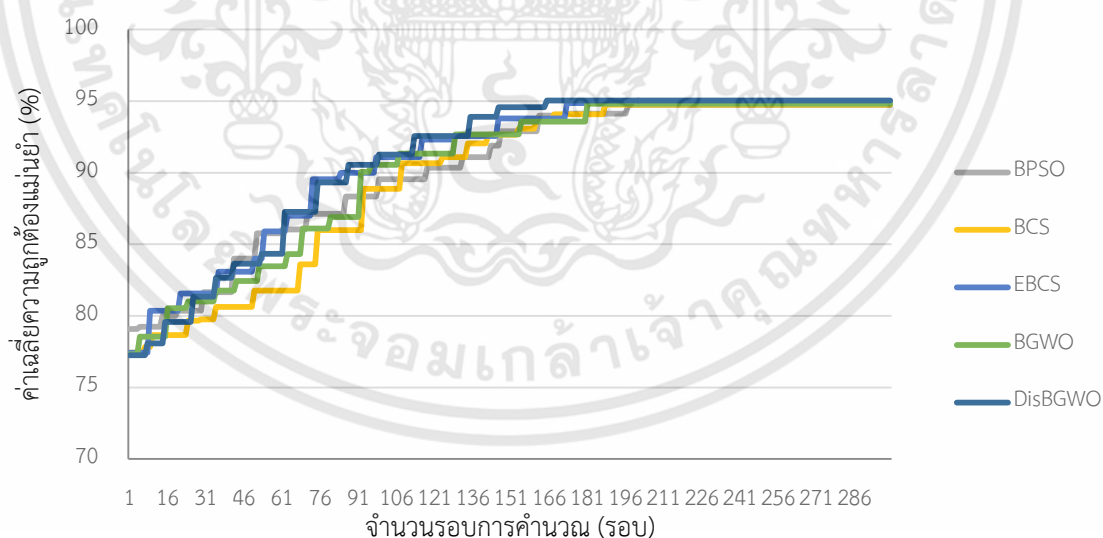
ตารางที่ 4.8 พารามิเตอร์สำหรับวิธีการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO

อัลกอริทึม	พารามิเตอร์
BPSO	ค่าคงที่ของสมการ 2.45 ได้แก่ $gbest (C_1 = 1)$, $lbest (C_2 = 2)$ และค่า $w = 0.72$
BCS	ค่าคงที่ของสมการ 2.47 ได้แก่ Levy ($\lambda = 1.25$), $\alpha = 1$ และค่าพารามิเตอร์ความน่าจะเป็นสำหรับการละทิ้งคำตอบ $P_a = 0.25$
EBCS	ค่าคงที่ของสมการ 2.47 ได้แก่ Levy ($\lambda = 1.25$), $\alpha = 1$, ค่าพารามิเตอร์ความน่าจะเป็นสำหรับการละทิ้งคำตอบ $P_a = 0.25$ จำนวนของการสร้างโซลูชันการ

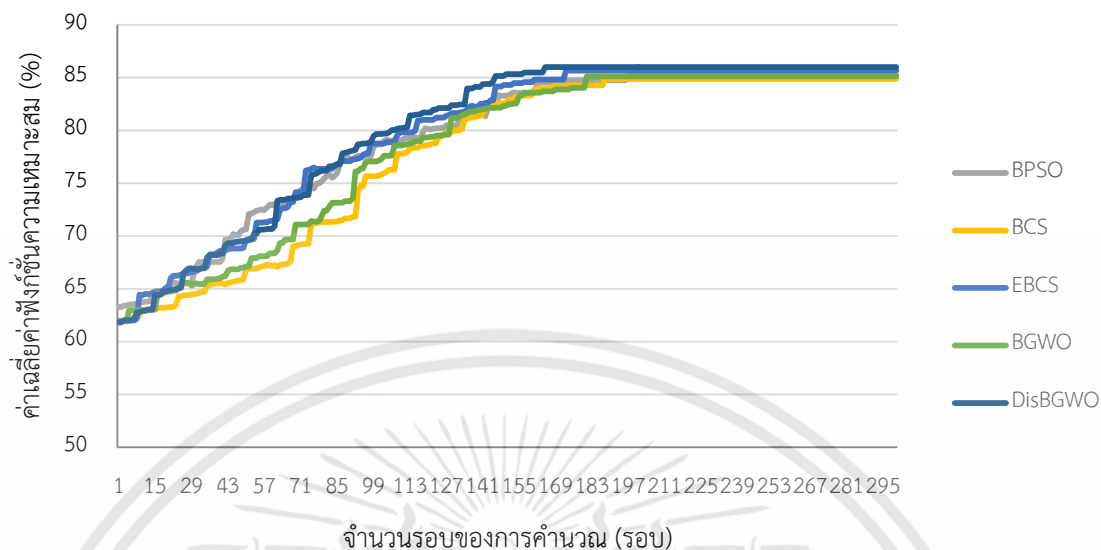
	กลายพันธุ์ $S_{mut} = 10$ และจำนวนค่าคำตอบที่ถูกกลายพันธุ์ $N_{mut} = 120$ ประมาณ $1/3$ ของจำนวนคุณลักษณะทั้งหมด [6]
BGWO	ค่าคงที่ของสมการ 2.54 ฟังก์ชัน sigmoid ($m = 10, n = 0.5$)
DisBGWO	ค่าคงที่ของสมการ 2.54 ฟังก์ชัน sigmoid ($m = 10, n = 0.5$) และ ค่าคงที่ของสมการ 3.6 ($\phi = 0.8$)

2. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการเรียนรู้การทดลองที่ 2

ผลการทดลองขั้นตอนการเรียนรู้การทดลองที่ 2 แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความถูกต้องแม่นยำและจำนวนรอบของการคำนวณแสดงดังรูปที่ 4.10 และกราฟความสัมพันธ์ระหว่างค่าเฉลี่ยฟังก์ชันความเหมาะสมและจำนวนรอบของการคำนวณแสดงดังรูปที่ 4.11 ผลการทดสอบประสิทธิภาพของวิธีการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO ในขั้นตอนการเรียนรู้ แสดงดังตารางที่ 4.9 ซึ่งแสดงด้วยค่าเฉลี่ยฟังก์ชันความเหมาะสมตามสมการ 3.2 (Fitness function value) จำนวนคุณลักษณะที่ถูกเลือก (The number of selected features) หรือจำนวนคุณลักษณะที่ทำให้ค่าฟังก์ชันความเหมาะสมเข้าสู่คำตอบที่เหมาะสม ค่าเฉลี่ยความถูกต้องแม่นยำของเซตคุณลักษณะที่เหมาะสม (Accuracy of selected features) จำนวนรอบการคำนวณที่เข้าสู่คำตอบที่เหมาะสม (The number of iterations) และเวลาในการประมวลผล (Time)



รูปที่ 4.10 ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธี BPSO, BCS, EBCS, BGWO และ DisBGWO



รูปที่ 4.11 ความสัมพันธ์ระหว่างค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธี BPSO, BCS, EBCS, BGWO และ DisBGWO

ตารางที่ 4.9 ผลการทดลองของขั้นตอนเรียนรู้ของการลดมิติด้วยวิธีการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO ในขั้นตอนเรียนรู้

Algorithms	Fitness function value (%±SD)	The number of selected features (Number±SD)	Accuracy of selected features (%±SD)	The number of iterations (Number±SD)	Time (sec.±SD)
BPSO	85.2423±0.53	193.33±0.57	94.7396±0.64	197.50±1.00	18,522.89±182.44
BCS	84.8779±0.72	200.00±0.81	94.7653±0.88	195.75±0.95	17,858.21±174.39
EBCS	85.6753±1.05	186.50±1.81	94.8689±0.61	176.00±2.16	15,511.71±191.48
BGWO	85.1355±0.62	194.75±1.25	94.8123±0.76	181.00±1.41	16,451.28±192.67
DisBGWO	85.9878±0.71	185.75±0.95	95.0534±0.56	165.25±2.50	14,301.45±221.81

จากผลการทดลองรูปที่ 4.9 - 4.10 และตารางผลการทดลองที่ 4.9 แสดงให้เห็นว่า วิธีการลดมิติด้วยการเลือกคุณลักษณะ DisBGWO เป็นวิธีที่แสดงค่าเฉลี่ยฟังก์ชันความเหมาะสมและค่าเฉลี่ยความถูกต้องแม่นยำสูงสุดเท่ากับร้อยละ 85.9878±0.71 และ 95.0534±0.56 ตามลำดับ นอกจากนี้ยังเป็นวิธีที่ลู่เข้าคำตอบที่เหมาะสมด้วยจำนวนรอบของการคำนวณน้อยที่สุดเท่ากับ 165.25±2.50 ทั้งนี้เนื่องจาก

การเลือกคุณลักษณะ DisBGWO เป็นวิธีที่มีการค้นหาค่าคำตอบทั้งจากข้อมูลปริภูมิครอบคลุมและปริภูมิเฉพาะที่ซึ่งทำให้การค้นหาคำตอบได้อย่างรวดเร็ว ซึ่งคล้ายกับเทคนิคการค้นหาด้วยวิธี EBCS โดยผลการทดลองจะเห็นได้ว่าวิธี EBCS สามารถลู่เข้าคำตอบที่เหมาะสมด้วยจำนวนรอบของการคำนวณน้อยเป็นลำดับที่สองเท่ากับ 176.00 ± 2.16 และยังค้นหาคำตอบได้เร็วกว่าวิธีการค้นหาคำตอบ BPSO, BCS และ BGWO ซึ่งเป็นการค้นหาคำตอบจากข้อมูลปริภูมิครอบคลุมอย่างเดียว แต่อย่างไรก็ตามการลดมิติด้วยวิธีการเลือกคุณลักษณะทั้งหมดนั้นยังมีข้อเสีย คือ ในขั้นตอนการเรียนรู้ใช้เวลาในการประมวลผลนาน เนื่องจากการค้นหาเซตคุณลักษณะทั้งหมดถูกนำมาพิจารณาร่วมกันทีเดียวทั้งหมด จึงทำให้การค้นหาคุณลักษณะที่เหมาะสมใช้เวลานาน

3. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการทดสอบการทดลองที่ 2

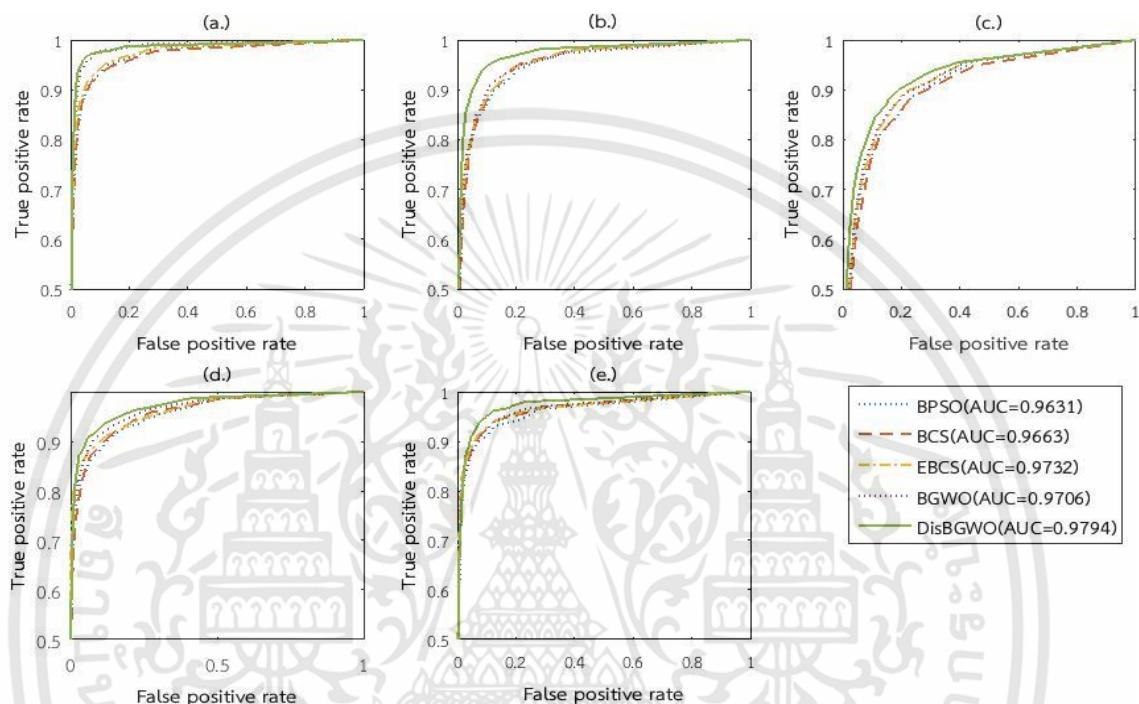
ผลการทดสอบประสิทธิภาพการจำแนกกระยะการนอนของวิธีการลดมิติด้วยการเลือกคุณลักษณะด้วยวิธีต่างๆ ของขั้นตอนการทดสอบแสดงดังตารางที่ 4.10 ซึ่งแสดงด้วยค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะ ค่าเฉลี่ยเวลาของการประมวลผล และค่าเฉลี่ยฟังก์ชันความเหมาะสมของการจำแนกข้อมูลตามสมการที่ 4.1

ตารางที่ 4.10 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลดมิติด้วยการเลือกคุณลักษณะ BPSO, BCS, EBCS, BGWO และ DisBGWO

Algorithms	ArgAcc. (%±SD)	ArgSent. (%±SD)	ArgSpec. (%±SD)	ArgTime (sec.±SD)	Fitness <i>classifier</i> (%±SD)
BPSO	93.4016±0.84	89.8311±0.66	97.6325±0.74	48.5227±3.03	75.1864±0.71
BCS	93.5567±0.68	91.0932±1.28	96.5201±0.93	49.4421±4.93	74.8443±0.67
EBCS	93.9422±1.01	90.6222±0.75	96.6212±0.95	46.8231±6.22	76.5592±0.59
BGWO	93.5889±0.86	91.1938±0.95	96.7341±0.63	47.6723±4.32	75.7631±0.71
DisBGWO	94.2961±0.96	92.2121±0.98	97.0901±0.83	46.0012±4.43	77.1807±0.69

จากผลการทดลองตารางที่ 4.10 เมื่อได้เซตของคุณลักษณะที่ถูกเลือกด้วยแต่ละวิธีแล้วจะถูกนำมาทดสอบประสิทธิภาพด้วยเซตข้อมูลการทดสอบซึ่งเป็นเซตข้อมูลที่ไม่เคยถูกเรียนรู้มาก่อน จากผลการทดลองแสดงให้เห็นว่า วิธีการเลือกคุณลักษณะ DisBGWO เป็นวิธีที่ให้ประสิทธิภาพสูงสุดด้วยค่าเฉลี่ยความถูกต้องแม่นยำเท่ากับ 94.2961 ± 0.96 ค่าเฉลี่ยความไวเท่ากับร้อยละ 92.2121 ± 0.98 และค่าเฉลี่ยจำเพาะเท่ากับ 97.0901 ± 0.83 นอกจากนี้ความเหมาะสมของฟังก์ชันการจำแนกแสดงค่าที่สูงสุดเท่ากับร้อยละ 77.1807 ± 0.69 เมื่อเปรียบเทียบกับผลการทดลองที่ 1 จะเห็นได้ว่าฟังก์ชันความเหมาะสมของ

การจำแนกของวิธี DisBGWO มีค่าสูงกว่าวิธี DCCCA ซึ่งเป็นวิธีที่เหมาะสมที่สุดของการทดลองที่ 1 จาก การเปรียบเทียบดังกล่าวสามารถสรุปได้ว่า วิธีการลดมิติด้วยวิธีการเลือกคุณลักษณะมีความเหมาะสมและมีประสิทธิภาพมากกว่าวิธีการลดมิติด้วยวิธีการแปลงคุณลักษณะสำหรับการจำแนกระยะการนอนโดยใช้หลายสัญญาณ



รูปที่ 4.12 กราฟ ROC ของวิธีการเลือกคุณลักษณะต่างๆ (BPSO, BCS, EBCS, BGWO และ DisBGWO) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.)

สำหรับผลการทดลองกราฟ ROC (Receiver operation characteristic curve) ของแต่ละวิธีการของการเลือกคุณลักษณะในการจำแนกในแต่ละระยะการนอน แสดงดังรูปที่ 4.12 ซึ่งผลการทดลองจะเห็นได้ว่าค่าพื้นที่ใต้กราฟหรือ Area Under Curve (AUC) ของวิธีการเลือกคุณลักษณะด้วยวิธีการของ DisBGWO ให้ค่าสูงสุดเท่ากับ 0.9794

4.1.2.3 การทดลองและผลการทดลองของประสิทธิภาพของการลดมิติแบบผสมผสานที่นำเสนอ

จากผลการทดลองที่ 2 แสดงให้เห็นว่าวิธีการลดมิติด้วยการเลือกคุณลักษณะเป็นวิธีที่ถึงแม้จะให้ค่าความถูกต้องแม่นยำสูง แต่ยังมีข้อเสียคือ เวลาที่ใช้ในการประมวลผลในขั้นตอนการเรียนรู้ นั่นเราจึงได้นำเสนอและออกแบบวิธีการลดมิติแบบผสมผสานใหม่ตามแบบที่ 3 และแบบที่ 4 ตามหัวข้อที่ 2.7 ซึ่งวิธีการลดมิติแบบผสมผสานที่งานวิจัยนี้เสนอเรียกว่า CCA/DCCA-NIAs และ NIAs-

DCCA/CCA ซึ่งการทดลองสำหรับการทดสอบประสิทธิภาพของวิธีการที่นำเสนอทั้งสองมีรายละเอียดดังต่อไปนี้

1. การทดลองที่ 3 การทดลองเปรียบเทียบประสิทธิภาพของวิธีการลมิติแบบผสมผสาน CCA/DCCA-NIAs สำหรับการจำแนกระยะการนอนโดยใช้สัญญาณ EEG EOG และ ECG

การทดสอบประสิทธิภาพของวิธีการลมิติแบบผสมผสาน CCA/DCCA-NIAs เป็นวิธีการลมิติที่เริ่มต้นด้วยวิธีการหลอมรวมคุณลักษณะ CCA/DCCA และตามด้วยวิธีการเลือกคุณลักษณะ NIAs วิธีการลมิติแบบผสมผสาน CCA/DCCA-NIAs ประกอบด้วย 10 รูปแบบ ดังนี้ CCA-BPSO, CCA-BCS, CCA-EBCS, CCA-BGWO, CCA-DisBGWO, DCCA-BPSO, DCCA-BCS, DCCA-EBCS, DCCA-BGWO และ CCA-DisBGWO โดยเซตของคุณลักษณะที่ได้จากการทดลองหัวข้อที่ 4.1.1 ถูกนำมาใช้ในการทดลอง เซตของคุณลักษณะแต่ละชุดถูกสุ่มแบ่งเป็น 2 เซต ได้แก่ เซตข้อมูลสำหรับการเรียนรู้จำนวนร้อยละ 80 ของคุณลักษณะทั้งหมด และเซตของข้อมูลสำหรับการทดสอบจำนวนร้อยละ 20 ของคุณลักษณะทั้งหมด รายละเอียดของอัลกอริทึม CCA/DCCA-NIAs ตามหัวข้อ 3.1.2.3 ข้อ 1 ซึ่งเริ่มต้นด้วยแต่ละเซตคุณลักษณะถูกหลอมรวมด้วยวิธี CCA/DCCA ซึ่งหลอมรวมทีละคู่จำนวนทั้งหมด 10 คู่และหลอมรวมแบบอนุกรม และเลือกขนาดของไอเกินเท่ากับค่าความเชื่อมั่นที่ร้อยละ 90 หลังจากนั้นเลือกเซตคุณลักษณะที่เหมาะสมด้วยวิธี NIAs และถูกประเมินด้วยค่าฟังก์ชันความเหมาะสมตามสมการ 3.2 ซึ่งกำหนดให้ค่าคงที่ γ เท่ากับ 0.8 สำหรับการขั้นตอนการทดสอบ โดยแต่ละวิธีจะถูกทดสอบประสิทธิภาพด้วยเซตข้อมูลของการทดสอบ สิ่งที่จะพิจารณาความเหมาะสมของวิธีการต่างๆ คือ ประสิทธิภาพความถูกต้องแม่นยำของการจำแนกระยะการนอน และเวลาที่ใช้ในการประมวลผล

จุดประสงค์การทดลอง เพื่อเปรียบเทียบประสิทธิภาพของการลมิติแบบผสมผสานที่นำเสนอ CCA/DCCA-NIAs สำหรับการจำแนกระยะการนอน Wake, REM, NREM1-3

เงื่อนไขการทดลอง

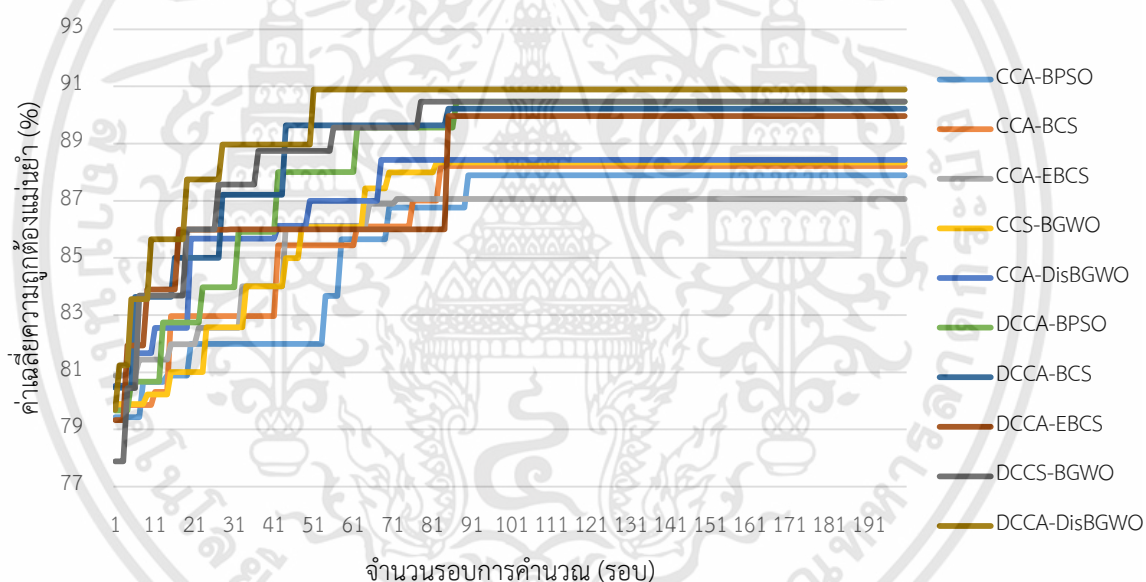
1. ข้อมูลสำหรับการเรียนรู้ถูกสุ่มแบ่งด้วย Cross-validation ที่ $k=5$ ประกอบด้วย 2 เซตคือ เซตของข้อมูลสำหรับการสร้างโมเดล และเซตข้อมูลสำหรับการสอบทวน
2. ขั้นตอนการเรียนรู้ เราได้ประเมินประสิทธิภาพของเซตคุณลักษณะที่ถูกเลือก ค่าเฉลี่ยฟังก์ชันความเหมาะสม ค่าเฉลี่ยความถูกต้องแม่นยำ จำนวนคุณลักษณะที่ถูกเลือก จำนวนรอบของการสุ่มเข้าคำตอบที่เหมาะสม และเวลาในการประมวลผลของขั้นตอนการเรียนรู้
3. กำหนดพารามิเตอร์ของวิธีเลือกคุณลักษณะของ NIAs มีดังนี้ จำนวนรอบของการค้นหาตอบของวิธี NIAs เท่ากับ 200 รอบ จำนวนคำตอบในแต่ละวิธีการค้นหาคำตอบเท่ากับ 364 และการกำหนดพารามิเตอร์สำหรับวิธีการ NIAs แสดงรายละเอียดตารางที่ 4.8 กำหนดเงื่อนไขของการสุ่มการค้นหาคำตอบเริ่มต้นด้วยวิธี Large initialization

4. การจำแนกข้อมูลด้วยอัลกอริทึม kNN ที่ $k=5$

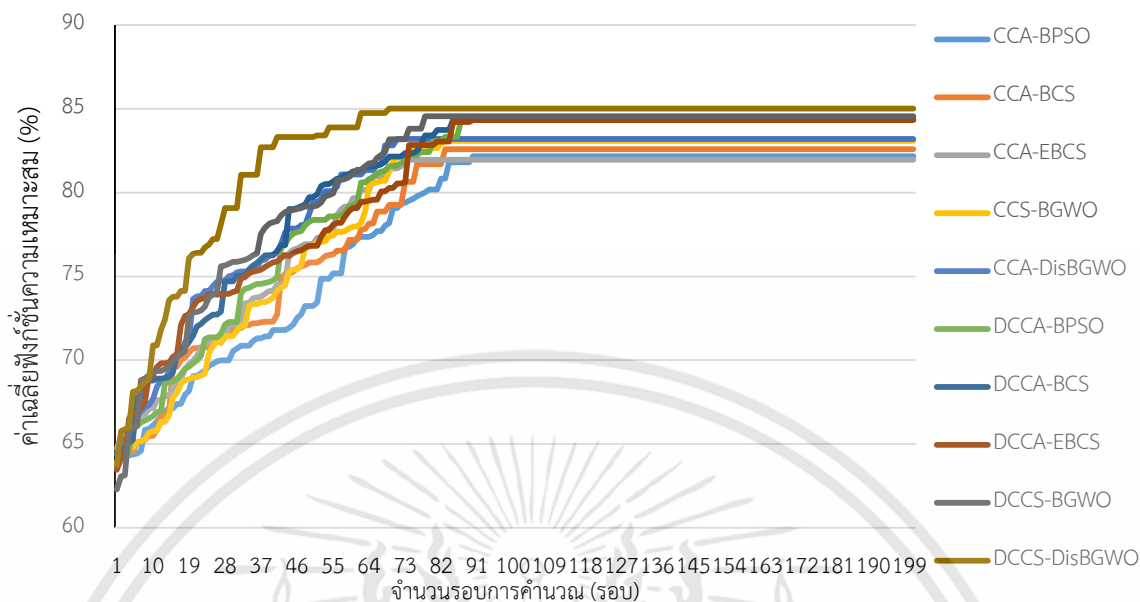
5. ผลการทดสอบประสิทธิภาพของแต่ละวิธีแสดงในรูปแบบของ ค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะ ค่าความแปรปรวน และเวลาในการประมวลผล ซึ่งผลการทดสอบประสิทธิภาพในขั้นตอนการทดสอบได้มาจากค่าเฉลี่ยของการทดลองจำนวน 10 ครั้ง

2. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการเรียนรู้ของการทดลองที่ 3

ผลการทดลองขั้นตอนการเรียนรู้แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความถูกต้องแม่นยำของการจำแนกกระยะการนอนและจำนวนรอบการคำนวณแสดงดังรูปที่ 4.13 และกราฟความสัมพันธ์ระหว่างค่าเฉลี่ยของฟังก์ชันค่าความเหมาะสมและจำนวนรอบของการคำนวณดังรูปที่ 4.14 ผลการทดสอบประสิทธิภาพการจำแนกกระยะการนอนของวิธีการลดมิติแบบผสมผสาน CCA/DCCA-NIAs ในขั้นตอนการเรียนรู้แสดงดังตารางที่ 4.13



รูปที่ 4.13 ความสัมพันธ์ระหว่างค่าเฉลี่ยของความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน CCA-NIAs และ DCCA-NIAs



รูปที่ 4.14 ความสัมพันธ์ระหว่างค่าเฉลี่ยของค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธี
มิติแบบผสมผสาน CCA-NIAs และ DCCA-NIAs

ตารางที่ 4.11 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองที่ 3 วิธีมิติแบบผสมผสาน CCA-NIAs
และ DCCA-NIAs

Algorithms	Fitness function value (%±SD)	The number of selected features (Number±SD)	Accuracy of selected features (%±SD)	The number of iterations (Number±SD)	Time (sec.±SD)
CCA-BPSO	82.1839±0.94	149.00±1.54	87.8947±0.87	90.00±1.41	8,134.43±221.81
CCA-BCS	82.6055±0.76	147.75±2.21	88.2157±1.33	84.250.95	7,411.51±176.36
CCA-EBCS	81.8920±0.69	139.25±0.95	87.0679±0.89	73.25±1.25	6,646.43±168.73
CCA-BGWO	83.0887±0.98	138.75±2.98	88.3389±1.03	83.25±1.50	6,131.82±192.65
CCA-DisBGWO	83.2159±0.69	136.00±0.81	88.4292±0.94	68.25±0.95	5,696.11±222.96
DCCA-BPSO	84.3540±0.78	148.00±1.82	90.4743±0.78	89.00±1.41	6,261.36±222.87
DCCA-BCS	84.3775±1.11	143.25±1.26	90.2246±1.06	85.75±0.95	5,961.45±231.53
DCCA-EBCS	84.3342±1.03	139.25±0.95	89.9644±0.75	73.75±0.95	5,326.55±241.98
DCCA-BGWO	84.5668±0.97	142.50±2.81	90.4612±0.89	79.50 ±1.29	5,506.93±236.45
DCCA-DisBGWO	85.0249±0.68	139.25±0.95	90.8965±0.85	63.00±1.81	4,921.44±236.76

จากผลการทดลองรูปที่ 4.13-4.14 และตารางผลการทดลองที่ 4.11 แสดงให้เห็นว่าวิธีการลดมิติแบบผสมผสาน DCCA-DisBGWO เป็นวิธีที่แสดงค่าเฉลี่ยฟังก์ชันความเหมาะสมสูงสุดเท่ากับร้อยละ 85.0249 ± 0.68 และค่าเฉลี่ยความถูกต้องแม่นยำที่สูงสุดเท่ากับร้อยละ 90.8965 ± 0.85 ขณะเดียวกันยังเป็นวิธีที่ลู่เข้าค่าตอบที่เหมาะสมอย่างรวดเร็วที่สุดด้วยจำนวนรอบของการคำนวณน้อยที่สุดเท่ากับ 63.00 ± 0.81 รอบ และเมื่อเทียบผลการทดลองกับการทดลองที่ 2 ดังตารางที่ 4.9 จะเห็นได้ว่าวิธีการลดมิติแบบผสมผสาน DCCA-DisBGWO แสดงค่าความถูกต้องแม่นยำน้อยกว่าวิธีการเลือกคุณลักษณะ DisBGWO แต่จำนวนของคุณลักษณะที่ใช้ในการลู่เข้าหาค่าตอบน้อยกว่า จึงทำให้ค่าความเหมาะสมของทั้งสองวิธีมีค่าใกล้เคียงกัน นอกจากนี้ผลการทดลองแสดงให้เห็นว่าจำนวนรอบของการคำนวณลู่เข้าค่าตอบของวิธี DCCA-DisBGWO เร็วกว่าวิธีการเลือกคุณลักษณะ DisBGWO ประมาณ 2.5 เท่า ทำให้เวลาในการประมวลผลของขั้นตอนการเรียนรู้เร็วกว่าถึง 2.91 เท่า ดังนั้นจะเห็นได้ว่าการลดมิติด้วยวิธี DCCA-DisBGWO สามารถลดจำนวนคุณลักษณะและสามารถลู่เข้าหาค่าตอบได้เร็วขึ้น ซึ่งสามารถช่วยปรับปรุงเวลาที่ในขั้นตอนการเรียนรู้ได้ แต่อย่างไรก็ตามวิธีดังกล่าวยังคงต้องการการปรับปรุงในด้านของความถูกต้องแม่นยำของการจำแนกข้อมูลต่อไป

3. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการทดสอบการทดลองที่ 3

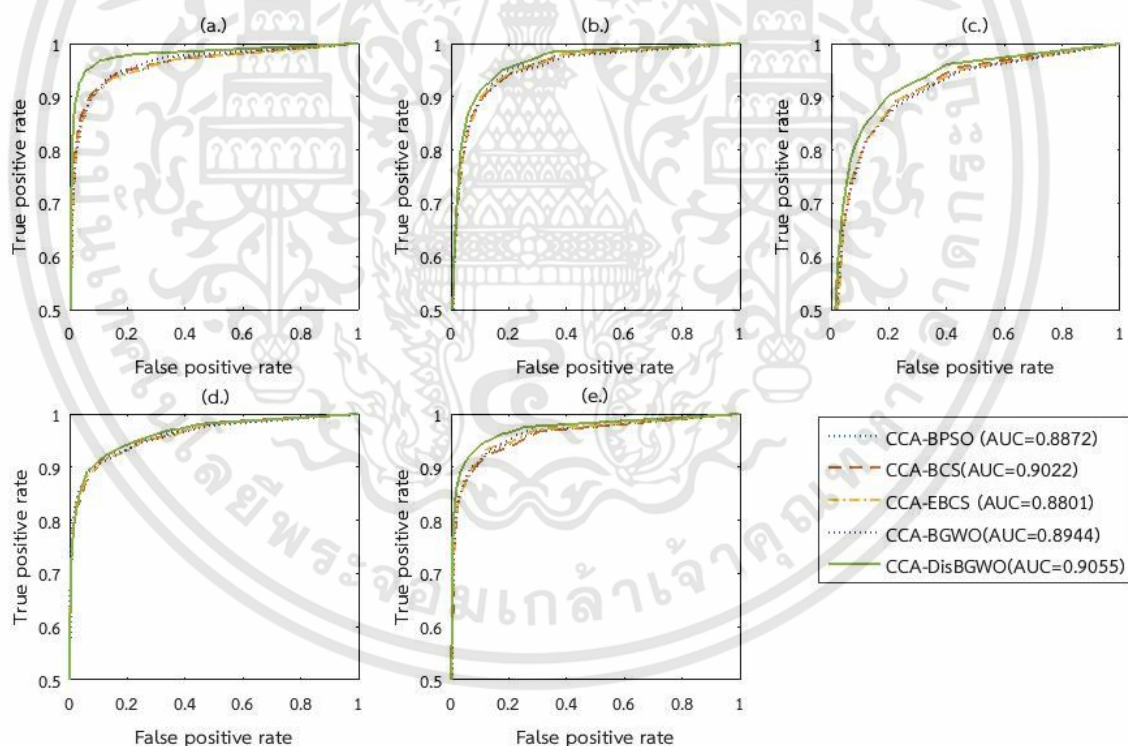
ผลการทดสอบประสิทธิภาพของวิธี CCA/DCCA-NIAs แสดงดังตารางที่ 4.12 ซึ่งแสดงด้วยค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะ ค่าเฉลี่ยเวลาของการประมวลผล และค่าเฉลี่ยฟังก์ชันความเหมาะสมของการจำแนกข้อมูลตามสมการ 4.1

ตารางที่ 4.12 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลดมิติแบบผสมผสานด้วยวิธี CCA-NIAs และ DCCA-NIAs

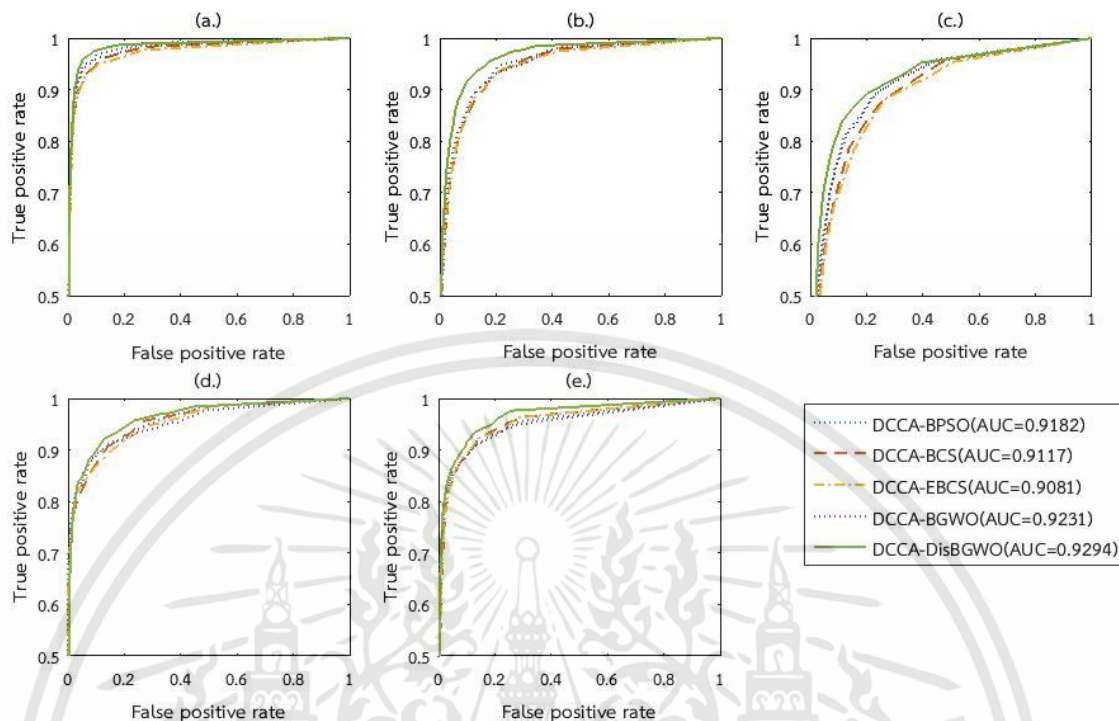
Algorithms	ArgAcc. (%±SD)	ArgSent. (%±SD)	ArgSpec. (%±SD)	ArgTime (sec.±SD)	Fitness _{classifier} (%±SD)
CCA-BPSO	86.9636±0.89	82.9762±0.97	91.7466±1.01	1,881.34±147.91	72.4756±0.95
CCA-BCS	87.5032±1.23	83.5568±0.99	90.6654±1.04	1,771.83±194.44	73.9023±0.57
CCA-EBCS	86.3634±1.34	82.5321±1.22	90.0442±1.12	1,742.23±206.85	73.2595±0.85
CCA-BGWO	87.4425±1.02	83.0856±0.86	91.6133±1.29	1,757.56±191.01	73.9834±0.73
CCA-DisBGWO	87.5722±0.89	83.2846±0.89	91.7952±0.78	1,734.39±172.43	74.2977±0.64
DCCA-BPSO	89.5377±0.79	85.8533±0.66	92.7553±0.96	2,201.21±211.46	71.6283±0.79
DCCA-BCS	89.4252±1.02	84.2964±1.03	93.8216±0.97	2,025.77±217.51	73.1324±0.88
DCCA-EBCS	88.5681±1.14	85.0457±1.21	92.0648±0.89	1,859.47±177.63	73.9579±0.67

DCCA-BGWO	89.5532±0.98	85.8981±1.06	92.5755±0.91	1,913.78±192.63	74.2525±0.77
DCCA-DisBGWO	90.1095±1.01	87.8563±0.88	92.27655±1.09	1,832.66±185.28	75.4346±0.54

จากผลการทดลองตารางที่ 4.12 เมื่อได้เซตคุณลักษณะที่ถูกเลือกด้วยวิธีต่างๆแล้วถูกนำมาทดสอบด้วยเซตคุณลักษณะสำหรับการทดสอบ ซึ่งผลการทดสอบแสดงให้เห็นว่า วิธีการลดมิติแบบผสมผสาน DCCA-DisBGWO เป็นวิธีที่แสดงค่าให้ประสิทธิภาพสูงสุดด้วยค่าเฉลี่ยความถูกต้องแม่นยำเท่ากับ 90.1095 ± 1.01 ค่าเฉลี่ยความไวเท่ากับร้อยละ 87.8563 ± 0.88 และค่าเฉลี่ยจำเพาะเท่ากับ 92.27655 ± 1.09 สำหรับเวลาในการประมวลผลของการทดสอบทุกอัลกอริทึมมีค่าใกล้เคียงกัน เมื่อเทียบกับผลการทดลองที่ 2 ดังตารางที่ 4.10 จะเห็นว่าค่าฟังก์ชันความเหมาะสมของการจำแนกของวิธี DCCA-BGWO น้อยกว่าและใช้เวลานานกว่าวิธีการเลือกคุณลักษณะ DisBGWO ทั้งนี้เนื่องจากการหลอมรวมคุณลักษณะด้วยฟังก์ชันการแปลงในขั้นตอนแรกทำให้คุณลักษณะมีคุณสมบัติเปลี่ยนไป ดังนั้นการค้นหาคุณลักษณะที่เหมาะสมจะถูกเลือกใหม่ทุกครั้ง จึงเป็นเป็นจุดด้อยของวิธีการลดมิติผสมผสานแบบที่ 3



รูปที่ 4.15 กราฟ ROC ของวิธีการลดมิติแบบผสมผสาน CCA-NIAs (CCA-BPSO, CCA-BCS, CCA-EBCS, CCA-BGWO และ CCA-DisBGWO) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.)



รูปที่ 4.16 กราฟ ROC ของวิธีการลดมิติแบบผสมผสาน DCCA-NIAs (DCCA-BPSO, DCCA-BCS, DCCA-EBCS, DCCA-BGWO และ DCCA-DisBGWO) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.)

ผลการทดลองกราฟ ROC ของแต่ละวิธีการของการลดมิติแบบผสมผสาน CCA-NIAs และ DCCA-NIAs ในการจำแนกในแต่ละระยะการนอน สามารถแสดงดังรูปที่ 4.15 และรูปที่ 4.16 ตามลำดับ ซึ่งผลการทดลองจะเห็นได้ว่าค่า AUC ของวิธีการลดมิติแบบผสมผสานด้วยของ DCCA-DisBGWO แสดงค่า ROC เท่ากับ 92.9421 ซึ่งเป็นค่าที่สูงที่สุดเมื่อเปรียบเทียบกับกราฟ ROC ของวิธีการลดมิติแบบผสมผสานอื่นๆ ทั้ง CCA-NIAs และ DCCA-NIAs

4. การทดลองที่ 4 การทดลองเปรียบเทียบประสิทธิภาพของการลดมิติแบบผสมผสานด้วยวิธี NIAs-CCA/DCCA สำหรับการจำแนกระยะการนอนโดยใช้สัญญาณ EEG EOG และ ECG

จากผลการทดลองที่ 2 แสดงให้เห็นว่าการเลือกคุณลักษณะมีข้อดี คือ เป็นวิธีที่มีความถูกต้องแม่นยำสูงสำหรับการจำแนกระยะการนอนด้วยการวิเคราะห์สัญญาณจากหลายแหล่งกำเนิด แต่มีข้อเสียคือ จำนวนรอบการคำนวณที่ใช้ในการลู่เข้าคำตอบที่เหมาะสมใช้จำนวนรอบมากทำให้เวลาในการประมวลผลนานด้วย ขณะเดียวกันการทดลองที่ 3 ได้นำข้อดีของการหลอมรวมแต่ละเซตคุณลักษณะด้วยความสามารถของ CCA/DCCA มาประยุกต์ใช้ ซึ่งสามารถแก้ไขเวลาในการลู่เข้าคำตอบในขั้นตอนการเรียนรู้ได้ แต่ความถูกต้องแม่นยำยังไม่สูงมากนัก เนื่องจากการหลอมรวมก่อนการเลือกคุณลักษณะทำให้

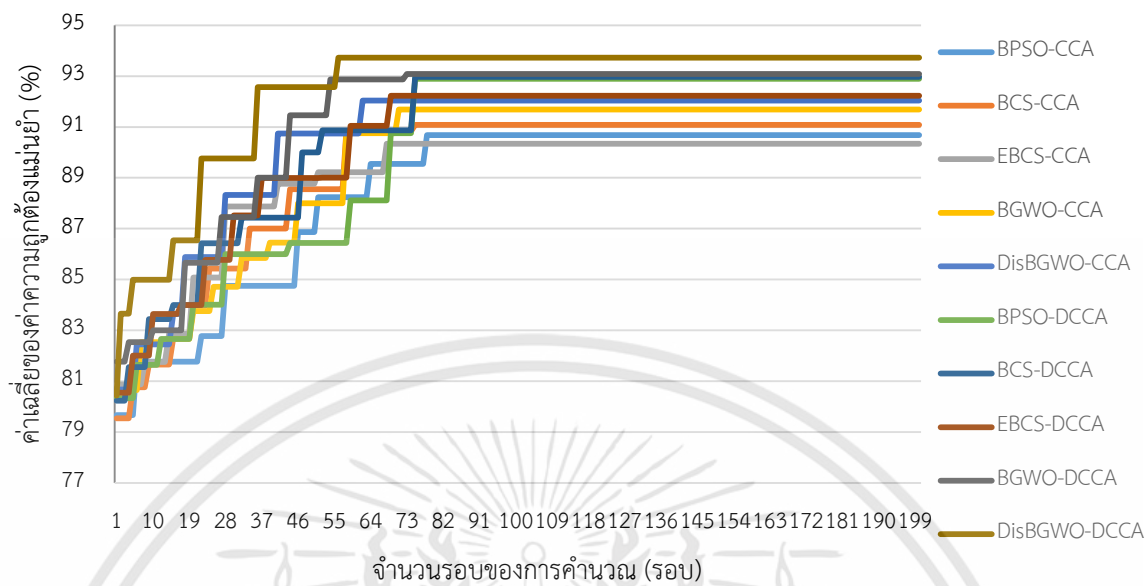
เกิดการสูญหายของข้อมูลบางส่วน นอกจากนี้ขั้นตอนการทดสอบยังใช้เวลานานอีกด้วย ดังนั้นเราจึงได้นำเสนอวิธีการลดมิติข้อมูลแบบผสมผสานแบบใหม่ ซึ่งใช้ข้อดีของการเลือกคุณลักษณะเพื่อลดเซตข้อมูลของการหลอมรวมคุณลักษณะ และขณะเดียวกันใช้วิธีการหลอมรวมคุณลักษณะระหว่างชุดข้อมูลช่วยลดความสามารถด้านเวลาการประมวลผลของวิธีการเลือกคุณลักษณะ วิธีการลดมิติแบบผสมผสานแบบที่ 4 ที่นำเสนอเรียกว่า NIAs-CCA/DCCA ประกอบด้วย 10 รูปแบบ ดังนี้ BPSO-CCA, BCS-CCA, EBCS-CCA, BGWO-CCA, DisBGWO-CCA, BPSO-DCCA, DCCA-BCS, EBCS-DCCA, BGWO-DCCA และ DisBGWO-DCCA เซตของคุณลักษณะที่ได้จากการทดลองหัวข้อที่ 4.1.1 ถูกนำไปใช้ในการทดลอง ซึ่งเซตของคุณลักษณะแต่ละแหล่งกำเนิดถูกสุ่มแบ่งเป็น 2 เซต ได้แก่ เซตข้อมูลสำหรับการเรียนรู้จำนวนร้อยละ 80 ของคุณลักษณะทั้งหมด และเซตของข้อมูลสำหรับการทดสอบจำนวนร้อยละ 20 ของคุณลักษณะทั้งหมด ขั้นตอนการทดลองมี 2 ขั้นตอนหลัก ขั้นตอนแรกคือขั้นตอนการเรียนรู้ ซึ่งเริ่มต้นด้วยแต่ละเซตคุณลักษณะของสัญญาณต่างๆ ถูกเลือกคุณลักษณะที่เหมาะสมด้วยวิธี NIAs จากนั้นแต่ละเซตคุณลักษณะที่ถูกเลือกถูกล้อมรวมด้วยวิธี CCA/DCCA แบบอนุกรมและเลือกขนาดของไอเกนเท่ากับค่าความเชื่อมั่นที่ร้อยละ 90 หลังจากนั้นเซตคุณลักษณะที่ถูกล้อมรวมแล้วจะเข้าสู่การประเมินด้วยฟังก์ชันความเหมาะสมของสมการ 3.2 เพื่อค้นหาวิธีการของการเลือกคุณลักษณะที่เหมาะสมที่สุด การทดลองนี้กำหนดให้ค่าคงที่ γ เท่ากับ 0.8 รายละเอียดของอัลกอริทึม NIAs-CCA/DCCA ตามหัวข้อที่ 3.1.2.3 ข้อ 2 จากนั้นจะเข้าสู่ขั้นตอนทดสอบประสิทธิภาพ โดยแต่ละวิธีถูกทดสอบด้วยเซตข้อมูลของการทดสอบ สิ่งที่จะพิจารณาความเหมาะสมของอัลกอริทึม คือ ประสิทธิภาพความถูกต้องแม่นยำของการจำแนกกระยะการนอน และเวลาที่ใช้ในการประมวลผล

จุดประสงค์การทดลอง เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมการลดมิติแบบผสมผสานที่นำเสนอ วิธี BPSO-CCA, BCS-CCA, EBCS-CCA, BGWO-CCA, DisBGWO-CCA, BPSO-DCCA, DCCA-BCS, EBCS-DCCA, BGWO-DCCA และ DisBGWO-DCCA

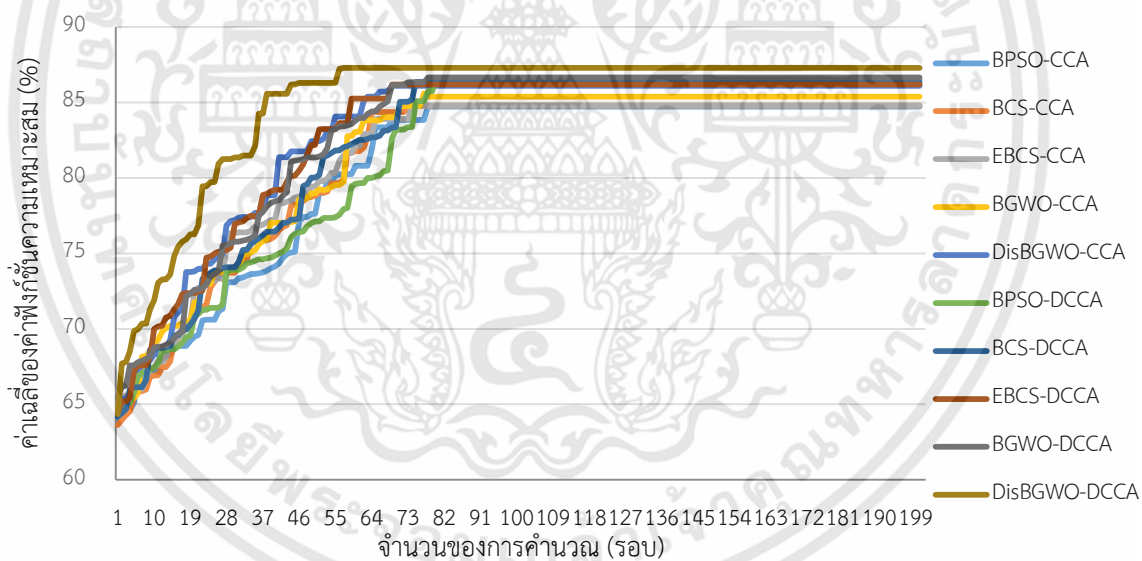
เงื่อนไขการทดลอง เงื่อนไขการทดลองของการลดมิติแบบผสมผสาน NIAs-CCA/DCCA คล้ายกับเงื่อนไขการทดลองที่ 3

5. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการเรียนรู้ของการทดลองที่ 4

ผลการทดลองขั้นตอนการเรียนรู้แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความถูกต้องแม่นยำของการจำแนกกระยะการนอนและจำนวนรอบของการคำนวณแสดงดังรูปที่ 4.17 และกราฟความสัมพันธ์ระหว่างค่าเฉลี่ยฟังก์ชันค่าความเหมาะสมและจำนวนรอบของการคำนวณแสดงดังรูปที่ 4.18 ผลการทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน NIAs-CCA/DCCA ในขั้นตอนการเรียนรู้ แสดงดังตารางที่ 4.13



รูปที่ 4.17 ความสัมพันธ์ระหว่างค่าเฉลี่ยของความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีการลดมิติแบบผสมผสาน NIAs-CCA/DCCA



รูปที่ 4.18 ความสัมพันธ์ระหว่างค่าเฉลี่ยของค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธีการลดมิติแบบผสมผสาน NIAs-CCA/DCCA

ตารางที่ 4.13 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน NIAs-CCA/DCCA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะโดยวิธีใดทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Algorithms	Fitness function value (%±SD)	The number of selected features (Number±SD)	Accuracy of selected features (%±SD)	The number of iterations (Number±SD)	Time (sec.±SD)
BPSO-CCA	84.7373±0.87	143.25±1.25	90.6744±0.65	79.00±1.15	3,572.53±131.64
BCS-CCA	84.7880±1.07	146.25±1.25	91.0811±0.49	74.50±1.29	3,463.76±124.74
EBCS-CCA	84.7947±0.87	138.50±1.29	90.3341±0.59	68.25±0.50	3,279.53±163.66
BGWO-CCA	85.3795±0.97	141.00±1.15	91.6832±0.58	70.50±1.91	3,224.87±146.04
DisBGWO-CCA	86.1054±0.44	138.75±0.95	92.0411±0.76	61.00±1.41	2,622.58±151.43
BPSO-DCCA	86.5115±0.87	143.50±1.29	92.8922±0.58	75.75±0.95	2,959.31±163.55
BCS-DCCA	86.3551±0.56	146.75±1.25	92.9714±0.77	77.00±1.41	3,024.34±142.75
EBCS-DCCA	86.1979±0.58	137.75±1.50	92.2255±0.87	69.25±0.95	2,796.03±140.39
BGWO-DCCA	86.6596±0.88	145.50±1.29	93.0772±0.67	73.25±1.50	2,742.92±152.81
DisBGWO-DCCA	87.2854±0.64	138.25±0.95	93.7222±0.48	59.75±0.95	2,325.76±153.76

จากผลการทดลองรูปที่ 4.17-4.18 และตารางผลการทดลองที่ 4.13 แสดงให้เห็นว่าวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA เป็นวิธีที่ลู่เข้าคำตอบที่เหมาะสมด้วยจำนวนรอบของการลู่เข้าที่น้อยที่สุดเท่ากับ 67.75 ± 0.95 รอบ ขณะเดียวกันวิธีการลดมิติแบบผสมผสาน DCCA-DisBGWO เป็นวิธีที่มีค่าเฉลี่ยของฟังก์ชันความเหมาะสมสูงสุดเท่ากับร้อยละ 87.2854 ± 0.64 และค่าเฉลี่ยของค่าความถูกต้องแม่นยำที่สุดเท่ากับร้อยละ 93.8222 ± 0.48 เมื่อเปรียบเทียบผลการทดลองกับผลการทดลองที่ 2 ดังตารางที่ 4.9 จะเห็นได้ว่า ค่าฟังก์ชันความเหมาะสมของวิธีการ DisBGWO-DCCA มีค่าสูงกว่าวิธี DisBGWO ซึ่งเป็นวิธีที่เหมาะสมที่สุดของวิธีการเลือกคุณลักษณะ นอกจากนี้จะเห็นได้ว่าการลู่เข้าของคำตอบที่เหมาะสมของวิธี DCCA-DisBGWO ใช้จำนวนในการลู่เข้าคำตอบที่เหมาะสมน้อยกว่าวิธีการเลือกคุณลักษณะ DisBGWO 2.76 เท่า และเวลาในการประมวลผลในขั้นตอนของการเรียนรู้ของเร็วกว่าถึง 6.14 เท่า ขณะเดียวกันเมื่อเปรียบเทียบกับผลการทดลองที่ 3 ดังตารางที่ 4.11 จะเห็นได้ว่าวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA แสดงค่าฟังก์ชันความเหมาะสมที่สูงกว่าวิธีการลดมิติแบบผสมผสาน DCCA-NIAS ซึ่งเป็นวิธีที่เหมาะสมที่สุดของการทดลองที่ 3 นอกจากนี้วิธี DisBGWO-DCCA สามารถลู่เข้าคำตอบที่เหมาะสมโดยใช้จำนวนรอบการลู่เข้าและเวลาในการประมวลผลน้อยกว่าอีกด้วย ซึ่งจากการเปรียบเทียบผลการทดลองดังกล่าวจะเห็นได้ว่าวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สามารถ

ปรับปรุงประสิทธิภาพของการจำแนกกระยะการนอนในขั้นตอนการเรียนรู้ทั้งค่าความถูกต้องแม่นยำและเวลาในการประมวลผลได้อย่างมีประสิทธิภาพ

6. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนการทดสอบการทดลองที่ 4

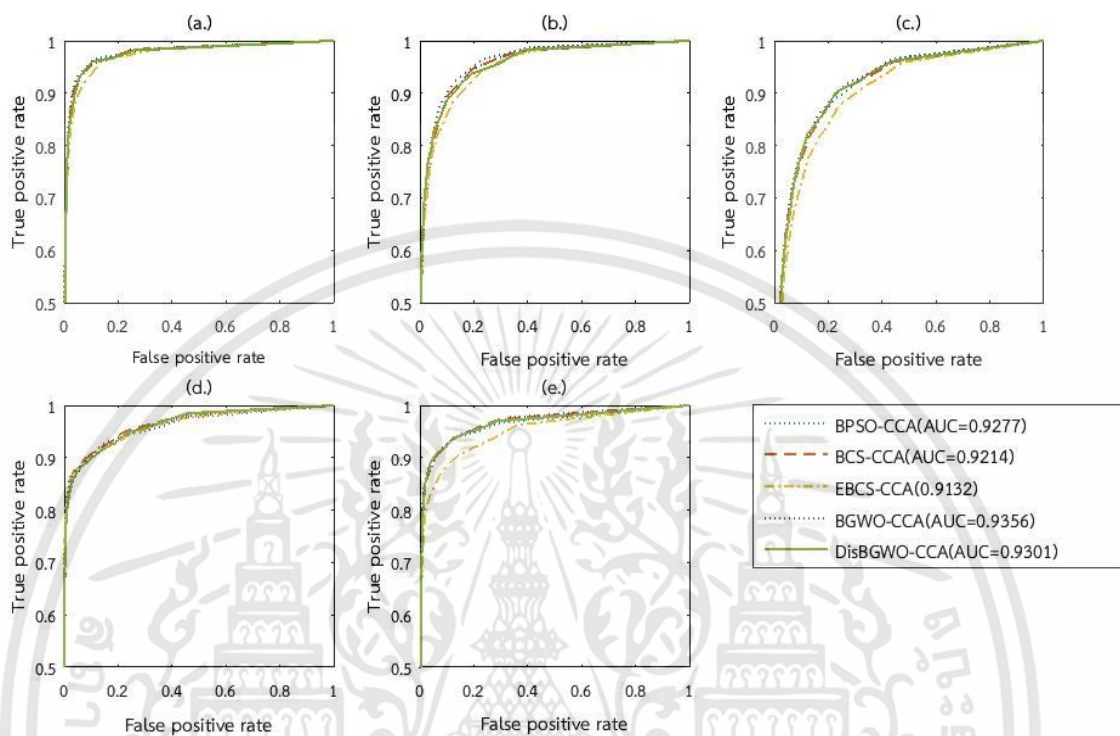
ผลการทดสอบประสิทธิภาพการจำแนกกระยะการนอนของวิธีการลดมิติแบบผสมผสาน CCA/DCCA-NIAs แสดงดังตารางที่ 4.14 ซึ่งแสดงด้วยค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของความไว ค่าเฉลี่ยของค่าจำเพาะ ค่าเฉลี่ยเวลาของการประมวลผล และฟังก์ชันความเหมาะสมของการจำแนกข้อมูลตามสมการที่ 4.1

ตารางที่ 4.14 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีการลดมิติแบบผสมผสาน ด้วยวิธี NIAs-CCA และ NIAs-DCCA

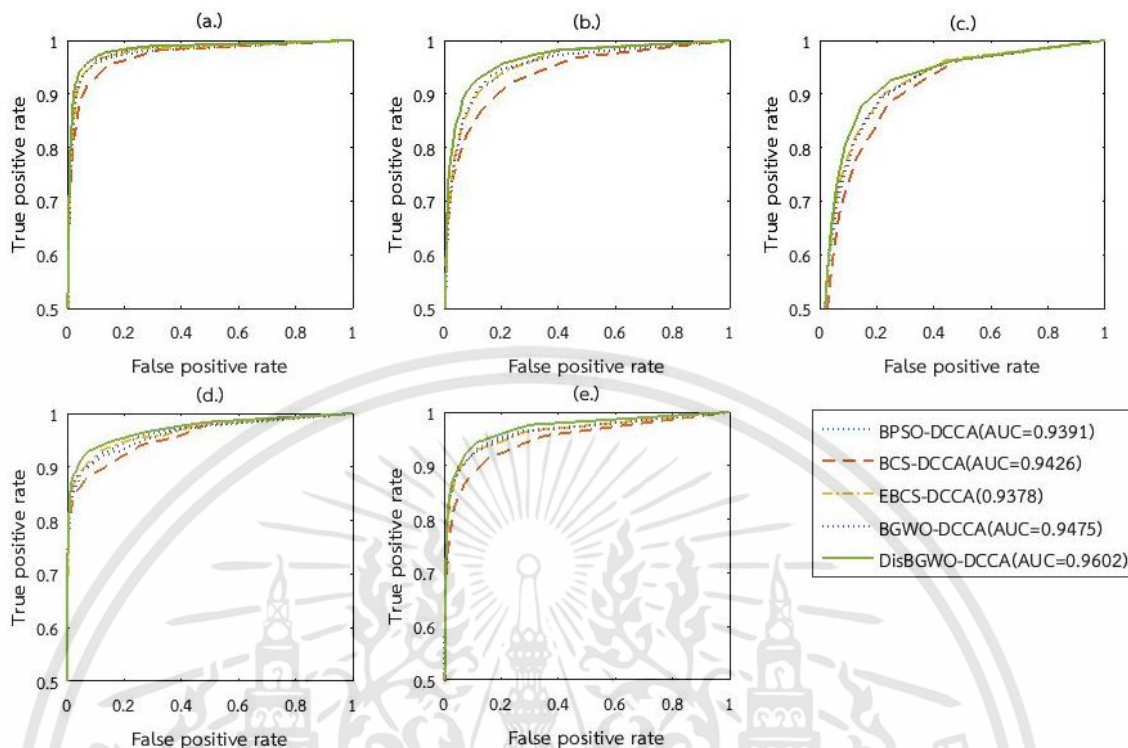
Algorithms	ArgAcc. (%±SD)	ArgSent. (%±SD)	ArgSpec. (%±SD)	ArgTime (sec.±SD)	Fitness _{classifier} (%±SD)
BPSO-CCA	90.0976±0.73	87.4557±0.65	92.5776±0.71	29.7431±4.32	72.076±0.57
BCS-CCA	90.4043±0.86	87.8954±0.59	92.7578±0.68	29.6269±5.31	72.3995±0.63
EBCS-CCA	89.7573±0.81	87.1054±0.92	92.0696±0.82	28.3296±4.12	72.7543±0.95
BGWO-CCA	92.0164±0.84	89.7892±0.61	93.8656±0.75	29.4168±4.92	74.503±0.94
DisBGWO-CCA	91.4643±0.76	90.6965±0.77	92.0643±1.11	24.7126±5.02	76.5523±0.80
BPSO-DCCA	92.1154±0.77	89.0311±0.53	94.8754±0.69	26.1295±4.71	76.1204±0.72
BCS-DCCA	92.2946±0.59	89.8765±0.88	93.9775±0.74	26.8961±4.22	75.7482±0.67
EBCS-DCCA	91.7487±0.80	89.5442±0.76	93.8655±1.02	26.8961±4.68	75.3115±0.55
BGWO-DCCA	92.5004±0.94	90.9344±0.59	93.3321±0.82	25.6637±5.12	76.7416±0.71
DisBGWO-DCCA	93.2854±0.97	91.7614±0.49	93.7466±0.94	24.4166±4.92	78.2083±0.63

จากผลการทดลองตารางที่ 4.14 แสดงให้เห็นว่า วิธีการลดมิติแบบผสมผสาน DisBGWO-NIAs เป็นวิธีที่แสดงค่าประสิทธิภาพสูงสุด ด้วยค่าเฉลี่ยความถูกต้องแม่นยำเท่ากับ 93.2854±0.97 ค่าความไวเท่ากับร้อยละ 91.7614±0.49 และค่าจำเพาะเท่ากับ 93.7466±0.94 สำหรับเวลาในการประมวลผลของการทดสอบทุกวิธีมีค่าใกล้เคียงกัน และเมื่อเปรียบเทียบผลการทดลองที่ 3 ดังตารางที่ 4.12 จะเห็นได้ว่า นอกจากค่าประสิทธิภาพความถูกต้องของวิธี DisBGWO-DCCA มีค่าสูงกว่าวิธี DCCA-DisBGWO แล้ว เวลาที่ใช้ในการประมวลผลของขั้นตอนการทดสอบยังใช้เวลาน้อยกว่า 75.05 เท่า นอกจากนี้เมื่อเทียบกับค่าความเหมาะสมของการจำแนกข้อมูลของผลการทดสอบของการทดลองที่ 1-3 จะเห็นได้ว่า ค่าความเหมาะสมของการจำแนกของวิธี DisBGWO-DCCA มีค่าสูงสุด ดังนั้นจะเห็นได้ว่าการลดมิติแบบผสมผสานด้วย

วิธีที่นำเสนอ DisBGWO-DCCA สามารถช่วยแก้ไขทั้งเวลาในการประมวลผลและเพิ่มประสิทธิภาพความถูกต้องแม่นยำของการจำแนกระยะการนอนที่หลายชุดสัญญาณได้อย่างมีประสิทธิภาพและเหมาะสม



รูปที่ 4.19 กราฟ ROC ของการลดมิติแบบผสมผสาน NIAS-CCA (BPSO-CCA, BCS-CCA, EBCS-CCA, BGWO-CCA และ DisBGWO-CCA) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.)



รูปที่ 4.20 กราฟ ROC ของการลตมิติแบบผสมผสาน NIAS-DCCA (BPSO-DCCA, BCS-DCCA, EBSC-DCCA, BGWO-DCCA และ DisBGWO-DCCA) ในการจำแนก Wake (a.), REM (b.), NREM1 (c.), NREM2 (d.) และ NREM3 (e.)

ผลการทดลองกราฟ ROC ของแต่ละวิธีการของการลตมิติแบบผสมผสาน NIAS-CCA และ NIAS-DCCA ในการจำแนกในแต่ละระยะการนอน สามารถแสดงดังรูปที่ 4.19 และรูปที่ 4.20 ตามลำดับ ซึ่งผลการทดลองจะเห็นได้ว่าค่า AUC ของวิธีการลตมิติแบบผสมผสานด้วยของ DisBGWO-DCCA แสดงค่า ROC เท่ากับ 0.9602 ซึ่งเป็นค่าที่สูงที่สุดเมื่อเปรียบเทียบกับกราฟ ROC ของวิธีการลตมิติแบบผสมผสานอื่นๆ ทั้ง NIAS-CCA และ NIAS-DCCA

7. การทดลองที่ 4.1 การทดลองเปรียบเทียบประสิทธิภาพของการลตมิติแบบผสมผสานของ DisBGWO-DCCA ด้วยเงื่อนไขการแบ่งข้อมูลที่ต่างกัน

จากการทบทวนวรรณกรรมที่ผ่านมาพบว่าเงื่อนไขในการแบ่งข้อมูลสำหรับการออกแบบเพื่อสร้างอัลกอริทึม มี 2 ประเภท ได้แก่ การแบ่งข้อมูลสำหรับการสร้างอัลกอริทึมโดยอ้างอิงตามเหตุการณ์ไม่ขึ้นกับบุคคล (Based on event) ซึ่งการแบ่งข้อมูลประเภทนี้เป็นการออกแบบทั่วไปสำหรับระบบการเรียนรู้ของเครื่อง และการแบ่งข้อมูลสำหรับการสร้างอัลกอริทึมโดยอ้างอิงตามตัวอย่าง (Based on

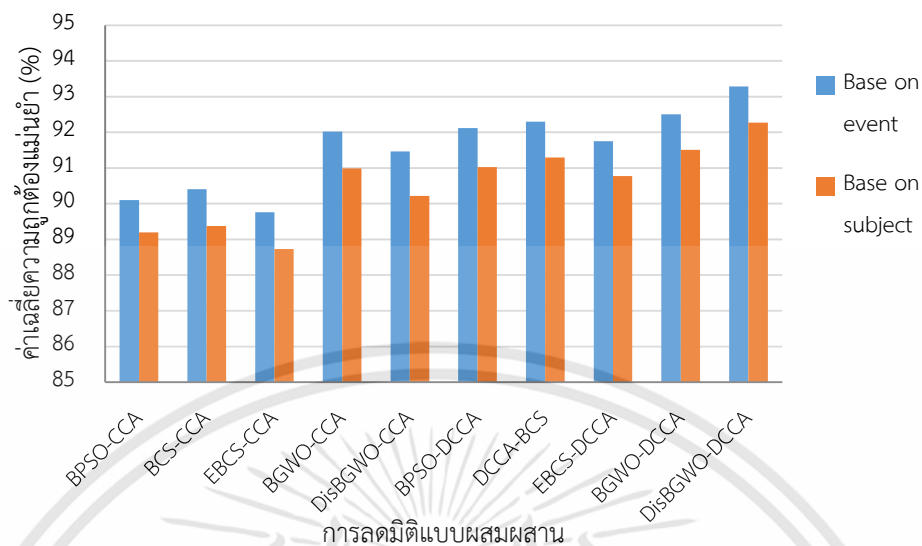
Subject) ซึ่งเป็นเทคนิคสำหรับการทดสอบประสิทธิภาพของงานวิจัยทางการแพทย์ ดังนั้นการทดลองนี้จึงมีจุดประสงค์ เพื่อแสดงความแตกต่างและเปรียบเทียบของเงื่อนไขการแบ่งข้อมูลทั้งสองประเภท

8. ผลการทดลองและวิเคราะห์ผลการทดลองที่ 4.1

ตารางผลการทดลองที่ 4.14 ของการทดลองที่ 4 แสดงผลการทดสอบของประสิทธิภาพการลดมิติด้วย DisBGWO-BGWO ด้วยเงื่อนไขการแบ่งข้อมูลการแบ่งข้อมูลตามวิธี Based on event และตารางผลการทดลองที่ 4.15 แสดงผลการทดลองของประสิทธิภาพการลดมิติด้วย DisBGWO-BGWO ด้วยเงื่อนไขการแบ่งข้อมูลการแบ่งตามวิธี Based on Subject และนำผลจากตารางที่ 4.14 และ 4.15 มาพล็อตกราฟดังรูปที่ 4.21 เพื่อเปรียบเทียบผลการทดลองของการแบ่งกลุ่มข้อมูลด้วยเงื่อนไขทั้งสองประเภท

ตารางที่ 4.15 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีลดมิติแบบผสมผสาน NIAs-CCA/DCCA ด้วยการแบ่งข้อมูลตามวิธี Based on subject

Algorithms	Acc. (%±SD)	Sent. (%±SD)	Spec. (%±SD)	Time (sec.±SD)	Fitness <i>classifier</i> (%±SD)
BPSO-CCA	89.1963±0.58	86.3769±0.66	91.5868±0.84	28.24±3.97	71.8879±1.93
BCS-CCA	89.3797±0.74	87.1012±0.73	91.7438±0.85	29.01±4.02	71.5038±1.67
EBCS-CCA	88.7327±0.65	86.2390±0.82	91.0552±0.57	27.83±3.65	71.7997±1.77
BGWO-CCA	90.9889±0.81	88.9020±0.49	92.8496±0.63	28.01±4.21	73.4805±2.05
DisBGWO-CCA	90.2193±1.03	89.5750±0.66	90.8308±0.89	26.56±4.69	73.8645±2.28
BPSO-DCCA	91.0284±0.49	87.7429±0.68	93.8009±0.91	27.01±3.89	74.2016±1.89
BCS-DCCA	91.2971±0.65	89.4647±0.72	92.9905±0.73	27.89±4.13	73.8098±2.11
EBCS-DCCA	90.7703±0.79	88.4713±0.54	92.9006±0.93	27.52±3.63	73.6435±1.76
BGWO-DCCA	91.5062±0.81	89.8607±0.93	92.3464±0.75	26.78±3.85	74.7424±1.73
DisBGWO-DCCA	92.2732±0.97	90.67854±0.49	92.7567±0.83	26.03±4.11	75.8729±1.95



รูปที่ 4.21 กราฟเปรียบเทียบประสิทธิภาพของการลดมิติ DisBGWO-DCCA ของเงื่อนไขการแบ่งข้อมูล สำหรับการออกแบบอัลกอริทึมด้วยการแบ่งข้อมูล Based on event และ Based on subject

จากการทดลองเปรียบเทียบผลการทดลองของการแบ่งข้อมูลด้วยเงื่อนไขทั้งสองประเภทจะเห็นได้ว่าการแบ่งข้อมูลด้วยวิธี Based on Subject แสดงค่าประสิทธิภาพการจำแนกกระยะการนอนน้อยกว่าการแบ่งข้อมูลด้วยวิธี Based on event ด้วยค่าเฉลี่ยความต่าง 1.02 ดังนั้นจะเห็นได้ว่าการแบ่งข้อมูลของออกแบบอัลกอริทึมที่แตกต่างกันแสดงค่าประสิทธิภาพที่แตกต่างกัน แต่อย่างไรก็ตามในกรณีการทดลองที่มีจุดประสงค์เพื่อเปรียบเทียบประสิทธิภาพของอัลกอริทึมต่างๆ สามารถเลือกแบ่งข้อมูลได้อย่างใดอย่างหนึ่ง แต่การแบ่งข้อมูลของทุกอัลกอริทึมในการทดลองนั้นจะต้องกำหนดการแบ่งข้อมูลรูปแบบที่เหมือนกันจึงจะสามารถนำมาเปรียบเทียบกันได้

9. การทดลองที่ 5 การทดสอบประสิทธิภาพการลดมิติแบบผสมผสานของ DisBGWO-DCCA สำหรับการจำแนกกระยะการนอนจากสัญญาณ EEG EOG และ ECG

การทดลองนี้มีจุดประสงค์เพื่อทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกกระยะการนอนกับข้อมูลการนอนที่ได้มาจากแหล่งข้อมูลอื่น คือ ข้อมูลการนอนที่มาจากฐานข้อมูล Dream database [13] มีจำนวนทั้งหมด 47 ตัวอย่าง ประกอบด้วย ข้อมูล PSG ของคนสุขภาพดีจำนวน 20 ตัวอย่าง และ ข้อมูล PSG ของผู้ป่วยที่มีความผิดปกติของการนอนจำนวน 27 ตัวอย่าง สัญญาณที่ใช้เพื่อจำแนกกระยะการนอน ได้แก่ สัญญาณ EEG จำนวน 3 ช่องสัญญาณ สัญญาณ EOG จำนวน 2 ช่องสัญญาณ และสัญญาณ ECG จำนวน 1 สัญญาณ ที่ความถี่สุ่มเท่ากับ 125 เฮิรตซ์ สัญญาณแต่ละช่องสัญญาณถูกแบ่งออกเป็นช่วงๆ ช่วงละ 30 วินาทีตามช่วงเวลา (Epoch) ของระยะเวลา

การนอน จำนวนตัวอย่างของระยะการนอนต่างๆ ของฐานข้อมูล Dream database แสดงดังตารางที่ 4.16

ตารางที่ 4.16 จำนวนตัวอย่างของระยะ Wake, ระยะ REM และระยะ NREM1-3 ที่มาจากฐานข้อมูล Dream Database

แหล่งข้อมูล	ระยะ Wake (Epoch)	ระยะ REM (Epoch)	ระยะ NREM1 (Epoch)	ระยะ NREM2 (Epoch)	ระยะ NREM3 (Epoch)
Dream Database	10,862	9,521	8,941	9,557	11,097

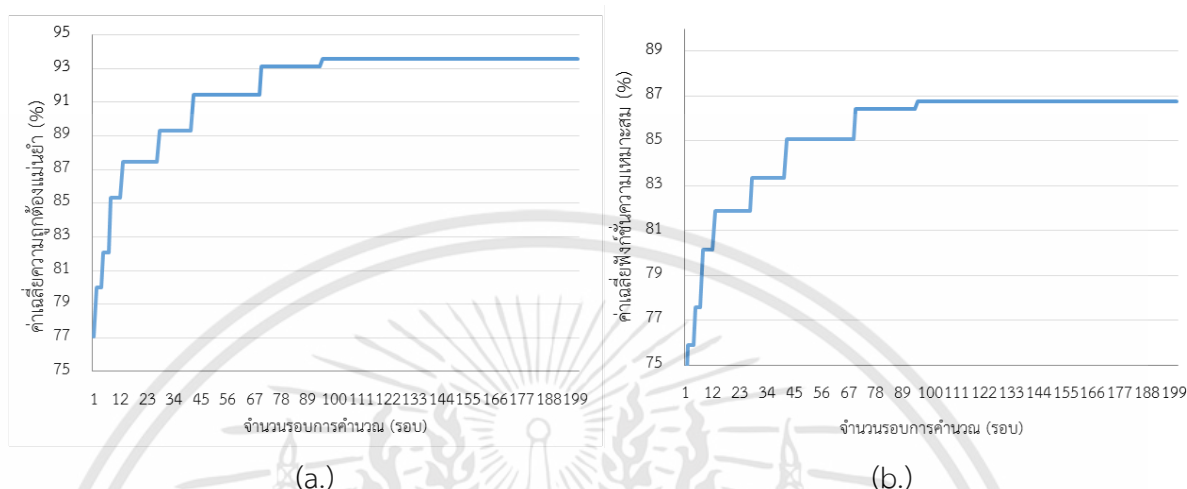
ก่อนเข้าสู่กระบวนการสกัดคุณลักษณะสัญญาณ แต่ละสัญญาณถูกประมวลผลสัญญาณก่อนเพื่อกรองสัญญาณรบกวน จากนั้นแต่ละสัญญาณถูกสกัดคุณลักษณะตามรายละเอียดหัวข้อที่ 3.1.1 จำนวนทั้งสิ้น 445 คุณลักษณะ ประกอบด้วย 3 ช่องสัญญาณของ EEG จำนวน $81 \times 3 = 243$ คุณลักษณะ 2 ช่องสัญญาณของ EOG จำนวน $73 \times 2 = 146$ คุณลักษณะ และ 1 ช่องสัญญาณของ ECG ซึ่งเป็นคุณลักษณะที่มาจากวิเคราะห์ HRV จำนวน $56 \times 1 = 56$ คุณลักษณะ เซตของคุณลักษณะที่ได้ในแต่ละสัญญาณถูกแบ่งออกเป็น 2 ส่วน คือ เซตข้อมูลสำหรับการเรียนรู้ร้อยละ 80 และเซตของข้อมูลสำหรับการทดสอบร้อยละ 20 ขั้นตอนการเรียนรู้เริ่มต้นด้วยเซตข้อมูลสำหรับการเรียนรู้ถูกแบ่งออกเป็น 2 เซตข้อมูลอีกครั้ง ประกอบด้วย เซตของข้อมูลสำหรับการสร้างโมเดล และเซตข้อมูลสำหรับการปรับแต่ง ซึ่งแต่ละเซตคุณลักษณะถูกเลือกคุณลักษณะที่เหมาะสมด้วยวิธีการเลือกคุณลักษณะ DisBGWO จากนั้นแต่ละเซตคุณลักษณะที่ถูกเลือกหลอมรวมด้วยวิธี CCA/DCCA ซึ่งหลอมรวมทีละคู่จำนวนทั้งหมด 15 คู่ที่มีการหลอมรวมแบบอนุกรม และเลือกขนาดของไอแกนเท่ากับค่าความเชื่อมั่นที่ร้อยละ 90 ขั้นตอนของการลดมิติแบบผสมผสาน DisBGWO-DCCA ตามอัลกอริทึมรูปที่ 3.10 สำหรับขั้นตอนการทดสอบจะเริ่มต้นเมื่อได้เซตของคุณลักษณะที่เหมาะสม และจะถูกทดสอบประสิทธิภาพวิธี DisBGWO-DCCA ด้วยเซตข้อมูลของการทดสอบ

เงื่อนไขการทดลอง เงื่อนไขในการทดลองของการทดลองนี้เหมือนกับเงื่อนไขการทดลองที่ 4 แต่การกำหนดจำนวนค่าคำตอบในแต่ละวิธีการค้นหาคำตอบเท่ากับ 445 ซึ่งเท่ากับจำนวนของคุณลักษณะทั้งหมดในการทดลองนี้

10. ผลการทดลองและวิเคราะห์ผลการทดลองขั้นตอนการเรียนรู้ของการทดลองที่ 5

ผลการทดลองขั้นตอนการเรียนรู้แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความถูกต้องแม่นยำของการจำแนกระยะการนอนและจำนวนรอบของการคำนวณแสดงดังรูปที่ 4.22(a.) และกราฟความสัมพันธ์ระหว่างค่าเฉลี่ยฟังก์ชันค่าความเหมาะสมและจำนวนรอบของการคำนวณดังรูปที่ 4.22(b.)

ผลการทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA ในขั้นตอนการเรียนรู้ แสดงดังตารางที่ดังตารางที่ 4.17



รูปที่ 4.22 (a.) ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน DisBGWO-DCCA (b.) ความสัมพันธ์ระหว่างค่าเฉลี่ยค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธีมิติแบบผสมผสาน DisBGWO-DCCA

ตารางที่ 4.17 ผลการทดลองของขั้นตอนการเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกกระยะการนอน

Algorithms	Fitness function value (%±SD)	The number of selected features (Number±SD)	Accuracy of selected features (%±SD)	The number of iterations (Number±SD)	Time (sec.±SD)
DisBGWO-DCCA	86.8631±0.32	175.50±2.64	93.0384±0.45	97.75±2.50	31.4281±3.39

จากตารางผลการทดลองที่ 4.17 จะเห็นได้ว่าการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกกระยะการนอนแสดงค่าเฉลี่ยฟังก์ชันความเหมาะสมเท่ากับร้อยละ 86.8631 ค่าเฉลี่ยจำนวนคุณลักษณะที่ถูกเลือกเท่ากับ 175.50±2.64 ค่าเฉลี่ยความถูกต้องแม่นยำของการจำแนกกระยะการนอนเท่ากับร้อยละ 93.0384±0.45 จำนวนรอบการคำนวณเท่ากับ 97.75±2.50 รอบและเวลาในการประมวลผลเท่ากับ 31.4281±3.39 วินาที ซึ่งผลการทดสอบมีความใกล้เคียงกับผลการทดลองที่ 4 ซึ่งแสดงให้เห็นว่าวิธี DisBGWO-DCCA มีประสิทธิภาพเมื่อนำไปใช้กับฐานข้อมูลอื่นๆ

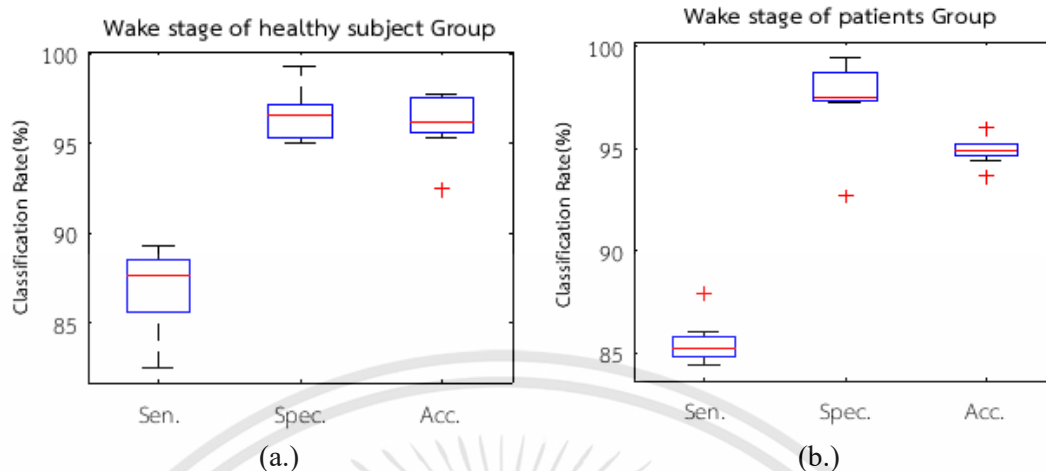
11. ผลการทดลองและวิเคราะห์ผลการทดลองของขั้นตอนทดสอบของการทดลองที่ 5

ประสิทธิภาพของวิธีการลตมิติแบบผสมผสาน DisBGWO-DCCA ของการจำแนกกระยะการนอนแต่ละระยะได้แก่ Wake, REM, NREM1-3 ของขั้นตอนการทดสอบแสดงดังตารางที่ 4.18 ซึ่งแสดงค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะของแต่ละระยะการนอน

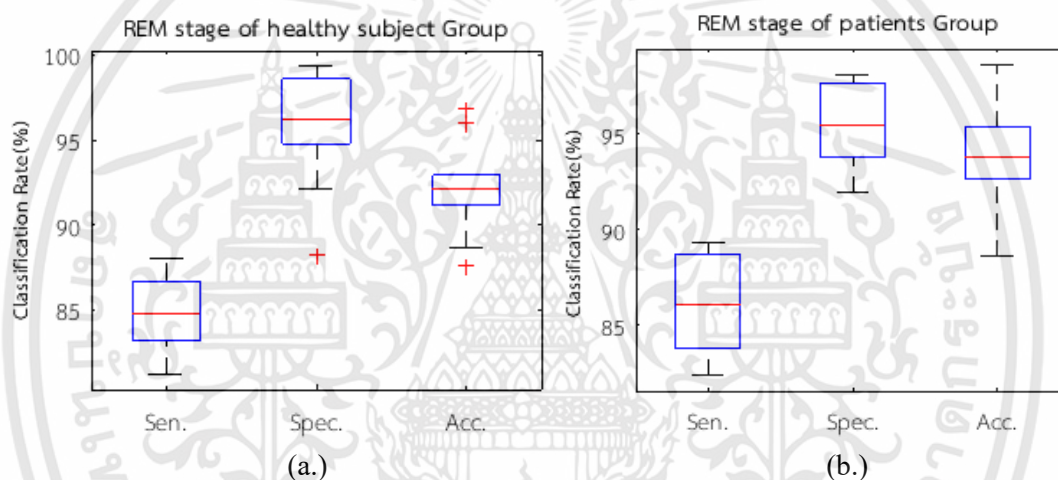
ตารางที่ 4.18 ผลการทดลองของขั้นตอนทดสอบการจำแนกกระยะการนอนด้วยวิธีการลตมิติแบบผสมผสาน DisBGWO-DCCA

Sleep stages	Acc. (%±SD)	Sent. (%±SD)	Spec. (%±SD)
Wake	95.8155±0.63	82.3654±0.46	97.6132±0.82
REM	94.6281±0.49	89.4724±0.52	97.6345±0.71
NREM1	89.4523±0.69	78.8511±0.52	96.6535±0.62
NREM2	90.9134±0.73	82.6341±0.77	94.3476±0.41
NREM3	94.7822±0.82	84.3214±0.64	96.2213±0.88
รวม	93.1183±0.76	84.1288±0.87	96.6940±0.65

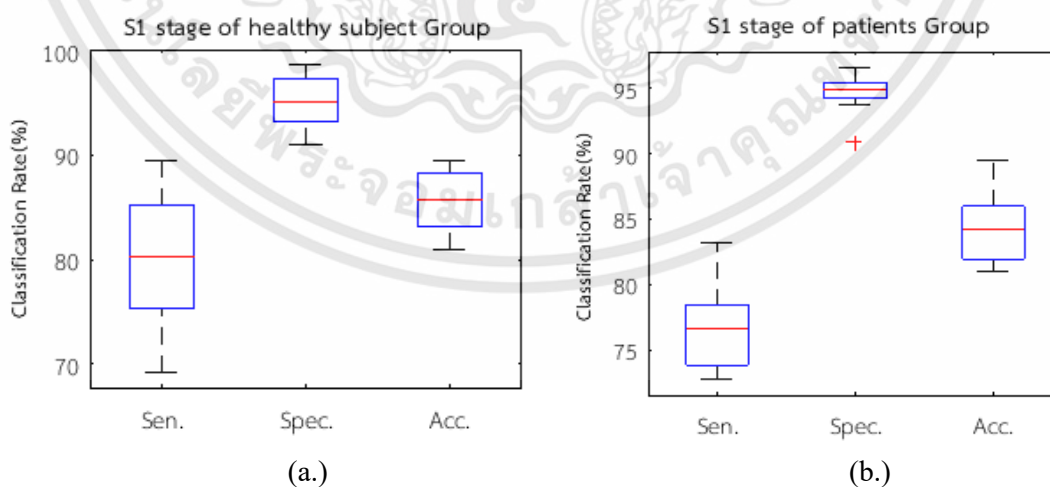
จากตารางที่ 4.18 ผลการทดสอบของวิธีการลตมิติแบบผสมผสาน DisBGWO-DCCA แสดงค่าเฉลี่ยความถูกต้องแม่นยำเท่ากับร้อยละ 93.1183±0.76 ค่าเฉลี่ยของค่าความไวเท่ากับร้อยละ 84.1288±0.87 และค่าเฉลี่ยของค่าจำเพาะเท่ากับร้อยละ 96.6940±0.65 ซึ่งผลการทดลองดังกล่าวค่อนข้างสูง ถึงแม้ว่าช่วง NREM1 จะมีค่าไม่สูงมากนัก ทั้งนี้เนื่องจากช่วง NREM1 เป็นช่วงเปลี่ยนจากการตื่นไปยังระยะหลับลึกในช่วงต้นของการนอนจึงทำให้คุณลักษณะของช่วง NREM1 และ Wake มีความคล้ายคลึงกัน นอกจากนี้ลักษณะทางกายภาพของสัญญาณของช่วง NREM1 และ NREM2 ยังมีลักษณะที่คล้ายกันอีกด้วย ดังนั้นจึงทำให้เกิดความผิดพลาดในการจำแนก ซึ่งผลการทดลองดังกล่าวคล้ายกับงานวิจัยที่ผ่านมา [51,53] แต่อย่างไรก็ตามเมื่อเปรียบเทียบกับผลการทดลองของงานวิจัยเหล่านั้นแล้วผลของประสิทธิภาพของวิธีที่งานวิจัยนี้ได้นำเสนอแสดงค่าความถูกต้องแม่นยำสูงสุด



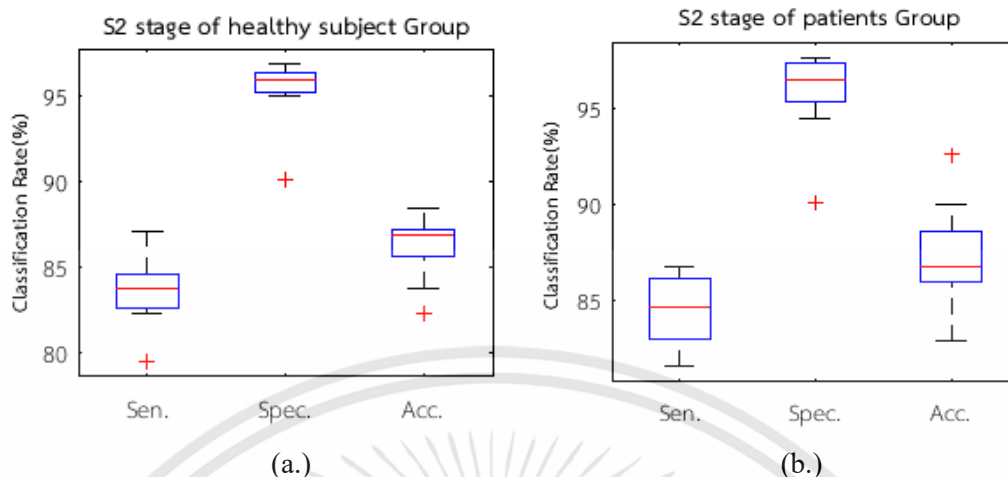
รูปที่ 4.23 กราฟ Boxplot ค่าประสิทธิภาพของระยะ wake ของกลุ่มคนปกติ (a.) และกลุ่มคนป่วย (b.)



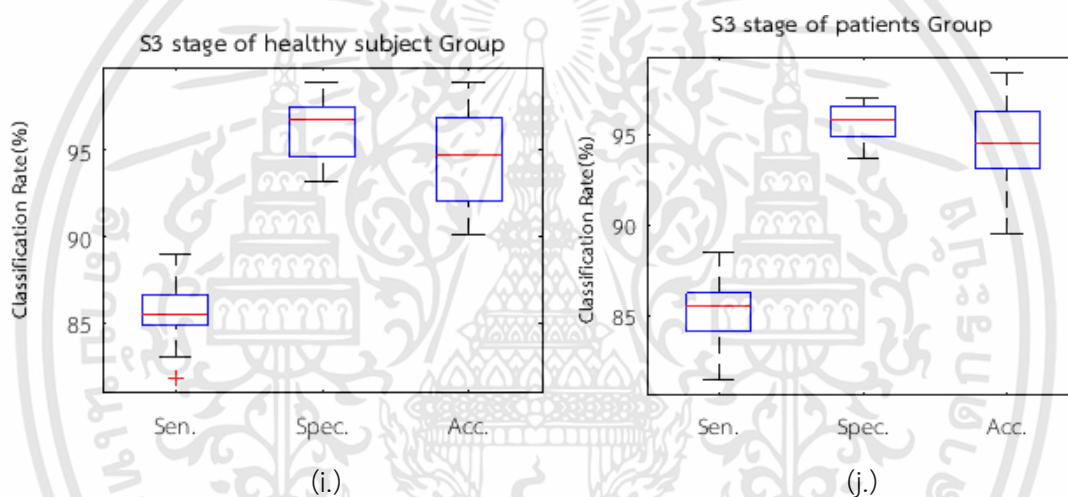
รูปที่ 4.24 กราฟ Boxplot ค่าประสิทธิภาพของระยะ REM ของกลุ่มคนปกติ (a.) และกลุ่มคนป่วย (b.)



รูปที่ 4.25 กราฟ Boxplot ค่าประสิทธิภาพของระยะ NREM1 ของกลุ่มคนปกติ(a.) และกลุ่มคนป่วย(b.)



รูปที่ 4.26 กราฟ Boxplot ค่าประสิทธิภาพของระยะ NREM2 ของกลุ่มคนปกติ(a.) และกลุ่มผู้ป่วย(b.)



รูปที่ 4.27 กราฟ Boxplot ค่าประสิทธิภาพของระยะ NREM3 ของกลุ่มคนปกติ(a.) และกลุ่มผู้ป่วย(b.)

ผลการทดสอบประสิทธิภาพการจำแนกระยะการนอนด้วยการลดมิติแบบผสมผสาน DisBGWO-DCCA รูปที่ 4.23 – 4.27 แสดงการเปรียบเทียบประสิทธิภาพของการจำแนกระยะการนอนต่างๆด้วยการลดมิติแบบผสมผสาน DisBGWO-DCCA ระหว่างข้อมูล PSG ของคนปกติและผู้ป่วยที่มีปัญหาการนอนจะเห็นได้ว่า กราฟของทั้งสองกลุ่มแสดงค่าประสิทธิภาพการจำแนกระยะการนอนในแต่ละระยะการนอนที่ไม่แตกต่างกันมาก ดังนั้นจึงแสดงให้เห็นว่าการจำแนกระยะการนอนด้วย DisBGWO-DCCA ซึ่งเป็นอัลกอริทึมที่งานวิจัยนี้นำเสนอสามารถประยุกต์ใช้ได้ทั้งการวิเคราะห์ระยะการนอนของคนปกติและผู้ป่วยได้อย่างมีประสิทธิภาพ

4.2 การทดลองจำแนกภาวะหยุดหายใจขณะนอนหลับแบบอัตโนมัติ

เนื่องจากการจำแนกภาวะหยุดหายใจขณะนอนหลับแบบอัตโนมัติในงานวิจัยที่ผ่านมาส่วนใหญ่แบ่งออกเป็น 2 ประเภท ได้แก่ การจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้หลายสัญญาณ และการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้เพียงสัญญาณเดียว ดังนั้นในงานวิจัยนี้จึงได้ประยุกต์ใช้วิธีการลตมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับทั้งสองประเภท ประกอบด้วย การจำแนกภาวะหยุดหายใจขณะนอนหลับด้วยสัญญาณ EEG และ ECG และการจำแนกภาวะหยุดหายใจขณะนอนหลับด้วยสัญญาณ ECG เพียงอย่างเดียว ขั้นตอนการทดลองนี้มีขั้นตอนเหมือนกับการทดลองที่ 5 ประกอบด้วย ขั้นตอนการเตรียมข้อมูล ขั้นตอนการประมวลผลสัญญาณ ขั้นตอนสกัดคุณลักษณะ ขั้นตอนการลตมิติข้อมูล และขั้นตอนการจำแนกข้อมูล

4.2.1 การทดลองและผลการทดลองประสิทธิภาพของวิธีการลตมิติแบบผสมผสาน DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับ

4.2.1.1 การทดลองที่ 6 การทดลองเปรียบเทียบประสิทธิภาพของการลตมิติแบบผสมผสานของ DisBGWO-DCCA ด้วยเงื่อนไขการแบ่งข้อมูลที่ต่างกันสำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG

การทดลองนี้มีจุดประสงค์เพื่อทดสอบประสิทธิภาพของวิธีการลตมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้หลายสัญญาณ ประกอบด้วย สัญญาณ EEG และสัญญาณ ECG ข้อมูล PSG ของการทดลองนี้ได้มาจากฐานข้อมูล Physionet ด้วยชุดข้อมูลของ Sleep Disorders Clinic at St Vincent's University Hospital (UCS sleep apnea database) [14] จำนวนทั้งหมด 28 ชุดข้อมูล ซึ่งเป็นข้อมูลการนอนผู้ป่วยที่หยุดหายใจขณะนอนหลับ สัญญาณที่นำมาใช้สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับ ได้แก่ สัญญาณ EEG จำนวน 2 ช่องสัญญาณ ณ ตำแหน่ง C2-A2 และ C4-A1 และสัญญาณ ECG ณ ตำแหน่ง lead V2 จำนวน 1 ช่องสัญญาณ ความถี่สุ่มเท่ากับ 128 เฮิรตซ์ ก่อนเข้าสู่กระบวนการสกัดคุณลักษณะสัญญาณ แต่ละสัญญาณถูกประมวลผลสัญญาณก่อนเพื่อกรองสัญญาณรบกวน จากนั้นแต่ละสัญญาณถูกสกัดคุณลักษณะตามรายละเอียดหัวข้อที่ 3.1.1 ในการทดลองนี้คุณลักษณะที่ได้จากสัญญาณ ECG มีความแตกต่างจากการทดลองที่ 1-5 คือ เราได้เพิ่มคุณลักษณะที่มาจากการสกัดสัญญาณการหายใจจากสัญญาณหัวใจหรือ EDR เข้ามาใช้ในการทดลอง เนื่องจากสัญญาณการหายใจมีผลต่อการจำแนกภาวะหยุดหายใจขณะนอนหลับ ขั้นตอนการสกัดสัญญาณ EDR เราได้ใช้เทคนิค Empirical Mode Decomposition (EMD) ตามหัวข้อ 2.3.3.2 ดังนั้นจำนวนคุณลักษณะที่ใช้การทดลองนี้จำนวนคุณลักษณะรวมทั้งสิ้น 249 คุณลักษณะ ประกอบด้วยคุณลักษณะของสัญญาณ EEG จำนวน 81×2 ช่องสัญญาณ = 162 คุณลักษณะ และคุณลักษณะของ

สัญญาณ ECG ซึ่งเป็นคุณลักษณะที่มาจากการวิเคราะห์ HRV จำนวน 56×1 ช่องสัญญาณ = 56 คุณลักษณะ และคุณลักษณะที่มาจากการสกัดสัญญาณ EDR จำนวน 31×1 ช่องสัญญาณ = 31 คุณลักษณะ เซตของคุณลักษณะที่ได้ในแต่ละสัญญาณถูกแบ่งออกเป็น 2 ส่วน เซตข้อมูล คือ เซตข้อมูลสำหรับการเรียนรู้ร้อยละ 80 และเซตของข้อมูลสำหรับการทดสอบร้อยละ 20 ขั้นตอนการเรียนรู้เริ่มต้นด้วยแต่ละเซตของคุณลักษณะจะถูกเลือกคุณลักษณะที่เหมาะสมด้วยวิธี DisBGWO จากนั้นแต่ละเซตเหล่านั้นจะถูกหลอมรวมด้วยวิธี DCCA ซึ่งหลอมรวมทีละคู่ จำนวนทั้งหมด 6 คู่ที่มีการหลอมรวมแบบอนุกรม และเลือกขนาดของไอเกนเท่ากับค่าความเชื่อมั่นร้อยละ 90 ขั้นตอนของการลดมิติแบบผสมผสาน DisBGWO-DCCA ตามอัลกอริทึมในหัวข้อที่ 3.1.2.3 สำหรับขั้นตอนการทดสอบเซตของคุณลักษณะที่เหมาะสม จะถูกทดสอบด้วยเซตข้อมูลสำหรับการทดสอบ

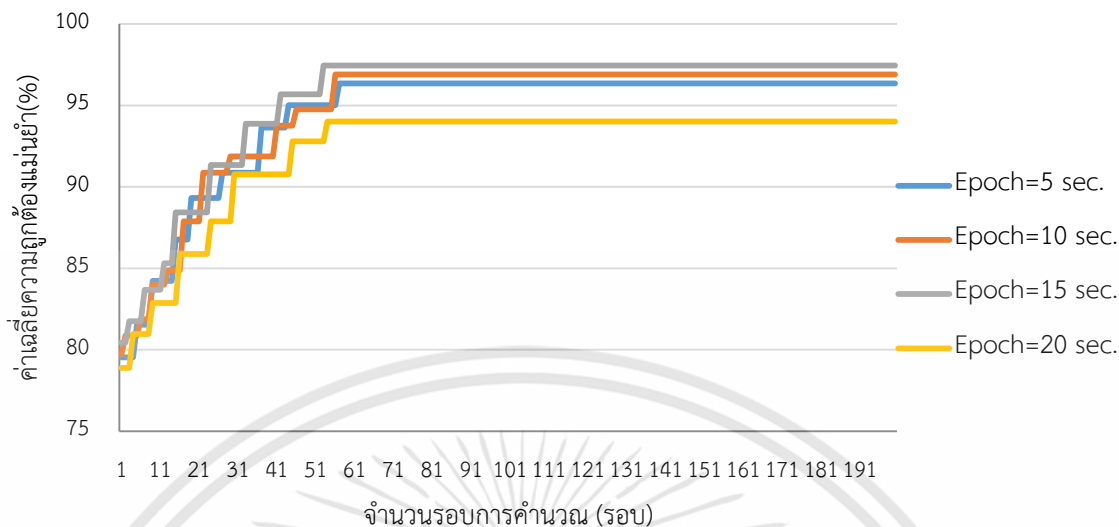
เนื่องจากการกำหนดการเกิดภาวะหยุดหายใจขณะนอนหลับหรือช่วง apnea และ hypopnea ด้วยวิธีแมนนวล คือ เมื่อมีการหยุดหายใจขณะนอนหลับเป็นช่วงเวลาที่ยาวนานอย่างน้อย 10 วินาที ซึ่งทำให้การแบ่งช่วงสัญญาณของการจำแนกภาวะหยุดหายใจขณะนอนจึงไม่เหมือนกับการจำแนกระยะการนอนที่กำหนดด้วยช่วงเวลาคงที่ ดังนั้นเราจึงได้ทำการทดลองแบ่งช่วงสัญญาณต่างๆ เพื่อหาช่วงเวลาการแบ่งสัญญาณที่เหมาะสมที่สุดสำหรับการจำแนกช่วง apnea, hypopnea และ normal โดยแบ่งออกเป็นช่วงๆ (Segment) ช่วงละ 5, 10, 15 และ 20 วินาที

เงื่อนไขการทดลอง เงื่อนไขในการทดลองของการทดลองนี้เหมือนกับเงื่อนไขการทดลองที่ 4 แต่แตกต่างที่การกำหนดจำนวนค่าคำตอบของแต่ละวิธีการค้นหาคำตอบหรือจำนวนคุณลักษณะทั้งหมดในการทดลองนี้เท่ากับ 249 จำนวน

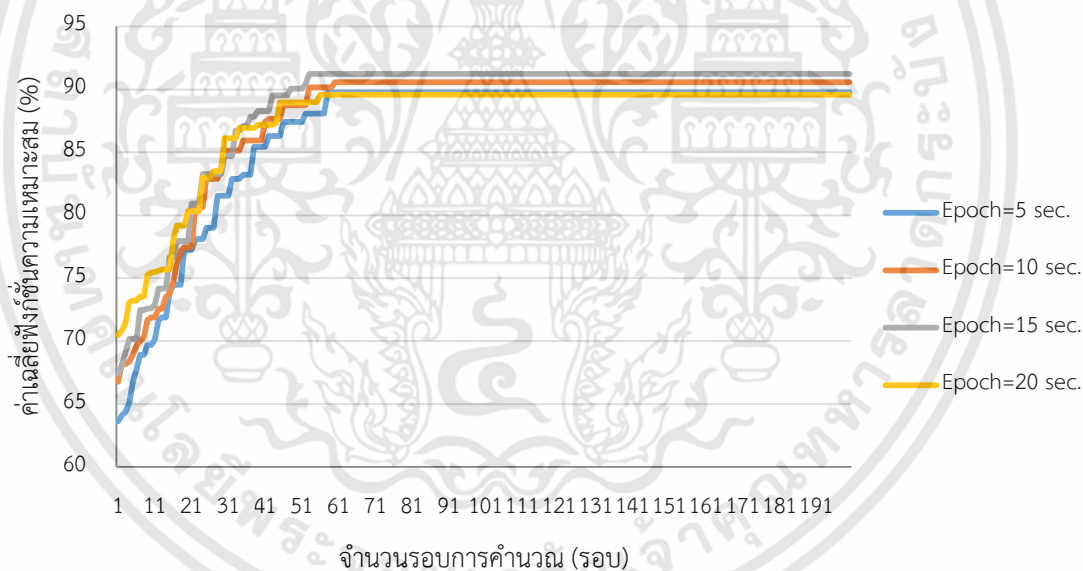
4.2.1.2 ผลการทดลองและวิเคราะห์ผลการทดลองของการทดลองที่ 6

ผลการทดลองขั้นตอนการเรียนรู้แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความถูกต้องแม่นยำของการจำแนกภาวะหยุดหายใจขณะนอนหลับและจำนวนรอบการคำนวณของแต่ละการแบ่งช่วงเวลา 5, 10, 15 และ 20 วินาที แสดงดังรูปที่ 4.28 และกราฟความสัมพันธ์ระหว่างค่าเฉลี่ยของฟังก์ชันค่าความเหมาะสมและจำนวนรอบการคำนวณแสดงดังรูปที่ 4.29 ผลการทดสอบประสิทธิภาพของวิธีลดมิติแบบผสมผสาน DisBGWO-DCCA ในขั้นตอนการเรียนรู้สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับแสดงดังตารางที่ 4.18

ผลการทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA ของการจำแนกระยะการนอนต่างๆ ประกอบด้วย Normal, Apnea และ Hypopnea ในขั้นตอนการทดสอบแสดงดังตารางที่ 4.20 ซึ่งแสดงค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะ เวลาในการประมวลผล และค่าฟังก์ชันความเหมาะสมตามสมการที่ 4.1



รูปที่ 4.28 ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนของการคำนวณของวิธี DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG



รูปที่ 4.29 ความสัมพันธ์ระหว่างค่าเฉลี่ยค่าฟังก์ชันความเหมาะสมและจำนวนรอบการคำนวณของวิธี DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG

ตารางที่ 4.19 ผลการทดลองของขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG

Algorithms	Fitness function value (%±SD)	The number of selected features (Number±SD)	Accuracy of selected features (%±SD)	The number of iterations (Number±SD)	Time (sec.±SD)
5 วินาที	89.0826±0.53	86.75±1.25	96.3532±0.66	54.25±0.95	1,879.62±135.76
10 วินาที	89.7155±0.64	85.50±1.29	96.8912±0.54	55.25±1.71	1,823.28±142.81
15 วินาที	90.1670±0.68	85.25±0.95	97.4556±0.73	55.25±1.50	1,751.14±161.32
20 วินาที	87.2604±0.75	88.50±1.91	94.0131±0.82	56.25±1.89	1,955.61±128.40
DisBGWO 15 วินาที	89.8644±0.66	91.50±1.22	97.5543±0.48	108.50±1.25	13,731.55±322.32

ตารางที่ 4.20 ผลการทดลองของขั้นตอนทดสอบการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ EEG และ ECG ด้วยวิธีมิติแบบผสมผสาน DisBGWO-DCCA

Segments	Acc. (%±SD)	Sent. (%±SD)	Spec. (%±SD)	Time (sec.±SD)	Fitness _{classifier} (%±SD)
5 sec.	96.0125±0.66	94.1712±0.78	97.5429±0.52	18.7153±2.87	83.9073±1.39
10 sec.	96.4525±0.93	94.5434±0.63	97.6325±0.69	18.4104±1.98	84.4695±1.06
15 sec.	97.0165±0.53	95.7743±0.71	97.8425±0.53	18.7611±2.06	84.6790±0.97
20 sec.	93.6540±0.83	92.6326±0.51	96.6212±0.75	18.9655±3.13	81.8481±1.52
DisBGWO 15 วินาที	96.7644±0.87	94.6431±0.77	97.4522±0.81	42.7213±5.23	83.5581±1.33

ผลการทดลองตารางที่ 4.19 ค่าประสิทธิภาพของวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับในขั้นตอนการเรียนรู้ จะเห็นได้ว่าการแบ่งสัญญาณในช่วงเวลา 15 วินาที แสดงค่าประสิทธิภาพสูงสุดด้วยค่าความเหมาะสมเท่ากับร้อยละ 90.1670±0.68 และผลการทดลองตารางที่ 4.20 แสดงค่าประสิทธิภาพของวิธีมิติแบบผสมผสาน DisBGWO-DCCA ในขั้นตอนการทดสอบจะเห็นได้ว่าการแบ่งช่วงสัญญาณ 15 วินาทีแสดงค่าประสิทธิภาพของการจำแนกข้อมูลสูงที่สุดเช่นกัน ด้วยค่าความถูกต้องแม่นยำเท่ากับ 97.0165±0.53 ค่าความไวเท่ากับ 95.7743±0.71 และค่าจำเพาะเท่ากับร้อยละ 97.8425±0.53 และแสดงค่าฟังก์ชันความเหมาะสมของการจำแนกข้อมูล

เท่ากับ 81.8481 ± 1.52 ซึ่งจากผลการทดลองนี้แสดงให้เห็นได้ว่าการแบ่งช่วงสัญญาณ 15 วินาทีเป็นการแบ่งช่วงสัญญาณที่เหมาะสมและมีประสิทธิภาพสำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับช่วง Normal, Apnea และ Hypopnea

4.2.1.3 การทดลองที่ 7 การทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว

การทดลองนี้มีจุดประสงค์เพื่อทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงสัญญาณเดียว ข้อมูล PSG ของการทดลองนี้ได้มาจากฐานข้อมูล Physionet ด้วยชุดข้อมูลของ ECG-Based Apnea detection [15] จำนวนทั้งหมด 70 ชุดข้อมูล ที่บันทึกสัญญาณมาจากผู้ป่วยที่มีความผิดปกติหยุดหายใจขณะนอนหลับโดยเฉลี่ยค่าประมาณ 7-10 ชั่วโมง สัญญาณ ECG ที่นำมาใช้สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับเป็นสัญญาณ ECG ณ ตำแหน่ง lead V2 ความถี่สุ่มเท่ากับ 100 เฮิร์ตซ์ ก่อนเข้าสู่กระบวนการสกัดคุณลักษณะสัญญาณ สัญญาณ ECG ผ่านกระบวนการกรองความถี่ จากนั้นถูกสกัดคุณลักษณะตามรายละเอียดหัวข้อที่ 3.1.1 ในการทดลองนี้คุณลักษณะที่ได้จากสัญญาณ ECG ประกอบด้วยคุณลักษณะที่ได้จากการวิเคราะห์ HRV จำนวน 56 คุณลักษณะและคุณลักษณะที่ได้จากสัญญาณ EDR จำนวน 31 คุณลักษณะ รวมทั้งหมด 87 คุณลักษณะ จากการลดมิติข้อมูลแบบผสมผสาน DisBGWO-DCCA มีเงื่อนไขคือ เป็นวิธีสำหรับการลดมิติข้อมูลที่มาจกหลายแหล่งข้อมูล ดังนี้เพื่อให้เป็นไปตามเงื่อนไขเราจึงได้ทดลองกำหนดแต่ละเซตของคุณลักษณะไว้ 2 รูปแบบ ดังนี้

รูปแบบที่ 1 เซตของคุณลักษณะมีจำนวน 2 เซตคุณลักษณะ ได้แก่ เซตของคุณลักษณะที่ได้จากการวิเคราะห์ HRV และเซตของคุณลักษณะที่ได้จากการวิเคราะห์ EDR

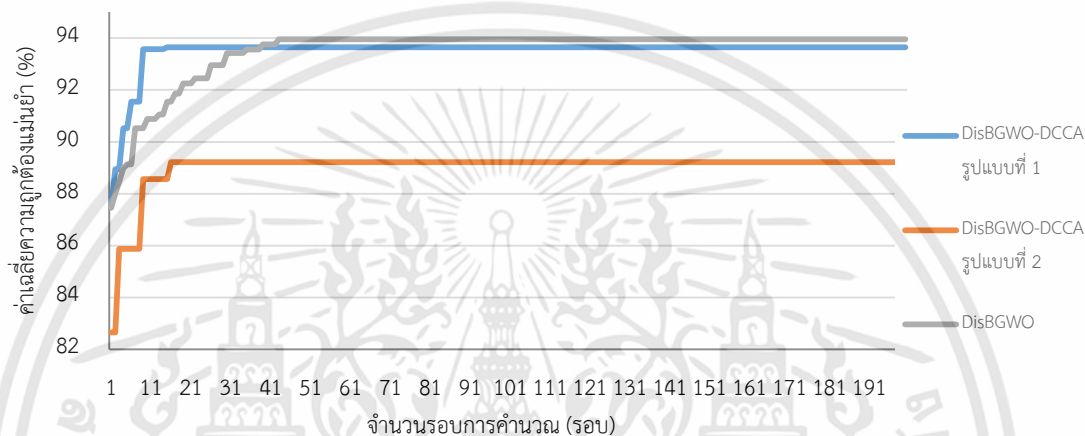
รูปแบบที่ 2 เซตของคุณลักษณะมีจำนวน 2 เซตคุณลักษณะ ได้แก่ เซตของคุณลักษณะที่ได้จากการวิเคราะห์ HRV และคุณลักษณะที่ได้จากการวิเคราะห์ EDR ด้วยเทคนิคเชิงเส้น และเซตของคุณลักษณะที่ได้จากการวิเคราะห์ HRV และคุณลักษณะที่ได้จากการวิเคราะห์ EDR ด้วยเทคนิคไม่เชิงเส้น

จากนั้นเซตของคุณลักษณะในแต่ละรูปแบบถูกดำเนินการทดลองด้วยวิธีการลดมิติ DisBGWO และ DisBGWO-DCCA ตามเงื่อนไขการทดลอง 3 และ 4 ตามลำดับ แต่แตกต่างกันที่การกำหนดจำนวนค่าคำตอบของแต่ละวิธีการค้นหาคำตอบหรือจำนวนคุณลักษณะทั้งหมดในการทดลองนี้เท่ากับ 87 จำนวน

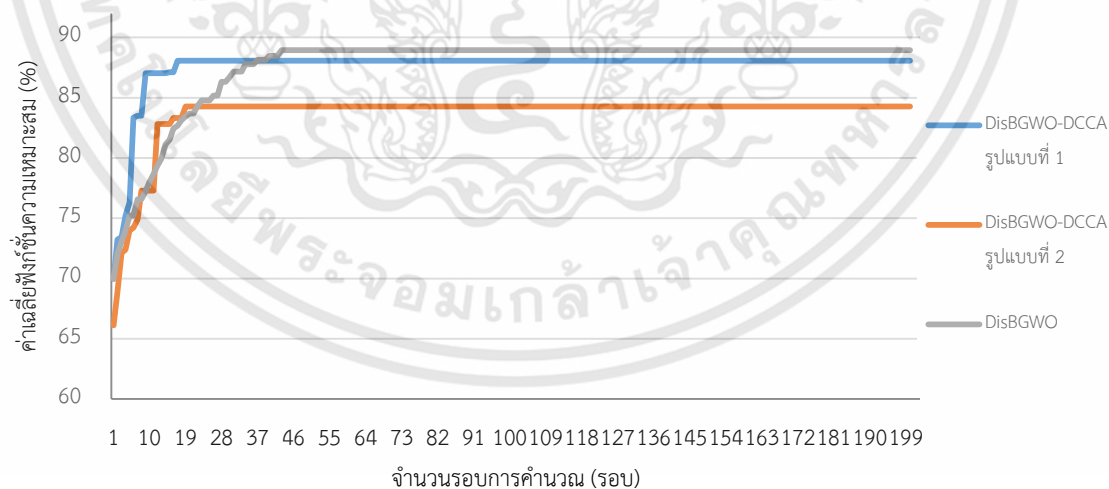
4.2.1.4 ผลการทดลองและวิเคราะห์ผลการทดลองของการทดลองที่ 7

ผลการทดลองขั้นตอนการเรียนรู้แสดงด้วยกราฟความสัมพันธ์ระหว่างค่าความถูกต้องแม่นยำของการจำแนกภาวะหยุดหายใจขณะนอนหลับและจำนวนรอบการคำนวณของเซตคุณลักษณะทั้ง 2 รูปแบบ แสดงดังรูปที่ 4.30 และกราฟความสัมพันธ์ระหว่างค่าเฉลี่ยของฟังก์ชันค่าความเหมาะสมและ

จำนวนรอบการคำนวณแสดงดังรูปที่ 4.31 ผลการทดสอบประสิทธิภาพของวิธีลตมิติด้วย DisBGWO และ DisBGWO-DCCA ในขั้นตอนการเรียนรู้สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียวแสดงดังตารางที่ 4.21 ผลการทดสอบประสิทธิภาพของวิธีลตมิติด้วย DisBGWO และ DisBGWO-DCCA เปรียบเทียบกับวิธีการลตมิติแบบที่ใช้ในงานวิจัยทั่วไป ได้แก่ PCA, LDA และ DCCA แสดงดังตารางที่ 4.22 ซึ่งแสดงค่าเฉลี่ยความถูกต้องแม่นยำ ค่าเฉลี่ยของค่าความไว ค่าเฉลี่ยของค่าจำเพาะ เวลาในการประมวลผล และค่าความเหมาะสมของการจำแนกข้อมูลตามสมการ 4.1



รูปที่ 4.30 ความสัมพันธ์ระหว่างค่าเฉลี่ยความถูกต้องแม่นยำและจำนวนรอบการคำนวณของวิธีลตมิติ DisBGWO และ DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว



รูปที่ 4.31 ความสัมพันธ์ระหว่างค่าเฉลี่ยฟังก์ชันความเหมาะสมและจำนวนรอบคำนวณของวิธีลตมิติ DisBGWO และ DisBGWO-DCCA สำหรับจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว

ตารางที่ 4.21 ผลการทดลองขั้นตอนเรียนรู้ของการทดลองวิธีมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียว

Algorithms	Fitness function value (%±SD)	The number of selected features (Number±SD)	Accuracy of selected features (%±SD)	The number of iterations (Number±SD)	Time (sec.±SD)
รูปแบบที่ 1 HRV Vs EDR	88.4813±0.64	28.25±0.95	93.6477±0.74	11.25±1.25	966.54±42.14
รูปแบบที่ 2 Linear Vs Non-linear	84.4734±0.82	30.00±0.81	89.2124±0.71	18.25±0.95	991.73±47.32
DisBGWO	88.9155±0.87	29.75±0.89	93.8136±0.74	39.45±1.55	6,371.42±105.42

ตารางที่ 4.22 ผลการทดลองของขั้นตอนทดสอบการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียวด้วยวิธีลดมิติ DisBGWO-DCCA, PCA, LDA, DCCA และ DisBGWO

Algorithms	ArgAcc. (%±SD)	ArgSent. (%±SD)	ArgSpec. (%±SD)	ArgTime (sec.±SD)	Fitness <i>classifier</i> (%±SD)
DisBGWO-DCCA รูปแบบที่ 1	92.4311±0.81	90.2231±0.91	94.2252±0.76	13.1121±2.31	76.2617±1.22
DisBGWO-DCCA รูปแบบที่ 2	88.9431±0.73	85.6574±0.82	93.6535±0.66	14.8301±1.88	71.1543±1.53
PCA	81.9013±0.52	81.0021±0.74	89.5521±0.83	12.0531±2.43	69.266±0.89
LDA	82.5411±0.86	81.2339±0.61	90.8821±0.59	12.9644±1.94	68.5489±1.03
DCCA	84.5526±0.64	83.6342±0.75	91.6625±0.94	13.6711±1.83	69.5051±1.33
DisBGWO	92.8132±0.73	90.8385±0.69	94.9542±0.81	10.7612±1.29	78.9865±0.89

ผลการทดลองดังตารางที่ 4.21 แสดงให้เห็นว่า ถึงแม้ว่าการวัดมิติด้วยการเลือกคุณลักษณะ DisBGWO แสดงค่าฟังก์ชันความเหมาะสมสูงสุดเท่ากับ 88.9155 ± 0.87 แต่เวลาที่ใช้ในการประมวลผลของขั้นตอนการเรียนรู้มากกว่าวิธีลดมิติแบบผสมผสาน DisBGWO-DCCA ทั้งนี้เนื่องจากการลู่อู่เข้าสู่คำตอบที่เหมาะสมของวิธีเลือกคุณลักษณะ DisBGWO ใช้จำนวนรอบการคำนวณที่มากกว่าวิธี DisBGWO-DCCA ดังนั้นวิธี DisBGWO จึงมีข้อเสียคือ ใช้เวลาการประมวลผลในขั้นตอนการเรียนรู้

จากผลการทดลองขั้นตอนการทดสอบดังตารางที่ 4.22 แสดงให้เห็นว่าวิธีการลดมิติ DisBGWO แสดงค่าความถูกต้องแม่นยำใกล้เคียงกับวิธี DisBGWO-DCCA ที่มีการหลอมรวมเซตของคุณลักษณะตามรูปแบบที่ 1 แต่เวลาการประมวลผลในผลขั้นตอนการทดสอบของวิธีการเลือกคุณลักษณะ DisBGWO ใช้เวลาน้อยที่สุด จึงเป็นสาเหตุที่ทำให้ค่าความเหมาะสมของวิธี DisBGWO แสดงค่าสูงสุดเท่ากับ 78.9865 ± 0.89 นอกจากนี้เมื่อเปรียบเทียบผลการทดสอบประสิทธิภาพการจำแนกภาวะหยุดหายใจขณะนอนหลับด้วยสัญญาณ ECG เพียงอย่างเดียวระหว่างวิธีการลดมิติที่งานวิจัยนี้นำเสนอทั้งสองวิธี (DisBGWO และ DisBGWO-DCCA) และวิธีการลดมิติที่ใช้งานวิจัยอื่น (PCA, LDA และ DCCC) [36-37, 56] จะเห็นได้ว่า ค่าฟังก์ชันความเหมาะสมของวิธีลดมิติที่เรานำเสนอแสดงค่าสูงกว่าวิธีการลดมิติ PCA, LDA และ DCCA ดังนั้นผลการทดลองนี้สามารถสรุปได้ว่าวิธีการลดมิติด้วยการเลือกคุณลักษณะ DisBGWO เป็นวิธีที่เหมาะสมและมีประสิทธิภาพสำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับของช่วง Normal, Apnea และ Hypopnea โดยใช้สัญญาณ ECG เพียงอย่างเดียว

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการทดลอง

วิทยานิพนธ์ฉบับนี้มุ่งเน้นการศึกษาเรื่อง การวิเคราะห์การนอนอัตโนมัติ โดยเฉพาะการจำแนก ระยะการนอนหลับ (Sleep stage) และการจำแนกภาวะหยุดหายใจขณะนอนหลับ (Sleep apnea disorder) โดยใช้หลายสัญญาณ เนื่องจากการศึกษาในงานวิจัยที่เกี่ยวข้องได้พบว่า แม้ว่าการใช้สัญญาณ เพียงชุดเดียวสำหรับการวิเคราะห์การนอนจะมีความสะดวกสบายในการใช้งานมากกว่า แต่ประสิทธิภาพ ด้านความถูกต้องแม่นยำของวินิจฉัยไม่สูงมากนัก ดังนั้นการใช้หลายชุดสัญญาณสำหรับการวิเคราะห์การ นอนจึงมีประสิทธิภาพที่ดีกว่าและเป็นที่ยอมรับทางการแพทย์มากกว่าการใช้สัญญาณเพียงสัญญาณเดียว แต่อย่างไรก็ตามการวิเคราะห์การนอนด้วยหลายสัญญาณนั้นทำให้เกิดปัญหามิติข้อมูลสูงซึ่งส่งผลทำให้ ประสิทธิภาพของระบบการเรียนรู้อัตโนมัติลดลง ดังนั้นงานวิจัยนี้จึงได้มุ่งเน้นการเพิ่มประสิทธิภาพของ การวิเคราะห์การนอนโดยได้ออกแบบแต่ละขั้นตอนอย่างเหมาะสม โดยเฉพาะอย่างยิ่งขั้นตอนการลดมิติ ซึ่งเป็นขั้นตอนที่สำคัญสำหรับการแก้ไขปัญหาลักษณะนี้ ขั้นตอนการออกแบบการวิเคราะห์การนอนแบบ อัตโนมัติประกอบด้วย ขั้นตอนการประมวลผลสัญญาณ ขั้นตอนการสกัดคุณลักษณะ ขั้นตอนการลดมิติ ข้อมูล และขั้นตอนการจำแนกข้อมูล งานวิจัยนี้ได้นำข้อมูล PSG เพื่อทดลองจำแนกระยะการนอนได้มา จากฐานข้อมูล SHHS จำนวน 100 ชุดข้อมูล และฐานข้อมูล Dream database จำนวน 47 ชุดข้อมูล สำหรับการทดลองการจำแนกภาวะหยุดหายใจขณะนอนได้นำข้อมูล PSG มาจากฐานข้อมูล Physionet ที่เป็นชุดข้อมูลของ UCS sleep apnea data จำนวน 28 ชุดข้อมูล และชุดข้อมูลของ ECG-Based Apnea detection จำนวน 70 ชุดข้อมูล

จากการศึกษางานวิจัยที่เกี่ยวข้องกับการวิเคราะห์การนอนพบว่า สัญญาณ EEG EOG และ ECG ซึ่งเป็นสัญญาณที่ถูกบันทึกขณะนอนหลับเป็นสัญญาณที่บ่งบอกการทำงานของร่างกายขณะนอนหลับ โดยเฉพาะความเกี่ยวเนื่องระหว่างการทำงานของสมองและหัวใจขณะนอนหลับ ตัวอย่างเช่น ขณะที่ ร่างกายอยู่ในระยะการนอนช่วง REM ขณะนั้นสมองจะมีการทำงานซึ่งสังเกตได้จากสัญญาณ EEG จะมีความถี่สูงและขนาดสัญญาณเล็ก ขณะเดียวกันมีการเคลื่อนไหวของลูกตาเป็นจังหวะรวดเร็ว และมีการ ทำงานของระบบซิมพาเทติกที่สูงขึ้น และขณะที่ร่างกายอยู่ในช่วงการนอน NREM หรือหลับลึก ซึ่ง ร่างกายจะผ่อนคลายมากที่สุด สัญญาณ EEG อยู่ในช่วงความถี่ต่ำและมีขนาดแอมพลิจูดใหญ่ การ เคลื่อนไหวของลูกตาจะช้ามาก และจังหวะการเต้นของหัวใจจะคงที่หรือการทำงานของระบบพาราซิมพา เทติกจะสูงขึ้น จากตัวอย่างดังกล่าวจะเห็นได้ว่าสัญญาณ EEG EOG และ ECG เป็นสัญญาณที่มี

ความสำคัญอย่างยิ่งสำหรับการนำมาใช้ในการวิเคราะห์การนอนและเป็นสัญญาณที่แสดงให้เห็นถึงการเปลี่ยนแปลงของร่างกายขณะนอนหลับได้อย่างชัดเจน การสกัดคุณลักษณะของสัญญาณ EEG และ EOG มีขั้นตอนการสกัดคุณลักษณะที่คล้ายคลึงกัน โดยใช้เทคนิคทางเวลา ทางความถี่ ทางเวลา-ความถี่ และเทคนิคไม่เชิงเส้น นอกจากนี้งานวิจัยนี้ได้นำเสนอคุณลักษณะของสัญญาณ EEG ใหม่ด้วยเทคนิคทางเวลา หรือเรียกว่า $avgMMD_{slope}$ และ $sumMMD_{slope}$ [15] ซึ่งเป็นคุณลักษณะหนึ่ง que แสดงให้เห็นถึงความแตกต่างระหว่างระยะการนอนต่างๆ ได้อย่างมีนัยสำคัญทางสถิติ สำหรับการสกัดคุณลักษณะของสัญญาณ ECG ประกอบด้วย คุณลักษณะที่ได้จากการวิเคราะห์ HRV และการวิเคราะห์สัญญาณการหายใจที่สกัดจากสัญญาณ ECG ด้วยเทคนิคเชิงเส้นและเทคนิคไม่เชิงเส้น การนำคุณลักษณะของสัญญาณ ECG ไปใช้ในการจำแนกระยะการนอนและการจำแนกภาวะหยุดหายใจขณะนอนหลับมีขั้นตอนที่ต่างกันเพียงเล็กน้อย คือ การจำแนกระยะการนอนจะใช้คุณลักษณะที่มาจากการวิเคราะห์ HRV เพียงอย่างเดียว แต่ในการจำแนกภาวะหยุดหายใจขณะนอนหลับจะใช้ทั้งคุณลักษณะที่ได้จากการวิเคราะห์ HRV และ EDR เนื่องจากสัญญาณการหายใจมีความสำคัญต่อการประเมินภาวะหยุดหายใจขณะนอนหลับ

จากขั้นตอนการสกัดคุณลักษณะจะเห็นได้ว่าจำนวนคุณลักษณะที่มีความสำคัญต่อการวิเคราะห์การนอนมีจำนวนมาก ซึ่งทำให้เกิดปัญหามิติข้อมูล ดังนั้นการวิจัยนี้จึงได้เน้นการนำเทคนิคการลดมิติมาใช้แก้ปัญหาดังกล่าว ซึ่งการลดมิติเป็นเทคนิคที่มีความสามารถในการเลือกคุณลักษณะและลดคุณลักษณะได้อย่างเหมาะสม งานวิจัยนี้ได้นำเสนออัลกอริทึมการลดมิติแบบผสมผสานใหม่ ประกอบด้วย อัลกอริทึมการลดมิติแบบผสมผสาน CCA/DCCA-NIAs และ NIAs-CCA/DCCA ซึ่งเป็นการลดมิติแบบผสมผสานรูปแบบที่ 3 และ 4 ตามที่กล่าวไว้ในหัวข้อที่ 2 อัลกอริทึมทั้งสองแตกต่างกันเพียงการสลับขั้นตอนก่อนหลัง โดยอัลกอริทึม CCA/DCCA-NIAs จะเริ่มต้นด้วยการลดมิติด้วยวิธีการแปลงคุณลักษณะ CCA/DCCA จากนั้นตามด้วยวิธีเลือกคุณลักษณะ NIAs สำหรับอัลกอริทึม NIAs-CCA/DCCA จะลำดับขั้นตอนก่อนหลังสลับกัน การทดสอบประสิทธิภาพของอัลกอริทึมที่นำเสนอถูกเปรียบเทียบกับเทคนิคการลดมิติที่ใช้ในงานวิจัยทั่วไป ได้แก่ การลดมิติด้วยการแปลงคุณลักษณะ และการลดมิติด้วยการเลือกคุณลักษณะ จากผลการทดลองที่ 1 เป็นการเปรียบเทียบประสิทธิภาพวิธีการลดมิติด้วยการแปลงคุณลักษณะ PCA, LDA, CCA และ DCCA จะเห็นได้ว่าการหลอมรวมด้วยเทคนิคการแปลงคุณลักษณะ DCCA ที่มีการหลอมรวมแบบอนุกรมบนพื้นฐานความเชื่อมั่นร้อยละ 90 แสดงค่าความถูกต้องแม่นยำของการจำแนกระยะการนอนได้สูงสุดเท่ากับร้อยละ 81.8763±0.55 นอกจากนี้ผลการทดลองที่ 1 ได้แสดงให้เห็นว่าการจำแนกข้อมูลด้วยอัลกอริทึม KNN ที่ k=5 มีความเหมาะสมสำหรับนำไปใช้ในการกำหนดเงื่อนไขของการวิเคราะห์การนอน เนื่องจากเป็นอัลกอริทึมแสดงค่าฟังก์ชันความเหมาะสมของการจำแนกสูงสุดเท่ากับร้อยละ 75.01 ซึ่งหมายถึงอัลกอริทึมดังกล่าวแสดงค่าความถูกต้องแม่นยำสูงและใช้เวลาในการประมวลผลน้อย สำหรับผลการทดลองที่ 2 เป็นผลการทดลองการเปรียบเทียบประสิทธิภาพระหว่าง

วิธีการเลือกคุณลักษณะทั่วไป ได้แก่ BPSO, BCS และ BGWO กับวิธีการเลือกคุณลักษณะที่งานวิจัยนี้ นำเสนอ (DisBGWO) เป็นวิธีการที่พัฒนามาจาก BGWO ซึ่งได้เพิ่มขั้นตอนการสร้างวิธีการค้นหาคำตอบใหม่จากวิธีการกลายพันธุ์ด้วยการใช้ค่าสัมประสิทธิ์ความต่างของ Jaccard (Dissimilarity Jaccard coefficient) หรือค่าสัมประสิทธิ์ความต่างระหว่างสองเวกเตอร์เข้าไปในกระบวนการค้นหา BGWO แบบเดิม จากผลการทดลองขั้นตอนการเรียนรู้เราได้เปรียบเทียบวิธีการกำหนดค่าคำตอบเริ่มต้นที่ แตกต่างกันด้วยวิธี Small initialization, Random initialization และ Large initialization จากรูปที่ 4.8-4.9 จะเห็นได้ว่าการกำหนดค่าเริ่มต้นด้วยวิธี Large initialization เป็นที่เหมาะสมที่สุดสำหรับการ กำหนดค่าเริ่มต้นของวิธีการค้นหาคำตอบของวิธีการเลือกคุณลักษณะ นอกจากนี้ผลการทดลองตารางที่ 4.9 แสดงให้เห็นว่า การเลือกคุณลักษณะ DisBGWO แสดงค่าความถูกต้องแม่นยำสูงสุดเท่ากับร้อยละ 95.0534±0.56 ถึงแม้ว่าค่าความถูกต้องแม่นยำจะมีค่าสูงแต่จำนวนของคุณลักษณะที่ใช้ยังสูงอีกด้วย จึง ทำให้วิธีการเลือกคุณลักษณะมีข้อเสียที่เด่นชัด คือ เวลาที่ใช้ในการประมวลผลในขั้นตอนการเรียนรู้ใช้ เวลานาน จากปัญหาดังกล่าวเราจึงได้นำเสนอและออกแบบวิธีการลดมิติแบบผสมผสานใหม่คือ อัลกอริทึม CCA/DCCA-NIAs โดยการทดลองที่ 3 เป็นการทดลองเปรียบเทียบประสิทธิภาพของวิธีที่ได้ นำเสนอ ผลการทดลองขั้นตอนของการเรียนรู้แสดงดังตารางที่ 4.11 แสดงให้เห็นว่าการลดมิติแบบ ผสมผสาน DCCA-DisBGWO เป็นอัลกอริทึมที่แสดงค่าความถูกต้องแม่นยำสูงสุดเท่ากับร้อยละ 90.8965±0.85 แต่เมื่อเทียบกับผลการทดลองที่ 2 แล้วค่าความถูกต้องแม่นยำของวิธีการเลือก DisBGWO มีค่าสูงกว่า อย่างไรก็ตามเมื่อเปรียบเทียบเวลาที่ใช้ในการประมวลผลจะเห็นได้ว่า วิธีการลด มิติแบบผสมผสาน DCCA-DisBGWO ใช้เวลาน้อยกว่าวิธีการเลือกคุณลักษณะ DisBGWO ถึง 2.91 เท่า และใช้จำนวนคุณลักษณะน้อยกว่าประมาณ 1.33 เท่า ถ้าแม้ว่าวิธี DCCA-DisBGWO จะสามารถช่วย แก้ไขปัญหาเวลาในการประมวลผลในขั้นตอนการเรียนรู้ได้แต่ในขั้นตอนการทดสอบยังใช้ในการประมวลผล นาน ทั้งนี้เนื่องจากวิธีการแปลงคุณลักษณะในขั้นตอนแรกส่งผลให้คุณสมบัติของคุณลักษณะเหล่านั้น เปลี่ยนไป ดังนั้นจึงต้องเลือกคุณลักษณะใหม่ทุกครั้ง จึงเป็นเหตุผลที่ทำให้ค่าฟังก์ชันความเหมาะสมของ การจำแนกข้อมูลแสดงค่าลดลงเท่ากับ 75.4346±0.54 จากผลการทดลองนี้แสดงให้เห็นว่า วิธีการลดมิติ แบบผสมผสาน DCCA-DisBGWO มีจุดด้อยในด้านของเวลาในการประมวลผลของขั้นตอนการทดสอบ และยังคงต้องปรับปรุงในส่วนของคุณค่าความถูกต้องแม่นยำอีกด้วย ซึ่งทำให้เราได้นำเสนอวิธีการลดมิติแบบ ผสมผสานใหม่อีกแบบเรียกว่า NIAs-DCCA/CCA ซึ่งเริ่มต้นด้วยวิธีการเลือกคุณลักษณะ NIAs และหลอม รวมแต่ละชุดคุณลักษณะด้วยวิธีแปลงคุณลักษณะ CCA/DCCA การทดสอบประสิทธิภาพของอัลกอริทึม NIAs-DCCA/CCA ถูกนำเสนอในการทดลองที่ 4 และผลการทดลองในขั้นตอนการเรียนรู้แสดงดังตารางที่ 4.13 จะเห็นได้ว่าการลดมิติแบบผสมผสาน DisBGWO-DCCA เป็นวิธีการลดมิติที่แสดงค่าความถูกต้อง แม่นยำสูงสุดเท่ากับร้อยละ 93.7222±0.48 และใช้จำนวนคุณลักษณะน้อยเท่ากับ 138.25±0.95 ซึ่งทำ

ให้ค่าฟังก์ชันความเหมาะสมของการลดมิติแบบผสมผสานเท่ากับร้อยละ 87.2875±0.64 ซึ่งเมื่อเปรียบเทียบกับวิธี DCCA, DisBGWO และ DCCA-DisBGWO ของการทดลองที่ 1-3 จะเห็นได้ว่าค่าฟังก์ชันความเหมาะสมของ วิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA แสดงค่าฟังก์ชันความเหมาะสมสูงสุด ดังนั้นจึงสามารถสรุปได้ว่า วิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA เป็นวิธีที่เหมาะสมสำหรับการจำแนกกระยะการนอน นอกจากนี้งานวิจัยนี้ได้ทำการทดลองทดสอบประสิทธิภาพของอัลกอริทึม DisBGWO-DCCA โดยนำไปใช้กับฐานข้อมูลอื่นซึ่งเป็นฐานข้อมูล Dream database ซึ่งประกอบด้วยข้อมูลของกลุ่มคนปกติและกลุ่มของผู้ป่วย ซึ่งผลการทดลองแสดงดังตารางที่ 4.17-4.18 แสดงให้เห็นว่าค่าประสิทธิภาพของอัลกอริทึม DisBGWO-DCCA ทั้งขั้นตอนการเรียนรู้และขั้นตอนการทดสอบมีค่าใกล้เคียงกับผลการทดลองที่ 4 และแสดงให้เห็นว่าสามารถนำไปใช้ในการวิเคราะห์ระยะการนอนทั้งข้อมูลการนอนของคนปกติและผู้ป่วยได้อย่างมีประสิทธิภาพ ดังนั้นเราสามารถสรุปได้ว่า วิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สามารถนำไปประยุกต์ได้กับการจำแนกกระยะการนอนกับฐานข้อมูลอื่นๆได้ สำหรับการทดสอบประสิทธิภาพของวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA สำหรับการจำแนกภาวะหยุดหายใจขณะนอนหลับโดยใช้หลายชุดสัญญาณถูกนำเสนอในการทดลองที่ 6 และผลการทดลองแสดงให้เห็นว่าวิธีการลดมิติ DisBGWO-DCCA สามารถนำไปประยุกต์ใช้ในการจำแนกภาวะหยุดหายใจขณะนอนหลับได้อย่างมีประสิทธิภาพ นอกจากนี้การทดสอบประสิทธิภาพอัลกอริทึมที่นำเสนอเพื่อจำแนกภาวะหยุดหายใจขณะหลับโดยใช้สัญญาณ ECG เพียงอย่างเดียวถูกนำเสนอในการทดลองที่ 7 ซึ่งผลการทดลองแสดงให้เห็นว่าการวิธีการเลือกคุณลักษณะ DisBGWO เป็นวิธีที่เหมาะสมที่สุดเมื่อเปรียบเทียบกับวิธีการที่ใช้ในงานวิจัยอื่นๆ ดังนั้นเราสามารถสรุปได้ว่าวิธีการลดมิติที่งานวิจัยนี้ได้นำเสนอสามารถนำไปใช้ในการจำแนกภาวะหยุดหายใจขณะนอนหลับได้อย่างมีประสิทธิภาพ

5.2 ข้อเสนอแนะ

เนื่องจากวิธีการลดมิติแบบผสมผสาน DisBGWO-DCCA เป็นวิธีที่ถูกนำเสนอเพื่อลดมิติข้อมูลสำหรับหลายชุดข้อมูล ซึ่งงานวิจัยนี้ได้ประยุกต์ใช้กับการวิเคราะห์การนอนเท่านั้น โดยผลการทดลองสามารถพิสูจน์ได้อย่างชัดเจนว่าเป็นวิธีที่สามารถปรับปรุงค่าความถูกต้องแม่นยำและเวลาในการประมวลผลได้อย่างมีประสิทธิภาพ หากต้องการนำวิธีที่นำเสนอไปใช้ในการวิเคราะห์ชุดข้อมูลอื่นๆ ข้อจำกัดของการนำไปใช้คือ 1) ชุดข้อมูลอินพุตจะต้องเป็นข้อมูลที่มากกว่าหนึ่งชุดข้อมูลเท่านั้น หากเป็นชุดข้อมูลอินพุตมีเพียงชุดข้อมูลเดียวอาจจะต้องออกแบบการแบ่งเซตของคุณลักษณะให้มีความแตกต่างกันเพื่อสามารถหาฟังก์ชันการแปลงของ DCCA ได้ดังตัวอย่างการทดลองที่ 6 หรือการเลือกใช้วิธีลดมิติด้วยการเลือกคุณลักษณะ DisBGWO อาจจะมีประสิทธิภาพที่ต่ำกว่าทั้งนี้ขึ้นอยู่กับรูปแบบของข้อมูลนั้นๆ 2) ค่าคงที่ของตัวแปรของต่างๆ อาจจะมีการเปลี่ยนแปลงขึ้นอยู่กับลักษณะรูปแบบของข้อมูลนั้นๆ 3) และการเลือกอัลกอริทึมของการจำแนกข้อมูลขึ้นอยู่กับผลของประสิทธิภาพการคำนวณ

เอกสารอ้างอิง

- [1] M. O. Mendez, M. Matteucci, and V. Castronovo, "Sleep staging from Heart Rate Variability: time-varying spectral features and Hidden Markov Models", **International Journal of Biomedical Engineering and Technology**, Vol. 3, No.3, pp. 246-263, 2010.
- [2] M. Ronzhina, M. Ronzhina, O. Janousek, J. Kolárová, M. Nováková, P. Honzík, and I. Provazník, "Sleep scoring using artificial neural networks" **Journal of Sleep medicine reviews**, Vol.16, pp. 251-263, 2012.
- [3] L. Zoubek, S. Charbonnier, S. Lesecq, A. Buguet, and Florian Chapotot "Feature selection for sleep/wake stages classification using data driven methods", **Journal of Biomedical Signal Processing and Control**, Vol.2, pp. 171-179, 2007.
- [4] L. Cervante, B. Xue, M. Zhang, and L. Shang "Binary Particle Swarm Optimization for Feature Selection: A Filter Based Approach", **WCCI 2012 IEEE World Congress on Computational Intelligence**, Brisbane, Australia, pp-10-15, 2012.
- [5] D. Rodrigues, L. A. M. Pereira, A. N. Souza, C. C. O. Ramos and Xin-She Yang "BCS: A Binary Cuckoo Search Algorithm for Feature Selection", **IEEE International Symposium on Circuits and Systems (ISCAS2013)**, Beijing, China, pp. 465-468, 2013.
- [6] S. Salesi and G. Cosma "A Novel Extended Binary Cuckoo Search Algorithm for Feature Selection", **2nd International Conference on Knowledge Engineering and Applications**, London, UK, 2017.
- [7] E. Emary, H.M. Zawbaa, and A.E. Hassanien, "Binary grey wolf optimization approaches for feature selection", **Journal of Neurocomputing**, Vol. 172, pp. 371-381, 2016.
- [8] Q. Li, H. Chen, H. Huang, X. Zhao, Z.N. Cai, C. Tong, W. Liu, and X. Tian "An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis", **Journal of Computational and Mathematical Methods in Medicine**, Vol. 2017 pp-1-15, 2017.
- [9] H. Hotelling, "*Relations between Two Sets of Variates*," **Journal of Biometrika**, Vol. 28, No. 3-4, pp. 321-377, 1936.

- [10] T. Sun, S. Chen, J. Yang, P. Shi “A Novel Method of Combined Feature Extraction for Recognition”, **8th IEEE International Conference on Data Mining**, pp. 1043-1048, 2008.
- [11] Y. Yuan, C. Ma and D. Pu, “A Novel Discriminant Minimum Class Locality Preserving Canonical Correlation Analysis And Its Applications” **Journal of Industrial and management optimization**, Vol. 12, No.1, pp.251-268, 2016.
- [12] National Sleep Research Resource, “Sleep Heart Health Study”. [Online].
Available: <https://sleepdata.org/datasets/shhs/pages/full-description.md>.
- [13] Dream Project, “The dreams subjects Database” [Online].
Available: <http://www.tcts.fpms.ac.be/~devuyst/Databases/DatabaseSubjects/>
- [14] Physionet, “St. Vincent's University Hospital Sleep Disorders Clinic” [Online].
Available: <https://physionet.org/cgi-bin/atm/ATM>.
- [15] Physionet, “Apnea-ECG Database” [Online].
Available: <https://physionet.org/physiobank/database/apnea-ecg/>
- [16] P. Moeynoi and Y. Kitjaidure, “Dimension reduction based on Canonical Correlation Analysis technique to classify sleep stages of sleep apnea disorder using EEG and ECG signals”, **Proc. of International Conf. On Electrical Engineering/ Electronic, Computer, Telecommunication and Information Technology**, Phuket, Thailand, pp.455-458, 2017.
- [17] P. Moeynoi and Y. Kitjaidure, “Hybrid Dimensionality Reduction of Multi-sets Using Nature Inspired Algorithms and Discriminant Canonical Correlation Analysis for Automatic Sleep Stage Classification”, **INASS: International Journal of Intelligence Engineering and system**, Vol.12, No.1, pp.277-289,2019.
- [18] Sharon Keenan, and Max Hirshkowitz. “Monitoring and Staging Human Sleep”, **Principles and practice of sleep medicine**. 5th edition ed. St. Louis: Elsevier Saunders. pp. 16-26.
- [19] R.K. Malhotra and A.Y. Avidan “Sleep stage and Scoring Techniques”, **Atlas of Sleep Medicine**, 1st edition, Butterworth-Heinemann, pp. 77-99, 2005.

- [20] M. H. Silber, S. Ancoli, M. H. Bonnet, S. Chokroverty, M. M. Grigg-Damberger and et al. "The Visual Scoring Sleep in Adults", **Journal of Clinical Sleep Medicine**, Vol. 3, No.2, pp.121-136, 2017.
- [21] American Academy of Sleep Medicine, A. The international classification of sleep disorders. In: Sateia M, editor. **Diagnostic and coding manual**, 2nd edition. Westchester (IL): American Academy of Sleep Medicine, p.1–297, 2005.
- [22] Carin Sahlin, **Sleep and Sleep apnea**, Department of Public Health and Clinical Medical, Respiratory Medicine an Allergy, Umea University, Sweden 2009.
- [23] E. Edson, and H. Nazeran, "EEG and HRV signal features for automatic sleep staging and apnea detection", **20th International Conference on Electronics Communications and Computers (CONIELECOMP)**, pp142-147, 2010
- [24] X. Kong, V. Goel, and N.Thakor. "Quantification of injury related EEG signal changes using Itakura distance measure", **International Conference on Acoustics, Speech, and Signal Processing**, pp.2947-2950, 1995
- [25] K.I. Aboalayon, M. Faesipour, W.S. Almuhammadi and S. Moslehpour, "Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation", **Journal of Entropy**, Vol 18, No. 272, 2016.
- [26] Hassan, A.R.; Bashar, S.K.; "Bhuiyan, M.I.H. On the classification of sleep states by means of statistical and spectral features from single channel electroencephalogram." **IEEE on Advances in Computing, Communications and Informatics (ICACCI)**, Kochi, India, pp. 2238–2243, 2015.
- [27] B. Sen, M. Peker, A. Cavusoglu and FV. Celebi, "A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms" **Journal of Medical System.**, Vol 38, No.18, pp. 1-21, 2014.
- [28] S. Aungsakul, A. Phinyomark, P. Phukpattaranont, C. Limsakul "Evaluating Feature Extraction Methods of Electrooculography (EOG) Signal for Human-Computer Interface", **Journal of Procedia Engineering**, Vol.32, pp.246-252, 2012.
- [29] J. Pan and W.J.Tomkins, "A real-time QRS detection algorithm" **IEEE Transection Biomedical Engineering**, Vol 32, pp 230-236, 1985.

- [30] Meng Xiao, Hong Yan, Jinzhong Song, Yuzhou Yang and Xianglin Yang, “Sleep Stage classification based on heart rate variability and random forest”, **Biomedical Signal Processing and Control**, Volume 8, pp. 624-633, 2013.
- [31] Ayyoob J, “Sleep apnea detection from ECG using features extracted from reconstructed phase space and frequency domain”, **Biomedical Signal Processing and control**, Vol8., pp.551-558, 2013.
- [32] P. Moeynoi and Y. Kitjaidure, “Canonical Correlation Analysis for Dimensionality Reduction of Sleep Apnea Features based on ECG single lead” **The 9th Biomedical Engineering International Conference**, December 7-9, 2016.
- [33] Steven V., “Heart Rate Variability: linear and non-linear analysis with applications in humans physiology”, Doctoral Thesis of Faculty of Electrical Engineering Kasteelpark Arenberg, Belgium, 2010.
- [34] Z. B. Khalad and G. Farges, “First approach for respiratory monitor by amplitude demodulation of electrocardiogram” **14th the Annual International Conference on IEEE Engineering in Medical and Biomedical Society**, 1992. Vol6 , pp. 2535-2536.
- [35] C. L. Mason and L. Tarassenko, “Quantitative assessment of respiratory derivation algorithms”, **23rd Annual IEEEEMBS International Conference**, 2001, pp.1998–2001.
- [36] B. Raymond, R.M.Cayton, R.A.Bates, M.Chappell, “Screening for obstructive sleep apnea based on the electrocardiogram” **The computers in cardiology challenge, Comput.Cardiol**, vol. 27, pp.267–270, 2000 .
- [37] J. Boyle, N. Bidargaddi, A. Sarela, M. Karunanithi, “Automatic detection of respiration rate from ambulatory single-lead ECG”, **13rd IEEE Transection of Information Technology of Biomedical**, pp.890–896, 2009 .
- [38] K. Kamble, V. Lohikpure, S. Sirsat, and N. Jadhav, “Estimation of a Respiratory Signal from Single Channel Electrocardiogram Signal using Empirical Mode Decomposition”, **International Journal of Scientific Engineering and Technology Research**, Vol. 3, No.17, pp. 3505-3508, 2014.

- [39] M. Campolo, D. Labate, F. La Foresta, F. C. Morabito, A. Lay-Ekuakille, and P. Vergallo, “ECG derived respiratory signal using empirical mode decomposition”, **IEEE International Symposium on Medical Measurements and Applications**, pp.399–403, May 2011.
- [40] D. Widjaja, C. Varon, A.C. Dorado, Johan A.K. Suykens, S.V. Huffel, “Application of kernel principal component analysis for single-lead-ECG derived respiration”, **IEEE Transactions on Biomedical Engineering**, vol.59, No.4, pp1169–1176, 2012.
- [41] D. Labate, F. L. Forestam, G. Occhiuto, F. C. Morabito, A. Lay-Ekuakille and P. Vergallo, “Empirical Mode Decomposition vs Wavelet Decomposition for the Extraction of Respiratory Signal from Single-Channel ECG: A Comparison”, **IEEE sensors journal**, Vol 13, no.7, pp. 2666-2674.
- [42] S.F. Liang, C. E. Kuo, Y.H.Hu, Y.H. Pan, and Y. H. Wang, “Automatic Stage Scoring of Single-Channel Sleep EEG by Using Multiscale Entropy and Autoregressive Models”, **IEEE Transactions On Instrumentation And Measurement**, Vol. 61, No.6, pp.1649-1657, 2012.
- [43] D. Popovic, M. Khoo, and P. Westbrook, “Automatic scoring of sleep stages and cortical arousals using two electrodes on the forehead: validation in healthy adults” **Journal of Sleep Respiratory**, Vol.23, No.2, pp.211-221, 2014.
- [44] A. Bhattacharjee, S. Saha, S. A. Fattah, W.P. Zhu and M. O. Ahmad, “Sleep Apnea Detection Based on Rician Modeling of Feature Variation in Multi-band EEG Signal”, **IEEE Journal of Biomedical and Health Informatics**, Vol 10, 2018.
- [45] V. Vimala, K. Ramar and M. Ettapan “An Intelligent Sleep apnea Classification System Based on EEG Signals”, **Journal of Medical systems**, Vol.43, 2019.
- [46] E. Magosso, M. Ursino, A. Zaniboni, F. Provini, and P. Montagna, “Visual and computer-based detection of slow eye movements in overnight and 24-h EOG recordings”, **Journal of Clinical Neurophysiology**, Vol.118, No.5, pp. 1122-1133, 2007.
- [47] A. Coskun, S. Ozsen, S. Yucelbas2, C. Yucelbas, G. Tezel, S. Kuccukturk, and S. Yosunkaya “Detection of REM in Sleep EOG Signals” **Journal of Science and Technology**, Vol.9, No.25, 2016.

- [48] Md. M. Rahman, M. I. Hassan Bhuiyan and A. R. Hassan “Sleep stage classification using single-channel EOG ” **Journal of Computer in Biology and Medicine**, Vol. 102, pp. 211-220, 2018.
- [49] F. Togo and Y. Yamamoto., “Decreased fractal component of human heart rate variability during non-REM sleep. American Journal of Physiology”, **Heart and Circulatory Physiology**, Vol. 280, No.1, pp.17–21, 2001.
- [50] G.J. Gates, S.E. Mateika, and J.H. Mateika, “Heart rate variability in nonapneic snorers and controls before and after continuous positive airway pressure”, **Bio Medical Central Pulmonary Medicine**, Vol.5, No.9, 2005.
- [51] S. Ozsen, “Classification of sleep stages using class-dependent sequential feature selection and artificial neural network” **Journal of Neural Computation and Application**, Vol.23, pp. 1239-1250, 2013.
- [52] S. Khalighi, T. Sousa, G. Pires, and U. Nunes “Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnographic channels”, **Expert Systems with Applications**, Vol.40, pp.7046-7059, 2013.
- [53] A. A. Gharbali, S. Najdi, J. M. Fonseca, “Investigating the contribution of distance-based features to automatic sleep stage classification”, **Journal of Computers in Biology and Medicine**, Vol.96, pp. 8-23, 2018.
- [54] P. Anderer, G. Gruber, S. Parapatics, M. Woertz, T. Miazhyńska, G. Klosch, G., et al. “An E-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24 x 7 utilizing the Siesta database”, **Journal of Neuropsychobiology**, Vol.51, No.3, pp. 115-133, 2005.
- [55] R. Yan, C. Zhang, K. Spruyt, L. Wei, Z. Wang, L. Tian, X. Li, T. Ristaniemi, J. Zhang, F. Cong, “Multi-modality of polysomnography signals’ fusion for automatic sleep scoring”, **Journal of Biomedical Signal Processing and Control**, Vol.49, pp.14-23, 2018.
- [56] S.M. Isa, M. I. Fanany, W. Jatmiko, and A. M. Arymurthy “Sleep Apnea Detection from ECG Signal Analysis on Optimal Features, Principal Components, and Nonlinearity” **5th International Conference on Bioinformatics and Biomedical Engineering**, Wuhan, China, 2011.

- [57] S. Rawat and V.K. Verma, “EEG Signal based Sleep Detection using PCA and Neural Network” **International Journal of Digital Application & Contemporary research**, Vol.3, No.3, 2014.
- [58] Rabia A., C.K. Verma, and N. Srivastava “Dimension reduction methods for microarray data: a review”, **Journal of AIMS Bioengineering**, Vol.4, No.1, pp.179-197, 2017.
- [59] J. Liu, H. Meng, M. Li, F. Zhang, R. Qin ,and A. K. Nandi, “Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction”, **Journal of Concurrency and Computation: Practice and Experience**, Vol.30, No.23, 2018.
- [60] Nicolle M. Correa¹, Tom Eichele², Tülay Adalı ¹, Yi-Ou Li¹, and Vince D. Calhoun “Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI”, **Journal of Neuroimage**, Vol.50, No.4, pp.1438-1445, 2010.
- [61] S.M. Uma, and E. Kirubakaran, “A hybrid heuristic dimensionality reduction technique for microarray gene expression data classification: a blending of GA, PSO and ACO” **International Journal of Data Mining, Modelling and Management**, Vol.8. No.2, 2016
- [62] K. K. Bharti , and P. K. Singh, “Hybrid dimension reduction by integrating feature selection with feature extraction method for text clustering”, **Journal of Expert Systems with Application**, Vol. 42, pp. 3105-3114, 2015.
- [63] S. Jacophine Susmi, H. Khanna Nehemiah, and A. Kannan, “Hybrid Dimension Reduction Techniques with Genetic Algorithm and Neural Network for Classifying Leukemia Gene Expression Data”, **Indian Journal of Science and Technology**, Vol.9, No.12, 2016.
- [64] S.Y. Xin and S. Deb “Cuckoo Search via Levy Flights”, **Proc. of World Congress on Nature & Biologically Inspired Computing**, pp. 210-214, 2009.
- [65] S. Mirjalili, S. Mohamad Mirjalili and A. Lewis “Grey Wolf Optimizer” **Journal of advances in engineering software**, Vol.69, pp. 46-61,2014.
- [66] Ahsan H. Khandoker, Marimuthu Palaniswami and Chandan K. Karmaker, “Support Machine for Automated Recognition of Obstructive Sleep Apnea Syndrome from

ECG Recordings”, *IEEE transection on Information Technology in biomedical*,
Vol.13, No.1, 2009.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก

รายละเอียดของคุณลักษณะที่ได้จากสัญญาณ EEG, EOG และ ECG

1. คุณลักษณะที่ได้จากสัญญาณ EEG จำนวนทั้งหมด 81 คุณลักษณะ

คุณลักษณะที่	คุณลักษณะ	จำนวน
Time Domain		
F1	Mean of EEG signal	1
F2	Median of EEG signal	1
F3	Standard deviation of EEG signal	1
F4-F5	Maximum and minimum value of EEG signal	2
F6	Variation of EEG signal	1
F7- F9	The 25 th , 50 th , and 75 th percentile of EEG signal	3
F10-F12	Hjorth Parameters as Eq. 2.1-1.3 and [23]	3
F13	Itakura distance as Eq. 2.8 and [24]	1
F14	Skewness	1
F15	Kurtosis	1
F16-F19	Maximum–Minimum distance (MMD) as Eq. 2.11 and [25] (sub-window =1 sec, 5 sec, 10 sec ,and 15 sec)	4
F20-F27	Average and Sum of the slope of maximum-minimum distance (MMD_{slope}) each sub-window as equation 2.14 and [16] (sub-window =1 sec, 5 sec, 10 sec ,and 15 sec)	8
Frequency Domain		
F28-F37	The unit of power spectrum density calculated by FFT and welch methods [26] (nDelta,nTheta,nAlpha,nBeta, nAll)	10
F38-F45	The relative percent of spectral energy band (RPEB) calculated by FFT and welch methods as Equation 2.15 (RPEBDelta, RPEBTheta, RPEBAlpha, RPEBBeta)	8
Time-Frequency Domain		
F46-50	Mean of coefficients derived from discrete wavelet transformation as [27] (argCoA5, argCoD5, argCoD4, argCoD3, argCoD2)	5
F51-55	Standard derivation of coefficients derived from discrete wavelet transformation (DWT) as [27] (varCoA5, varCoD5, varCoD4, varCoD3, varCoD2)	5
F56-F60	Power spectrum density using FFT method of coefficients DWT as [26-27] (PSD _{fft} of CoA5, PSD _{fft} of CoD5, PSD _{fft} of CoD4, PSD _{fft} of CoD3, PSD _{fft} of CoD2)	5
F61-F65	Power spectrum density using Welch method of coefficients DWT as [26-27] (PSD _{welch} CoA5, PSD _{welch} CoD5, PSD _{welch} CoD4, PSD _{welch} CoD3, PSD _{welch} CoD2)	5
F66-F70	Ratio of power spectrum density using FFT method of coefficients (rPSD _{fft} of CoA5, rPSD _{fft} of CoD5, rPSD _{fft} of CoD4, rPSD _{fft} of CoD3, rPSD _{fft} of CoD2)	5
F71-F75	Ratio of power spectrum density using Welch method of coefficients (PSD _{welch} CoA5, PSD _{welch} CoD5, PSD _{welch} CoD4, PSD _{welch} CoD3, PSD _{welch} CoD2)	5
Non-linear		
F76-F79	Auto regression coefficients for EEG epoch	4
F80-F81	Mean zero crossing index and standard derivation of zero crossing index	2

2. คุณลักษณะที่ได้จากสัญญาณ EOG จำนวนทั้งหมด 73 คุณลักษณะ

คุณลักษณะที่	คุณลักษณะ	จำนวน
Time Domain		
F1	Mean of EOG signal	1
F2	Median of EOG signal	1
F3	Standard deviation of EOG signal	1
F4-F5	Maximum and minimum value of EOG signal	2
F6	Variance of EOG signal as Eq. 2.16	1
F7	Skewness	1
F8	Kurtosis	1
F9	Area under curve of EOG signal as Eq. 2.17 and [28]	1
F10	Threshold crossing value of EOG signal as Eq. 2.18 and [28]	1
F11-F14	Maximum–Minimum distance (MMD) as Eq. 2.11 and [25] (sub-window =1 sec., 5 sec., 10 sec., and 15 sec.)	4
F15-F22	Average and Sum of MMD_{slope} each intervals of sub-window as Eq. 2.14 and [16] (sub-window =1 sec., 5 sec., 10 sec., and 15 sec.)	8
Frequency Domain		
F23-F30	The unit of power spectrum density calculated by FFT and welch methods (0.3-2, 2-5, 5-7 and 8-10 Hz.)	8
F31-F38	The relative percent of spectral energy band (RPEB) calculated by FFT and welch methods as Eq. 2.15 (0.3-2, 2-5, 5-7 and 8-10 Hz.)	8
Time-Frequency Domain		
F39-F43	Mean of coefficients derived from DWT (argCoA5, argCoD5, argCoD4, argCoD3, argCoD2)	5
F44-F48	Standard derivation of coefficients derived from DWT (varCoA5, varCoD5, varCoD4, varCoD3, varCoD2)	5
F49-F53	Power spectrum density using FFT method of coefficients DWT (PSD_{fft} of CoA5, PSD_{fft} of CoD5, PSD_{fft} of CoD4, PSD_{fft} of CoD3, PSD_{fft} of CoD2)	5
F54-F59	Power spectrum density using Welch method of coefficients DWT (PSD_{welch} CoA5, PSD_{welch} CoD5, PSD_{welch} CoD4, PSD_{welch} CoD3, PSD_{welch} CoD2)	5
F60-F64	Ratio power spectrum density using FFT method of coefficients DWT ($rPSD_{fft}$ of CoA5, $rPSD_{fft}$ of CoD5, $rPSD_{fft}$ of CoD4, $rPSD_{fft}$ of CoD3, $rPSD_{fft}$ of CoD2)	5
F65-F69	Ratio power spectrum density using welch method of coefficients DWT (PSD_{welch} CoA5, PSD_{welch} CoD5, PSD_{welch} CoD4, PSD_{welch} CoD3, PSD_{welch} CoD2)	5
Non-linear		
F70-F71	Auto regression coefficients for EOG epoch	2
F72-F73	Mean zero crossing index and Standard derivation of zero crossing index	2

3. คุณลักษณะที่ได้จากสัญญาณ ECG ของการวิเคราะห์ HRV จำนวนทั้งหมด 56 คุณลักษณะ

คุณลักษณะที่	คุณลักษณะ (จำนวน)	จำนวน
Time Domain		
F1	Mean of RRI	1
F2	Median of RRI	1
F3	Standard deviation value of RRI	1
F4	Variance of RRI	1
F5-F6	Maximum and Minimum of RRI	2
F7	Skewness of RRI	1
F8	Kurtosis of RRI	1
F9	Root mean square value of successive differences of RRI	1
F10	The number of RRI of more or less than 50ms.	1
F11	The percentage of RRI of more/less than 50ms.	1
F12	Mean of heart rate	1
F13	Standard deviation value of heart rate	1
Frequency Domain		
F14-F22	The unit of power spectrum density calculated by FFT and welch methods as Eq. 2.18-2.19 (VLF, LF, HF, the ratio of LF and HF)	8
F23-F30	The mean frequency using FFT and Welch methods (VLF, LF, HF, the ratio of LF and HF)	8
Time-Frequency Domain		
F31-F39	Mean of coefficients DWT (argCoA8, argCoD8-D2)	9
F40-F48	Standard derivation of coefficients DWT (stdCoA8, stdCoD8-D2)	9
Non-linear		
F49	Correlation dimension analysis as Eq. 2.22 and [31]	1
F50	Auto regression coefficients of RR interval	1
F51-F52	Detrended Fluctuation Analysis as Eq. 2.24 and [31] (δ_1, δ_2)	2
F53-54	Mutual Information as Eq. 2.25 and [30] (Beat Decay and Peak Decay)	2
F55-F56	Mean zero crossing index and standard derivation of zero crossing index	2

4. คุณลักษณะที่ได้จากสัญญาณ ECG ของการวิเคราะห์ EDR จำนวนทั้งหมด 31 คุณลักษณะ

คุณลักษณะที่	คุณลักษณะ	จำนวน
Time Domain		
F1	Mean of EDR	1
F2	Median of EDR	1
F3	Standard deviation value of EDR	1
F4	Variance of EDR	1
F5-F6	Maximum and Minimum of EDR	2
F7	Skewness of EDR	1

F8	Kurtosis of EDR	1
F9	Root mean square value of successive differences of EDR	1
F10	The number of EDR of more or less than 50ms.	1
F11	The percentage of EDR of more/less than 50ms.	1
Frequency Domain		
F12-F13	PSD in unit from FFT and Welch methods	2
F14-F15	The mean frequency FFT and Welch methods	2
Time-Frequency Domain		
F16-F20	Mean of coefficients (argCoA5, argCoD5-D2)	5
F21-F25	SD of coefficients (stdCoA5, stdCoD5-D2)	5
Non-linear		
F26	Correlation dimension analysis as Eq. 2.22	1
F27	Detrended Fluctuation Analysis as Eq. 2.24 (δ_1)	2
F28-F29	Mutual Information as Eq. 2.25 and [30] (Beat Decay and Peak Decay)	2
F30-F31	Mean zero crossing index and Standard derivation of zero crossing index	2

ประวัติผู้เขียน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อ-นามสกุล	นางสาวพิมพ์พร เหมยน้อย
วัน เดือน ปีเกิด	28 มกราคม 2531 ที่เชียงใหม่
ที่อยู่	71/1 หมู่ที่ 6 ตำบล ร้องวัวแดง อำเภอ สันกำแพง เชียงใหม่ 50130 โทร.087-197-1468
ประวัติการศึกษา	2552 วิศวกรรมศาสตรบัณฑิต สาขาวิชาโทรคมนาคม สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง 2555 วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมชีวการแพทย์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ความชำนาญเฉพาะด้าน	1.) Machine learning and Artificial Intelligence system 2.) Signal and Image processing of Medical data 3.) Data analysis
ประสบการณ์การทำงานและผลงาน	
พ.ศ.2554-2556	ตำแหน่งผู้ช่วยนักวิจัย ศูนย์ AIMC คณะแพทยศาสตร์ โรงพยาบาลรามาธิบดี - Signal and image processing in the advanced MRI (fMRI, DTI, T2*, and MRs) techniques for the Thai brain mapping project. - True innovation award 2012, Idea seed silver award, 2012, “iDoctor ข้อมูลโรงพยาบาลส่วนตัว”
พ.ศ.2557-2559	ตำแหน่ง Signal processing engineer (R&D) บริษัท Zensorium Ltd. - Conduct algorithms for mental states detection, sleep stage classification, blood pressure detection, HRV, and motion artifact. - Patient ID FP7868: PCT/SG2015/050500, 2015 “Method and apparatus for deriving a mental state of a subject”.
พ.ศ.2560-2561	ตำแหน่งนักวิศวกร และ Intelligence Unit Manager บริษัท Atapy Ltd. - ออกแบบอัลกอริทึม activity tracking, stress event detection, sleep stage classification and health care predictive - 3 rd Runner up Award in Medical Product Design of NIA 2016 Thailand: “Sookjai and Somjai wearable for elderly healthcare”.