การจำแนกประเภทคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัส โดยใช้ซิงกูลาร์แวลูดีคอมโพซิชันและอัลกอริทึมเอสเอ็มโอ

SURFACE ELECTROMYOGRAPHY SIGNAL CLASSIFICATION USING SINGULAR VALUE DECOMPOSITION AND SMO ALGORITHM

ยศภัทร เรืองไพศาล

YOTSAPAT RUANGPAISARN

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2558

KMITL-2015-SC-M-002-066

การจำแนกประเภทคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัส โดยใช้ซิงกูลาร์แวลูดีคอมโพซิชันและอัลกอริทึมเอสเอ็มโอ

SURFACE ELECTROMYOGRAPHY SIGNAL CLASSIFICATION USING SINGULAR VALUE DECOMPOSITION AND SMO ALGORITHM

ยศภัทร เรืองไพศาล

YOTSAPAT RUANGPAISARN

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
ปริญญาวิทยาศาสตรมหาบัณฑิต
ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
พ.ศ. 2558

KMITL-2015-SC-M-002-066

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SURFACE ELECTROMYOGRAPHY SIGNAL CLASSIFICATION USING SINGULAR VALUE DECOMPOSITION AND SMO ALGORITHM

YOTSAPAT RUANGPAISARN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN COMPUTER SCIENCE

FACULTY OF SCIENCE

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2015

KMITL-2015-SC-M-002-066

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

COPYRIGHT 2015

FACULTY OF SCIENCE

KING MOGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คณะวิทยาศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ใบรับรองวิทยานิพนธ์

หัวข้อวิทยานิพนธ์ "การจำแนกประเภทคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัสโดยใช้ซิงกูลาร์แวลูดีคอมโพซิชันและอัลกอริทึมเอสเอ็มโอ"
"SURFACE ELECTROMYOGRAPHY SIGNAL CLASSIFICATION USING SINGULAR VALUE DECOMPOSITION AND SMO ALGORITHM"

ชื่อนักศึกษา นายยศภัทร เรืองไพศาล
รหัสประจำตัว 56605027
ปริญญา วิทยาศาสตรมหาบัณฑิต (สาขาวิทยาการคอมพิวเตอร์)
ภาควิชา วิทยาการคอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์ ดร.สายชล ใจเย็น
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม (ถ้ามี) -----

คณะกรรมการสอบวิทยานิพนธ์	ลายมือชื่อ
ผศ.ดร.ศรัณย์ อินทโกสุม ประธานกรรมการ	
ผศ.ดร.อนันตพร หรรษคุณาฒัย อาจารย์บัณฑิตประจำ (ในสาขาวิชาที่เกี่ยวข้อง)	
ผศ.ดร.ศุภกานต์ พิมลธเรศ ผู้ทรงคุณวุฒิจากภายนอกสถาบันฯ	
ดร.สายชล ใจเย็น อาจารย์ที่ปรึกษาวิทยานิพนธ์	

วัน/ เดือน/ ปี ที่สอบ 1 ธันวาคม พ.ศ. 2558 เวลา 17.00 - 19.00 น.
สถานที่สอบ ณ ห้อง 306 ตึกปฏิบัติการหลังใหม่

คณะวิทยาศาสตร์รับรองแล้ว

(รองศาสตราจารย์ ดร.ดุษณี ธนะบริพัฒน์)
คณบดีคณะวิทยาศาสตร์
วันที่ ๒๑ **เดือน** ธันวาคม **พ.ศ.** ๒๕๕๘

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การจำแนกประเภทคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัสโดยใช้ซิงกูลาร์แวลูดีคอมโพซิชันและอัลกอริทึมเอสเอ็มโอ
ชื่อนักศึกษา	ยศภัทร เรืองไพศาล
รหัสประจำตัว	56605027
ปริญญา	วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์
ภาควิชา	วิทยาการคอมพิวเตอร์
พ.ศ.	2558
อาจารย์ที่ปรึกษาวิทยานิพนธ์	อาจารย์ ดร.สายชล ใจเย็น

บทคัดย่อ

การวิเคราะห์คลื่นสัญญาณไฟฟ้ากล้ามเนื้อ นับได้ว่าเป็นงานที่ท้าทายอย่างหนึ่งในสาขาวิชาประสาทวิทยา โดยคลื่นสัญญาณไฟฟ้าเหล่านี้มีความเกี่ยวข้องกับกิจกรรมการทำงานของระบบกล้ามเนื้อในร่างกายมนุษย์ ซึ่งสามารถนำไปประยุกต์ใช้ในการควบคุมสั่งงานแขนจักรกลสำหรับผู้พิการได้ ดังนั้นวิทยานิพนธ์ฉบับนี้ จึงได้นำเสนอวิธีการวิเคราะห์รูปแบบใหม่โดยใช้พื้นฐานการทำงานของวิธีซิงกูลาร์แวลูดีคอมโพซิชันและอัลกอริทึมเอสเอ็มโอ เพื่อจำแนกคลื่นสัญญาณไฟฟ้ากล้ามเนื้อออกเป็นการเคลื่อนไหวพื้นฐานของมือจำนวน 6 ประเภท ในการทดลองนี้คลื่นสัญญาณไฟฟ้ากล้ามเนื้อจากผู้รับการทดลองจำนวน 5 คน จะถูกสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับวิธีซิงกูลาร์แวลูดีคอมโพซิชัน และแบ่งข้อมูลที่ได้ออกเป็นชุดฝึกสอนและชุดทดสอบเป็นการตรวจสอบไขว้แบบ 2 ส่วน จากนั้นจึงส่งข้อมูลไปยังเทคนิคการจำแนกประเภทจำนวน 5 เทคนิค ได้แก่ ต้นไม้ตัดสินใจ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว ขั้นตอนวิธีแบบเบย์อย่างง่าย โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และอัลกอริทึมเอสเอ็มโอ ผลการทดลอง พบว่าการใช้การสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับวิธีซิงกูลาร์แวลูดีคอมโพซิชันช่วยในการลดจำนวนคุณลักษณะของข้อมูลลงไปได้มากและเมื่อใช้เทคนิคจำแนกประเภทด้วยอัลกอริทึมเอสเอ็มโอแล้วให้ประสิทธิภาพการจำแนกประเภทการเคลื่อนไหวพื้นฐานของมือได้ดีที่สุดเมื่อเปรียบเทียบกับวิธีการจำแนกอื่น ๆ ดังที่กล่าวมาข้างต้น

คำสำคัญ : การสกัดคุณลักษณะ การจำแนกประเภท คลื่นไฟฟ้ากล้ามเนื้อแบบผิวสัมผัส ซิงกูลาร์แวลูดีคอมโพซิชัน อัลกอริทึมเอสเอ็มโอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Surface Electromyography Signal Classification Using Singular Value Decomposition and SMO Algorithm
Student Name	Yotsapat Ruangpaisarn
Student ID	56605027
Degree	Master of Science Computer
Department	Computer Science
Year	2015
Thesis Advisor	Dr.Saichon Jaiyen

Abstract

Surface Electromyography (sEMG) signal analysis is a challenging task in neuroscience. The signal is associated with an activity of muscles in human body, thus it becomes an essential way to control the robotic arm for people with disabilities. Therefore, this thesis has proposed a new method based on Singular Value Decomposition (SVD) and SMO algorithm for classifying sEMG signals into six basic hand movements. For this experiment, The features of sEMG signals from five subjects are extracted by vector to matrix and-SVD feature extraction, divided into training and testing sets by two fold cross-validation, and subsequently fed to five classifiers including Decision Tree, K-nearest Neighbor, Naïve Bayes, Radial Basis Function Network, and Sequential Minimal Optimization. The experimental results have shown that vector to matrix and SVD feature extraction with SMO algorithm can achieve the best performance for the classification of basic hand movements when compared to the other classification of the above mentioned.

Keywords : Feature Extraction, Classification, Surface Electromyography, Singular Value Decomposition, SMO algorithm.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

วิทยานิพนธ์นี้มิอาจจะสำเร็จลุล่วงไปได้ด้วยดี หากปราศจากความอนุเคราะห์ คำแนะนำและความเอาใจใส่จากอาจารย์ที่ปรึกษาซึ่งคือ ดร.สายชล ใจเย็น โดยท่านได้สละเวลาให้คำปรึกษาเสนอแนะและแก้ไขการทำวิทยานิพนธ์ให้กับข้าพเจ้าอย่างเต็มที่ และขอขอบพระคุณคณะกรรมการสอบหัวข้อและโครงร่างวิทยานิพนธ์ ผศ.ดร.ศรัณย์ อินทโกสุม ผศ.ดร.อนันตพร หรรษคุณาฒัย และผศ.ดร.ศุภกานต์ พิมลธเรศ ที่กรุณาให้คำแนะนำตลอดจนชี้แนะแนวทางแก้ไขจนทำให้วิทยานิพนธ์เล่มนี้สามารถสำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณนายณัฐวุฒิ ชัยรัตนทรงพร นายนิพัทธ์ คล้ายโพธิ์ นางสาวภูมิธารา เรืองทอง นายอนุสรณ์ เจริญนาน นายทักษ์ดนัย สุวรรณ เพื่อนมหาบัณฑิต ๆ ร่วมห้องค้นคว้าและวิจัย ภาควิชาวิทยาการคอมพิวเตอร์ สถาบันสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณลาดกระบัง ซึ่งในระหว่างการศึกษาของข้าพเจ้าได้รับการช่วยเหลือพร้อมคำปรึกษาและได้รับอำนวยความสะดวกในด้านต่าง ๆ

ขอขอบคุณมหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก วิทยาเขตบางพระที่ได้ให้โอกาสและสนับสนุนทุนการศึกษาให้แก่ข้าพเจ้า

สุดท้ายนี้ข้าพเจ้าขอขอบพระคุณบิดา มารดา และพี่ ๆ น้อง ๆ ที่เป็นกำลังใจให้แก่ข้าพเจ้า และขอขอบคุณบุคคลอื่น ๆ ที่มิได้กล่าวถึงมา ณ ที่นี้ด้วย

นายยศภัทร เรืองไพศาล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ก

บทคัดย่อภาษาอังกฤษ ..... ข

กิตติกรรมประกาศ ..... ค

สารบัญ ..... ง

สารบัญตาราง ..... ฌ

สารบัญรูป ..... ฏ

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของงานวิจัย 2

1.3 สมมติฐานของงานวิจัย 2

1.4 ขอบเขตงานวิจัย 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 4

2.1 ความรู้พื้นฐานเกี่ยวกับคลื่นสัญญาณไฟฟ้าในร่างกาย 4

2.2 เครื่องมือและอุปกรณ์สำหรับตรวจวัดคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ 4

2.3 ชนิดของกล้ามเนื้อที่ใช้ทดลองในงานวิจัย 5

2.4 การลดจำนวนมิติของข้อมูล (Dimensionality Reduction) 6

2.4.1 ขั้นตอนวิธีแบบเลือกคุณลักษณะของข้อมูล (Feature Selection) 6

2.4.2 ขั้นตอนวิธีแบบสกัดคุณลักษณะของข้อมูล (Feature Extraction) 7

2.5 ความรู้พื้นฐานเกี่ยวกับเวกเตอร์และเมทริกซ์ 7

2.5.1 เวกเตอร์ (Vector) 7

2.5.2 เมทริกซ์ (Matrix) 7

2.6 การวิเคราะห์องค์ประกอบ (Factor Analysis) 9

2.6.1 ซิงกูลาร์แวลูดีคอมโพซิชัน (Singular Value Decomposition : SVD) 9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.2 การวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ (Linear Discriminat Analysis : LDA)....10
2.7 การเรียนรู้ของเครื่อง (Machine Learning)....11
2.7.1 ขั้นตอนวิธีแบบซีโฟพอยต์ไฟว์ (C4.5)....12
2.7.2 ขั้นตอนวิธีแบบเบย์อย่างง่าย (Naïve bayes)....14
2.7.3 ขั้นตอนวิธีแบบเพื่อนบ้านใกล้สุด *k* ตัว (*k*-Nearest neighbour : *k*-NN) ...15
2.7.4 โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial basis Function Neural Network : RBF)....17
2.7.5 อัลกอริทึมเอสเอ็มโอ (Sequential Minimal Optimizatio : SMO)....18
2.8 วิธีการวัดประสิทธิภาพของตัวจำแนกประเภท (Evaluation of Classifiers)....20
2.8.1 การตรวจสอบไขว้ (Cross-Validation)....20
2.8.2 Confusion Matrix....21
2.8.3 ค่าความถูกต้อง (Accuracy)....22
2.8.4 ค่าความแม่นยำ (Precision)....23
2.8.5 ค่าความระลึก (Recall)....23
2.8.6 ค่าเอฟเมเซอร์ (F-measure)....23
2.9 งานวิจัยที่เกี่ยวข้อง....24
2.9.1 sEMG based classification of basic hand movements based on time-frequency features....24
2.9.2 Towards sEMG Classification Based on Bayesian and *k*-NN to Control a Prosthetic Hand....25
บทที่ 3 การสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน....27
3.1 ขั้นตอนการเตรียมข้อมูล....27
3.2 วิธีการแปลงข้อมูลแบบเวกเตอร์เป็นข้อมูลแบบเมทริกซ์....29
3.3 การสกัดคุณลักษณะของข้อมูลด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน....30
3.4 ขั้นตอนการวัดประสิทธิภาพ....33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4 การทดลองและผลการทดลอง 34

4.1 การทดลอง 34

4.1.1 โปรแกรมที่ใช้ในงานวิจัยและการกำหนดค่าพารามิเตอร์ 34

4.2 ผลการทดลอง.....35

4.2.1 การทดลองกับข้อมูลผู้ทดลองที่ 1 35

4.2.1.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 35

4.2.1.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 35

4.2.1.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 38

4.2.2 ผลการทดลองกับข้อมูลผู้ทดลองที่ 2.....40

4.2.2.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 40

4.2.2.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 40

4.2.2.1 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 43

4.2.3 ผลการทดลองกับข้อมูลผู้ทดลองที่ 3.....45

4.2.3.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 45

4.2.3.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 45

4.2.3.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.4 ผลการทดลองกับข้อมูลผู้ทดลองที่ 4.....50

4.2.4.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 50

4.2.4.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 50

4.2.4.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 53

4.2.5 ผลการทดลองกับข้อมูลผู้ทดลองที่ 5.....55

4.2.5.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 55

4.2.5.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 55

4.2.5.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 58

4.2.6 ผลการทดลองโดยรวมของข้อมูลผู้ทดลองทั้งหมด 60

4.2.6.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 60

4.2.6.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ 60

4.2.6.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 62

4.2.6.4 ผลลัพธ์ของการวัดค่าความระลึกของข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองทั้ง 5 คน.....63

บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ..... 65

5.1 สรุปผลการวิจัย.....65

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 ข้อเสนอแนะ.....66
เอกสารอ้างอิง.....67
ภาคผนวก ก งานวิจัยที่ตีพิมพ์.....68
ประวัติผู้เขียน.....75

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่ **หน้า**

ตารางที่ 2.1 เปรียบเทียบชนิดของสัญญาณไฟฟ้าของ สมอง หัวใจ และกล้ามเนื้อ....4
ตารางที่ 2.2 รายชื่อกล้ามเนื้อและหน้าที่ของกล้ามเนื้อที่ใช้ในทดลองในงานวิจัย....6
ตารางที่ 2.3 ตัวอย่างข้อมูล Confusion Matrix ขนาด 4 X 4....24
ตารางที่ 4.1 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 1....35
ตารางที่ 4.2 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 1....36
ตารางที่ 4.3 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 1....37
ตารางที่ 4.4 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 1 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ....38
ตารางที่ 4.5 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 1 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน....38
ตารางที่ 4.6 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 2....40
ตารางที่ 4.7 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 2....41
ตารางที่ 4.8 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 2....42
ตารางที่ 4.9 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 2 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ....43
ตารางที่ 4.10 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 2 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน....43
ตารางที่ 4.11 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 3....45
ตารางที่ 4.12 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 3....46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.13 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ และข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน กับข้อมูลผู้ทดลองที่ 3 47
ตารางที่ 4.14 ผลลัพธ์การจำแนกข้อมูลผู้ทดลอง 3 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 48
ตารางที่ 4.15 ผลลัพธ์การจำแนกข้อมูลผู้ทดลอง 3 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน 48
ตารางที่ 4.16 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 4 50
ตารางที่ 4.17 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 4 51
ตารางที่ 4.18 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ และข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน กับข้อมูลผู้ทดลองที่ 4 52
ตารางที่ 4.19 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 4 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 53
ตารางที่ 4.20 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 4 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน 53
ตารางที่ 4.21 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 5 55
ตารางที่ 4.22 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 5 56
ตารางที่ 4.23 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ และข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน กับข้อมูลผู้ทดลองที่ 5 57
ตารางที่ 4.24 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 5 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ 58
ตารางที่ 4.25 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 5 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน 58
ตารางที่ 4.26 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของข้อมูลโดยรวม 60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.27 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองทั้งหมด.....61
ตารางที่ 4.28 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองทั้งหมดที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ.....62
ตารางที่ 4.29 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองทั้งหมดที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน.....62
ตารางที่ 4.30 ผลลัพท์ค่าความระลึกของคลาสที่ได้ 100% ในชุดข้อมูลที่ให้ผลการจำแนกดีที่สุด 2 อันดับของแต่ละข้อมูลผู้ทดลอง.....64

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่ **หน้า**

รูปที่ 2.1 ตัวอย่างขั้วไฟฟ้าแบบเข็ม (ด้านซ้าย) และขั้วไฟฟ้าแบบติดที่ผิวหนัง (ด้านขวา).....5
รูปที่ 2.2 แผนภาพแสดงกล้ามเนื้อแขนด้านนอก (ด้านซ้าย) และกล้ามเนื้อแขนด้านใน (ด้านขวา).....6
รูปที่ 2.3 รูปแบบและตัวอย่างของเวกเตอร์แถว (ด้านบน) และเวกเตอร์หลัก (ด้านล่าง).....7
รูปที่ 2.4 รูปแบบเมทริกซ์.....8
รูปที่ 2.5 ตัวอย่างเมทริกซ์จัตุรัส (ด้านบนซ้าย) เมทริกซ์ที่ไม่เป็นจัตุรัส (ด้านบนขวา).....8
รูปที่ 2.6 ตัวอย่างข้อมูลเมทริกซ์ A ขนาด 4 X 4.....9
รูปที่ 2.7 ตัวอย่างการแยกองค์ประกอบของข้อมูลด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน.....10
รูปที่ 2.8 แผนภาพแสดงข้อมูลจริง.....12
รูปที่ 2.9 แผนภาพแสดงโครงสร้างต้นไม้ตัดสินใจจากข้อมูลในรูปที่ 2.8.....13
รูปที่ 2.10 ขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัว.....16
รูปที่ 2.11 แผนภาพแสดงการระบุคลาส เมื่อกำหนดให้ค่าสมาชิกเพื่อนบ้านเป็น 5.....16
รูปที่ 2.12 โครงสร้างของเรเดียลเบสิสฟังก์ชัน (Radial Basis Neural Network).....18
รูปที่ 2.13 ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน.....19
รูปที่ 2.14 ขั้นตอนการเลือกชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบด้วยวิธี (*k*-fold cross-validation).....20
รูปที่ 2.15 ตารางเปรียบเทียบ (Confusion Matrix).....22
รูปที่ 3.1 การเคลื่อนไหวพื้นฐานของการจับใน 6 รูปแบบ.....27
รูปที่ 3.2 ตัวอย่างคลื่นสัญญาณไฟฟ้ากล้ามเนื้อก่อนและหลังรวมช่องสัญญาณ.....28
รูปที่ 3.3 ตัวอย่างการเปลี่ยนข้อมูลแบบเวกเตอร์เป็นข้อมูลแบบเมทริกซ์.....30
รูปที่ 3.4 ขั้นตอนการสกัดคุณลักษณะด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน.....30
รูปที่ 4.1 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 1.....36
รูปที่ 4.2 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 1.....39
รูปที่ 4.3 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 2 ในแต่ละวิธีการจำแนกประเภท.....41

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 4.4 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 2 44
รูปที่ 4.5 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 3 ในแต่ละวิธีการจำแนกประเภท 46
รูปที่ 4.6 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 3 49
รูปที่ 4.7 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 4 ในแต่ละวิธีการจำแนกประเภท 51
รูปที่ 4.8 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 4 54
รูปที่ 4.9 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 5 ในแต่ละวิธีการจำแนกประเภท 56
รูปที่ 4.10 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 5 59
รูปที่ 4.11 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองทั้งหมด ในแต่ละวิธีการจำแนกประเภท 61
รูปที่ 4.12 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองทั้งหมด 63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในทางวิทยาศาสตร์การแพทย์ สมอง ประสาทและกล้ามเนื้อประกอบด้วยเซลล์ที่มีความไวต่อสิ่งเร้า เมื่อเซลล์เหล่านั้นถูกกระตุ้นโดยสิ่งเร้าก็จะทำการสร้างคลื่นสัญญาณไฟฟ้าแล้วส่งออกมา โดยมีเครื่องมือที่สามารถวัดผลและเก็บบันทึกเพื่อนำคลื่นสัญญาณไฟฟ้ามาวิเคราะห์ได้ คลื่นสัญญาณไฟฟ้ามีประโยชน์ในการช่วยวินิจฉัยและรักษาทางการแพทย์ได้ โดยคลื่นสัญญาณไฟฟ้าในร่างกายของคนเรามีดังนี้ Electrooculography (EOG) เป็นคลื่นสัญญาณไฟฟ้าที่เกิดขึ้นในดวงตา Electrocardiogram (ECG หรือ EKG) เป็นคลื่นสัญญาณไฟฟ้าจากการทำงานของหัวใจ Electroencephalogram (EEG) เป็นคลื่นสัญญาณไฟฟ้าจากสมองและ Electromyography (EMG) เป็นคลื่นสัญญาณไฟฟ้าจากกล้ามเนื้อ ดังนั้นในกิจกรรมต่าง ๆ ของมนุษย์จะมีคลื่นสัญญาณไฟฟ้าเกิดขึ้นเสมอ แม้กระทั่งการนอนหลับก็สามารถเกิดคลื่นสัญญาณไฟฟ้าได้ คลื่นสัญญาณไฟฟ้าสามารถนำมาวิเคราะห์ได้หลายอย่าง เช่น การวิเคราะห์คลื่นสัญญาณไฟฟ้าสมองเพื่อช่วยในการวินิจฉัยโรคลมชัก การวิเคราะห์การทำดาวสวิงด้วยคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ การวิเคราะห์คลื่นสัญญาณไฟฟ้าหัวใจที่ผิดปกติ เป็นต้น นอกเหนือจากนี้ยังสามารถช่วยเหลือบุคคลผู้พิการที่สูญเสียอวัยวะบางส่วน ซึ่งถือเป็นบุคคลที่ควรจะได้รับความช่วยเหลือด้านใดด้านหนึ่งเพื่อให้สามารถปฏิบัติกิจกรรมในชีวิตประจำวันหรือเข้าไปมีส่วนร่วมทางสังคมได้อย่างบุคคลทั่วไปได้

ในงานวิจัยนี้มุ่งเน้นปัญหาการลดปริมาณข้อมูลที่ใช้ในการจำแนกการเคลื่อนไหวมือโดยใช้คลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัส คลื่นสัญญาณไฟฟ้าจะถูกเก็บผ่านเครื่องวัดโดยจะเก็บข้อมูลเป็นความแรงของสัญญาณคลื่นไฟฟ้ารวมกับความถี่ของสัญญาณคลื่นไฟฟ้าในหนึ่งหน่วยเวลา จำนวนของคุณลักษณะของข้อมูลนั้นจะเพิ่มขึ้นตามปริมาณเวลาที่ใช้ในการเก็บ ซึ่งจะทำให้ใช้เวลาในวิเคราะห์คลื่นสัญญาณไฟฟ้านานขึ้น ดังนั้นเพื่อลดเวลาในการประมวลผลลง จึงจำเป็นต้องลดจำนวนคุณลักษณะของข้อมูลลง โดยผู้วิจัยได้ใช้วิธีการลดคุณลักษณะของข้อมูลด้วยวิธี ซิงกูลาร์แวลูดีคอมโพซิชัน (Singular Value Decomposition) ซึ่งจะทำให้จำนวนคุณลักษณะของข้อมูลที่ใช้ในการวิเคราะห์มีขนาดเล็กลง และเพิ่มประสิทธิภาพในการจำแนกข้อมูลมากขึ้นโดยเปรียบเทียบกับวิธีการเรียนรู้ของเครื่อง 5 วิธี คือ ต้นไม้ตัดสินใจ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว ขั้นตอนวิธีแบบเบย์อย่างง่าย ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และอัลกอริทึมเอสเอ็มโอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 วัตถุประสงค์ของงานวิจัย

1) เพื่อนำเสนอวิธีการลดจำนวนคุณลักษณะของข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ โดยใช้วิธีการแปลงข้อมูลคุณลักษณะเดิมให้เป็นข้อมูลคุณลักษณะใหม่ที่ง่ายต่อการจำแนกประเภท
2) เพื่อนำเสนอวิธีการจำแนกประเภทคลื่นสัญญาณไฟฟ้ากล้ามเนื้อโดยใช้อัลกอริทึมเอสเอ็มโอ

1.3 สมมติฐานของงานวิจัย

การจำแนกประเภทของข้อมูลด้วยอัลกอริทึมเอสเอ็มโอ (SMO) มีพื้นฐานมาจากขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน (SVM) ซึ่งมีการใช้ฟังก์ชันแมพข้อมูลเข้าสู่ปริภูมิคุณลักษณะ (Feature space) ที่สูงขึ้นและสร้างไฮเปอร์เพลนที่เป็นเส้นตรงที่สามารถตัดแบ่งข้อมูลในแต่ละกลุ่มเป้าหมายได้อย่างเหมาะสมที่สุด ซึ่งหากนำมาใช้ในการจำแนกคลื่นไฟฟ้ากล้ามเนื้อที่ใช้วิธีการสกัดคุณลักษณะของข้อมูลแล้วน่าจะจำแนกประเภทได้ดี โดยที่วิธีการที่ใช้ในการสกัดคุณลักษณะของข้อมูลนั้น ผู้วิจัยคาดว่าหากเปลี่ยนข้อมูลเวกเตอร์เป็นข้อมูลเมทริกซ์แล้วใช้วิธีซิงกูลาร์แวลูดีคอมโพซิชันในการดึงคุณลักษณะของข้อมูลออกมา ก็น่าจะเป็นการช่วยลดจำนวนคุณลักษณะที่มีอยู่เดิมลดลงในปริมาณที่มากได้และน่าจะเป็นการช่วยเพิ่มประสิทธิภาพของการจำแนกข้อมูลให้สูงขึ้นอีกด้วย

1.4 ขอบเขตงานวิจัย

1) ชุดข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อที่ใช้ในการทดลองเกิดจากการเคลื่อนไหวมือใน 6 รูปแบบ คือ
    (1) การจับวัตถุทรงกลม (Spherical)
    (2) การจับวัตถุขนาดเล็ก (Tip)
    (3) การจับวัตถุที่แบนบาง (Lateral)
    (4) การจับวัตถุทรงกระบอก (Cylindrical)
    (5) การจับวัตถุที่มีน้ำหนักมาก (Hook or Snap)
    (6) การจับวัตถุโดยหันฝ่ามือเข้าหาตัวเอง (Palmar)
2) เทคนิคการเรียนรู้ของเครื่องที่ใช้ในงานวิจัยนี้ได้แก่ อัลกอริทึมเอสเอ็มโอ (SMO) ต้นไม้ตัดสินใจ (Decision Tree: *C4.5*) ตัวจำแนกแบบเบย์อย่างง่าย (Naive Bayes Classifier) เพื่อนบ้านใกล้สุด *k* ตัว (*k*-Nearest neighbor) และโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial Basis Function Neural Networks)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1) สามารถประยุกต์ใช้วิธีการลดคุณลักษณะกับปัญหาอื่นได้
2) สามารถนำไปประยุกต์ใช้ได้ในหลาย ๆ อุตสาหกรรม เช่น การแพทย์ เป็นต้น
3) สามารถนำไปประยุกต์ใช้ในงานเกี่ยวกับการควบคุมอุปกรณ์โดยใช้คลื่นสัญญาณไฟฟ้าได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในส่วนของบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องอันประกอบด้วยความรู้พื้นฐานเกี่ยวกับคลื่นสัญญาณไฟฟ้าในร่างกาย เครื่องมือและอุปกรณ์สำหรับตรวจวัดคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ ชนิดของกล้ามเนื้อที่ใช้ทดลองในงานวิจัย การลดจำนวนมิติของข้อมูล ความรู้พื้นฐานเกี่ยวกับเวกเตอร์และเมทริกซ์ การวิเคราะห์องค์ประกอบ การเรียนรู้ของเครื่อง และวิธีการวัดประสิทธิภาพของตัวจำแนกประเภท

2.1 ความรู้พื้นฐานเกี่ยวกับคลื่นสัญญาณไฟฟ้าในร่างกาย

คลื่นสัญญาณไฟฟ้าในร่างกายเกิดจากเนื้อเยื่อส่วนที่ไวต่อสิ่งเร้า เช่น สมอง เส้นประสาท และกล้ามเนื้อ ซึ่งเมื่อถูกกระตุ้นด้วยสิ่งเร้าจะสามารถสร้างสัญญาณไฟฟ้าและส่งผ่านสัญญาณไปตามเส้นใยประสาทกล้ามเนื้อ โดยที่คลื่นสัญญาณไฟฟ้าหัวใจ คลื่นสัญญาณไฟฟ้าสมองและคลื่นสัญญาณไฟฟ้ากล้ามเนื้อนั้นจะมีความแตกต่างกันขึ้นอยู่กับขนาดของความถี่ที่ตอบสนองและศักย์ไฟฟ้า [1] ดังตารางที่ 2.1

ตารางที่ 2.1 เปรียบเทียบชนิดของสัญญาณไฟฟ้าของ สมอง หัวใจ และกล้ามเนื้อ

ชนิดของสัญญาณไฟฟ้า	ศักย์ไฟฟ้า (mV)	ความถี่ตอบสนอง (Hz)
สมอง (Electroencephalography : EEG)	0.001-0.10	0.02-100
หัวใจ (Electrocardiogram : ECG or EKG)	0.02-3.0	0.1-30
กล้ามเนื้อ (Electromyography : EMG)	0.003-5.0	2-10,000

2.2 เครื่องมือและอุปกรณ์สำหรับตรวจวัดคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ

เครื่องมือและอุปกรณ์ที่ใช้ในการตรวจวัดสัญญาณไฟฟ้าประกอบด้วยตัวเครื่องและอิเล็กโทรดบันทึกสัญญาณเข้า โดยตัวเครื่องมักจะมีอุปกรณ์วงจรไฟฟ้าที่ทำหน้าที่ กรอง ขยายสัญญาณและแสดงผล ซึ่งอิเล็กโทรดหรือขั้วไฟฟ้าสำหรับการตรวจวัดสัญญาณไฟฟ้าสามารถแบ่งได้เป็น 2 ชนิด คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) ขั้วไฟฟ้าสำหรับกระตุ้น มักเป็นขั้วไฟฟ้าขนาดเล็กและเป็นแบบ 2 ขั้ว (Bipolar Electrode) ยึดติดกัน
2) ขั้วไฟฟ้าสำหรับบันทึกสัญญาณไฟฟ้า ซึ่งสามารถแบ่งออกเป็น 2 ชนิด คือ ขั้วไฟฟ้าแบบเข็ม (Needle Electrode) ซึ่งจะใช้การแทงขั้วเข็มรับสัญญาณไฟฟ้าเข้าไปในกล้ามเนื้อเพื่อบันทึกสัญญาณหรือศักย์ไฟฟ้าที่ใยกล้ามเนื้อโดยตรง ส่วนอีกแบบหนึ่งนั้นจะเป็นแบบที่ขั้วไฟฟ้าจะถูกแนบติดที่บริเวณผิวหนัง (Surface Electrode) เพื่อทำการวัดสัญญาณไฟฟ้า ซึ่งขั้วไฟฟ้าแบบเข็มและแบบติดที่ผิวหนังจะมีลักษณะที่แสดงได้ดังรูปที่ 2.1

รูปที่ 2.1 ตัวอย่างขั้วไฟฟ้าแบบเข็ม (ด้านซ้าย) และขั้วไฟฟ้าแบบติดที่ผิวหนัง (ด้านขวา)

2.3 ชนิดของกล้ามเนื้อที่ใช้ทดลองในงานวิจัย

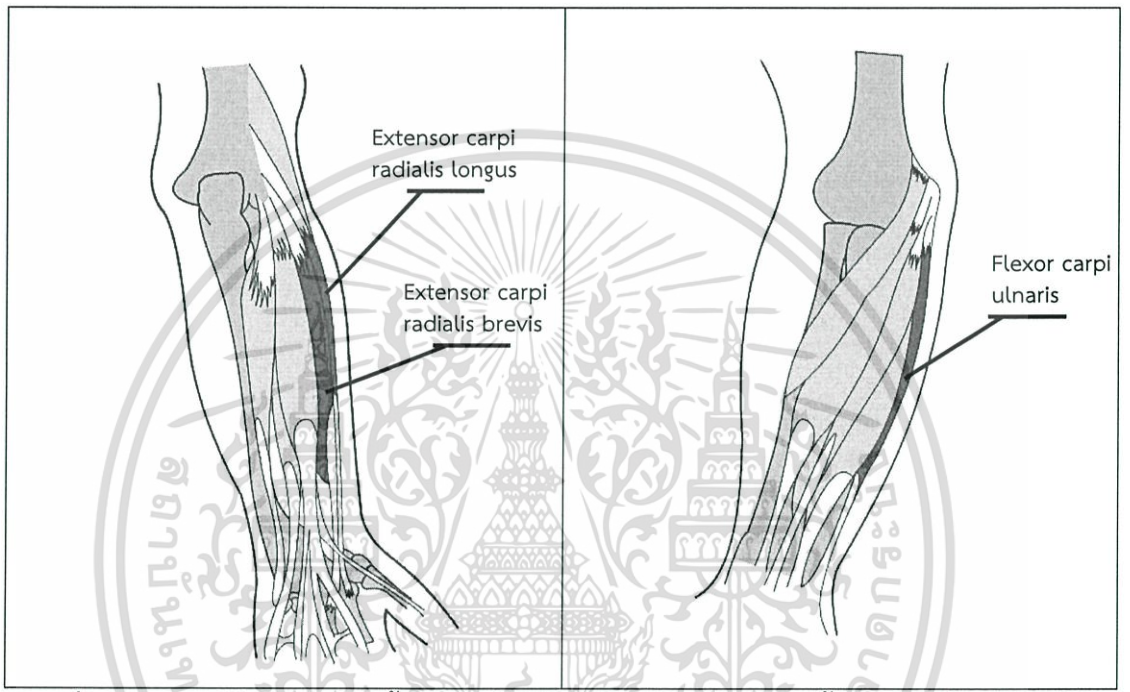
ระบบกล้ามเนื้อ (Muscular System) ในร่างกายของมนุษย์มีอยู่ด้วยกัน 3 ชนิด ได้แก่ กล้ามเนื้อลาย (Skeletal Muscle) กล้ามเนื้อเรียบ (Smooth Muscle) และกล้ามเนื้อหัวใจ (Cardiac Muscle) ซึ่งในส่วนนี้จะขอกล่าวถึงเฉพาะกล้ามเนื้อลายซึ่งเป็นชนิดที่มีการจัดเก็บชุดข้อมูลมาใช้ในงานวิจัยนี้

กล้ามเนื้อลาย (Skeletal Muscle) หรือ กล้ามเนื้อโครงร่าง (Skeletal Muscle) เป็นกล้ามเนื้อที่ยึดระหว่างกระดูก 2 ชิ้น หรืออาจยึดระหว่างกระดูกกับผิวหนังหรือผังพืด ซึ่งในร่างกายของคนเราจะมีกล้ามเนื้อลายทั้งหมด 792 มัดและมีน้ำหนักของกล้ามเนื้อลายรวมกันทั้งหมดประมาณ 40% ของน้ำหนักตัว กล้ามเนื้อลายมีหน้าที่เกี่ยวข้องกับการเคลื่อนไหวของร่างกายและเป็นส่วนประกอบของกล้ามเนื้อบริเวณขา แขน และลำตัว โดยกล้ามเนื้อชนิดนี้จะอยู่ภายใต้การควบคุมด้วยอำนาจจิตใจ (Voluntary Control) [2] ซึ่งในงานวิจัยนี้ได้ใช้ชุดข้อมูลที่ได้จากการเก็บคลื่นสัญญาณไฟฟ้าจากกล้ามเนื้อลายของแขนท่อนล่างดังแสดงในตารางที่ 2.2 และรูปที่ 2.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.2 รายชื่อกล้ามเนื้อและหน้าที่ของกล้ามเนื้อที่ใช้ในทดลองในงานวิจัย

ชื่อกล้ามเนื้อ	หน้าที่
1. Flexor carpi ulnaris	งอข้อมือ และหุบมือ เมื่อทำงานร่วมกับกล้ามเนื้อ extensor carpi ulnaris
2. Extensor carpi radialis longus	เหยียดข้อมือ และกางมือ
3. Extensor carpi radialis brevis	เหยียดข้อมือ และกางมือ



รูปที่ 2.2 แผนภาพแสดงกล้ามเนื้อแขนด้านนอก (ด้านซ้าย) และกล้ามเนื้อแขนด้านใน (ด้านขวา)

2.4 การลดจำนวนมิติของข้อมูล (Dimensionality Reduction)

มิติของข้อมูล (Dimension) หรือที่เรียกอีกอย่างว่า คุณลักษณะของข้อมูล (Feature) เป็นองค์ประกอบอย่างหนึ่งที่ใช้บอกถึงขนาดและความซับซ้อนของข้อมูล ซึ่งการมีมิติของข้อมูลที่มากจะส่งผลให้ขั้นตอนวิธีในการจำแนกประเภทของข้อมูลต้องใช้เวลานานและสิ้นเปลืองพื้นที่ในการจัดเก็บข้อมูล ดังนั้นการลดจำนวนมิติของข้อมูลจะทำให้ข้อมูลมีความซับซ้อนที่น้อยลง ซึ่งขั้นตอนวิธีในการลดจำนวนมิติของข้อมูลสามารถแบ่งเป็น 2 ประเภท ได้แก่

2.4.1 ขั้นตอนวิธีแบบเลือกคุณลักษณะของข้อมูล (Feature Selection)

ขั้นตอนวิธีแบบเลือกคุณลักษณะของข้อมูล จะเลือกสับเซต (Subset) ของคุณลักษณะที่เหมาะสมที่สุดออกมา โดยจะไม่มีการเปลี่ยนแปลงค่าใด ๆ ในคุณลักษณะของชุดข้อมูลที่ทำการลดมิติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.2 ขั้นตอนวิธีแบบสกัดคุณลักษณะของข้อมูล (Feature Extraction)

ขั้นตอนวิธีแบบสกัดคุณลักษณะของข้อมูลจะปรับปรุงหรือเปลี่ยนแปลงค่าของคุณลักษณะ ซึ่งจะทำให้ได้คุณลักษณะใหม่แทนที่คุณลักษณะเดิมของข้อมูลโดยจะมีการลดจำนวนมิติไปในตัวด้วย

2.5 ความรู้พื้นฐานเกี่ยวกับเวกเตอร์และเมทริกซ์

2.5.1 เวกเตอร์ (Vector)

ในทางคณิตศาสตร์ เวกเตอร์ [3] คือ ลำดับของชุดตัวเลขที่สามารถเขียนตามแนวนอนได้ ซึ่งเรียกว่า เวกเตอร์แถว (Row Vector) โดยมีขนาดเป็น n X 1 หรือเขียนตามแนวตั้ง เรียกว่า เวกเตอร์หลัก (Column Vector) โดยมีขนาดเป็น 1 X n ซึ่งเวกเตอร์ 1 เวกเตอร์จะประกอบด้วยตัวเลข n จำนวน และตัวเลขแต่ละตัวจะเรียกว่าสมาชิก (Element) ของเวกเตอร์ ตัวอย่างของเวกเตอร์แถวและเวกเตอร์หลักสามารถเขียนได้ดังรูปที่ 2.3

$$A = [a_1 \quad a_2 \quad \dots \quad a_n] \qquad A = [5 \quad 0 \quad 2 \quad 3 \quad 7]$$

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \qquad B = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 2 \end{bmatrix}$$

รูปที่ 2.3 รูปแบบและตัวอย่างของเวกเตอร์แถว (ด้านบน) และเวกเตอร์หลัก (ด้านล่าง)

จากรูปที่ 2.3 เวกเตอร์ A คือเวกเตอร์แถวที่สมาชิกตั้งแต่ a ลำดับที่ 1 ถึง a ลำดับที่ n ด้านบนขวา เวกเตอร์ A มีสมาชิกทั้งหมด 5 ตัวประกอบด้วย 5 0 2 3 และ 7 ส่วนเวกเตอร์ B คือ เวกเตอร์หลักที่มีข้อมูลสมาชิกตั้งแต่ b ลำดับที่ 1 ถึง b ลำดับที่ n มีการเรียงลำดับข้อมูลตามแนวตั้ง ดังตัวอย่างในรูปที่ 2.3 ด้านล่างขวา เวกเตอร์ B มีสมาชิกทั้งหมด 4 ตัว ประกอบด้วย 2 3 0 และ 2

2.5.2 เมทริกซ์ (Matrix)

เมทริกซ์ (Matrix) คือ การแสดงข้อมูล หรือตัวเลขชุดหนึ่ง ในรูปสี่เหลี่ยมมุมฉากที่ประกอบด้วยแถว (Row) และหลัก (Column) การระบุขนาดของเมทริกซ์จะใช้จำนวนแถวและ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวนหลัก โดยทั่วไปเมทริกซ์จะประกอบด้วย m แถวและ n หลัก ซึ่งสามารถเขียนเป็นรูปแบบของเมทริกซ์ได้ดังรูปที่ 2.4

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & a_{ij} & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

รูปที่ 2.4 รูปแบบเมทริกซ์

สมาชิกแต่ละตัวในเมทริกซ์จะมีซับสคริปต์ (subscript) อยู่ 2 ตัว เป็นตัวอ้างอิงตำแหน่งสมาชิกบนแถวและหลักนั้น ๆ ซึ่งตัวแรกคือแถวที่สมาชิกนั้นอยู่ และตัวที่สองคือหลักที่สมาชิกนั้นอยู่ เช่น กรณีของเมทริกซ์ A ในรูปที่ 2.5 a_{21} = 3 หมายถึงสมาชิกของเมทริกซ์ A ในแถวที่ 2 หลักที่ 1 มีค่าเป็น 1 และกรณีของเมทริกซ์ C ในรูปที่ 2.5 c_{22} = -3 หมายถึงสมาชิกของเมทริกซ์ C ในแถวที่ 2 หลักที่ 2 มีค่าเป็น -3 เป็นต้น ชนิดของเมทริกซ์มีหลายแบบ เช่น เมทริกซ์จัตุรัสเป็นเมทริกซ์ที่มีจำนวนแถวและหลักเท่ากัน เมทริกซ์ไม่จัตุรัสเป็นเมทริกซ์ที่มีจำนวนแถวและหลักไม่เท่ากัน เมทริกซ์เฉียงเป็นเมทริกซ์ที่มีสมาชิกในแนวเฉียง (Diagonal Element) มีค่าไม่เท่ากับ 0 แต่สมาชิกนอกแนวเฉียง (Off Diagonal Element) มีค่าเป็น 0 เมทริกซ์สลับเปลี่ยน (Transpose Matrix) เป็นเมทริกซ์ที่สลับสมาชิกจากแถวไปเป็นหลัก ทั้งหมดดังแสดงได้ดังรูปที่ 2.5

$$A = \begin{bmatrix} 5 & 0 \\ 3 & 1 \end{bmatrix} \qquad B = \begin{bmatrix} 5 & 6 \\ 0 & 1 \\ -2 & 4 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 8 \end{bmatrix} \qquad B^T = \begin{bmatrix} 5 & 0 & -2 \\ 6 & 1 & 4 \end{bmatrix}$$

รูปที่ 2.5 ตัวอย่างเมทริกซ์จัตุรัส (ด้านบนซ้าย) เมทริกซ์ที่ไม่เป็นจัตุรัส (ด้านบนขวา) เมทริกซ์เฉียง (ด้านล่างซ้าย) และเมทริกซ์สลับเปลี่ยน (ด้านล่างขวา)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 การวิเคราะห์องค์ประกอบ (Factor Analysis)

2.6.1 ซิงกูลาร์แวลูดีคอมโพซิชัน (Singular Value Decomposition : SVD)

ซิงกูลาร์แวลูดีคอมโพซิชัน [4] เป็นขั้นตอนวิธีหนึ่งที่จัดอยู่ในหมวดหมู่ของการวิเคราะห์องค์ประกอบ (Factor Analysis) ซึ่งเป็นเทคนิคที่ใช้ในการจัดกลุ่มหรือรวมคุณลักษณะที่มีความสัมพันธ์กันเข้าไว้ด้วยกัน ซิงกูลาร์แวลูดีคอมโพซิชันจะทำการแยกองค์ประกอบของเมทริกซ์ด้วยการแยกเมทริกซ์ออกเป็นผลคูณของ 3 เมทริกซ์ดังสมการที่ (2.1)

$$A_{m\times n} = U_{m\times m} S_{m\times n} V^{T}_{n\times n} \tag{2.1}$$

โดยที่ $A_{m\times n}$ คือ เมทริกซ์ที่ต้องการแยกค่าองค์ประกอบ ขนาด m x n

$U_{m\times m}$ คือ เมทริกซ์ยูนิแทรี (Unitary matrix) ขนาด m x m

$S_{m\times n}$ คือ เมทริกซ์เฉียง (Diagonal Matrix) ขนาด m x n

$V^{T}_{n\times n}$ คือ เมทริกซ์สลับเปลี่ยน (Transpose Matrix) ขนาด n x n

สำหรับค่าองค์ประกอบที่นำมาใช้ในงานวิจัย คือ ค่าเมทริกซ์ S โดยจะนำค่าข้อมูลเฉียงมาทำให้ข้อมูลอยู่ในรูปแบบเวกเตอร์

ตัวอย่าง กำหนดให้ข้อมูลเมทริกซ์ A มีขนาด 4 X 4 ดังรูปที่ 2.6

A =

20	5	15	1
32	8	10	72
11	16	7	8
10	0	0	0

รูปที่ 2.6 ตัวอย่างข้อมูลเมทริกซ์ A ขนาด 4 X 4

เมื่อคำนวณหาค่าองค์ประกอบด้วยขั้นตอนวิธีซิงกูลาร์แวลูดีคอมโพซิชันตามสมการที่ (2.1) จะได้ เมทริกซ์ใหม่ 3 เมทริกซ์คูณกัน ดังรูปที่ 2.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Matrix of A					Matrix of U					Matrix of S					Matrix of V				
20	5	15	1	=	-0.15	0.81	0.41	-0.37	X	82.17	0	0	0	X	-0.44	0.62	0.40	0.49
32	8	10	72		-0.96	-0.23	0.07	-0.02		0	26.24	0	0		-0.13	0.37	-0.90	0.15
11	16	7	8		-0.18	0.47	-0.83	0.18		0	0	11.90	0		-0.16	0.50	0.09	-0.84
10	0	0	0		-0.05	0.23	0.34	0.90		0	0	0	5.49		-0.86	-0.47	-0.08	-0.12

รูปที่ 2.7 ตัวอย่างการแยกองค์ประกอบของข้อมูลด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน

2.6.2 การวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ (Linear Discriminat Analysis : LDA)

การวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ [5] เป็นขั้นตอนวิธีหนึ่งที่จัดอยู่ในหมวดหมู่ของการวิเคราะห์องค์ประกอบ (Factor Analysis) เช่นเดียวกับซิงกูลาร์แวลูดีคอมโพซิชัน ซึ่งการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ยังสามารถใช้ในการจำแนกประเภท (Classification) ได้อีกด้วย แต่ในบทนี้จะกล่าวถึงการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์เฉพาะในส่วนของการลดจำนวนมิติ(Dimensionlity Reduction) ของข้อมูลเท่านั้น ในการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์เป็นการหาปริภูมิย่อยใหม่ของข้อมูลที่ง่ายต่อการจำแนก โดยข้อมูลใหม่ในปริภูมิย่อยใหม่สามารถคำนวณได้จากสมการที่ (2.2)

$$Z = W^T X \tag{2.2}$$

โดยที่ Z คือ ข้อมูลเวกเตอร์ใหม่

X คือ ข้อมูลเวกเตอร์เดิม

W คือ เมทริกซ์ของไอเกนเวกเตอร์ (eigenvector) ที่เรียงตามค่าไอเกนแวลู (eigenvalue) จากมากไปน้อย

จากสมการที่ (2.2) เมทริกซ์ W คำนวณได้จากการหาค่าไอเกนเวกเตอร์ (eigenvector) และค่าไอเกนแวลู (eigenvalue) จาก $S_w^{-1}S_b$ โดยที่ S_w หาได้จากสมการที่ (2.3)

$$S_w = \sum_{i=1}^{K} S_i \tag{2.3}$$

เมทริกซ์ S_i สามารถคำนวณได้ดังสมการที่ (2.4)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$S_i = \sum_{t=1}^{N_i} (x_t - m_i)(x_t - m_i)^T \quad (2.4)$$

โดยที่ N_i คือ จำนวนข้อมูลทั้งหมดของกลุ่มที่ i

x_t คือ ข้อมูลเวกเตอร์ตัวที่ t

m_i คือ ค่าเฉลี่ยของกลุ่มที่ i

และ S_B คำนวณได้จากสมการที่ (2.5)

$$S_B = \sum_{i=1}^{K} N_i (m_i - m)(m_i - m)^T \quad (2.5)$$

โดยที่ N_i คือ จำนวนข้อมูลทั้งหมดของกลุ่มที่ i

K คือ จำนวนของกลุ่มเป้าหมาย

m_i คือ ค่าเฉลี่ยของกลุ่มที่ i

m คือ ค่าเฉลี่ยรวมของข้อมูล

เมื่อคำนวณค่าของเมทริกซ์ $S_w^{-1}S_b$ แล้วให้คำนวณหาไอเกนเวกเตอร์และไอเกนแวลูจากเมทริกซ์ดังกล่าว จากนั้นนำไอเกนเวกเตอร์มาเรียงต่อกันโดยเรียงตามค่าไอเกนแวลูที่มีค่ามากไปหาค่าไอเกนแวลูที่มีค่าน้อย ซึ่งจะได้เมทริกซ์ W สำหรับการลดมิติของข้อมูลจะทำการตัดไอเกนเวกเตอร์ที่มีค่าไอเกนแวลูที่น้อยออกจากเมทริกซ์ W หรือเลือกไอเกนเวกเตอร์ตามจำนวนมิติของข้อมูลที่ต้องการแล้วทำการตัดไอเกนเวกเตอร์ในส่วนที่เหลือออกจากเมทริกซ์ W จากทางด้านขวา

2.7 การเรียนรู้ของเครื่อง (Machine Learning)

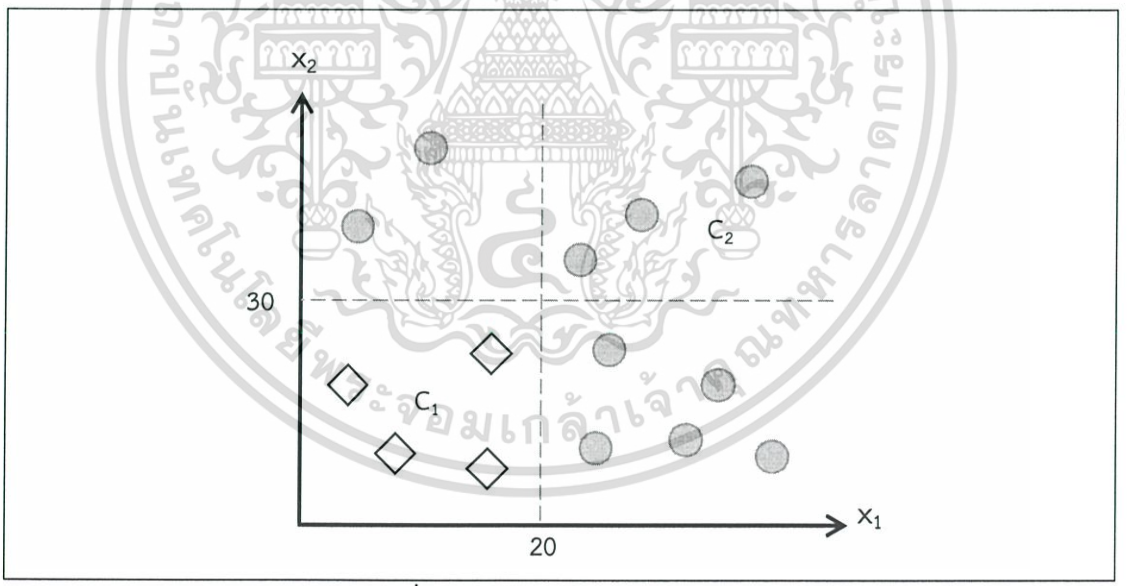
การเรียนรู้ของเครื่องเป็นสาขาหนึ่งของสาขาปัญญาประดิษฐ์ ซึ่งมีวัตถุประสงค์ในการทำให้เครื่องจักรกลกลหรือระบบคอมพิวเตอร์มีความสามารถในการเรียนรู้ หรือแม้กระทั่งทำให้เครื่องสามารถเรียนรู้เพื่อทำการตัดสินใจหรือทำนายผลจากข้อมูลที่ป้อนไปให้ได้ โดยการเรียนรู้ของเครื่องนั้นสามารถแบ่งได้เป็น 3 ประเภท คือ การเรียนรู้แบบมีผู้สอน (Supervised Learning) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) และการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) สำหรับในส่วนนี้จะกล่าวถึงเฉพาะการเรียนรู้แบบมีผู้สอน (Supervised Learning) ซึ่งมีความเกี่ยวข้องกับงานวิจัยนี้เท่านั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การเรียนรู้แบบมีผู้สอน (Supervised Learning) คือ การสอนเครื่องจักรกลกลหรือเครื่องคอมพิวเตอร์ โดยการป้อนหรือให้ข้อมูลที่ใช้สำหรับเรียนรู้ (Training Data) กับเครื่องนั้น ๆ โดยจะมีการป้อนข้อมูลผลลัพธ์เป้าหมาย (Target) ไปด้วยเพื่อสอนให้เครื่องรู้ว่าข้อมูลนั้นอยู่ในกลุ่มอะไร โดยเมื่อเครื่องได้ผ่านขั้นตอนการเรียนรู้แล้ว เมื่อถึงขั้นตอนการทดสอบ ผู้สอนก็จะใช้ข้อมูลที่ได้เตรียมไว้สำหรับทดสอบ (Testing Data) มาป้อนให้กับเครื่อง ซึ่งเมื่อเครื่องให้ผลลัพธ์กลับออกมาก็จะนำผลลัพธ์ที่ได้นั้นมาทำการเปรียบเทียบกับผลลัพธ์ที่คาดหวังเพื่อวัดผลประสิทธิภาพในการเรียนรู้ของเครื่องต่อไป

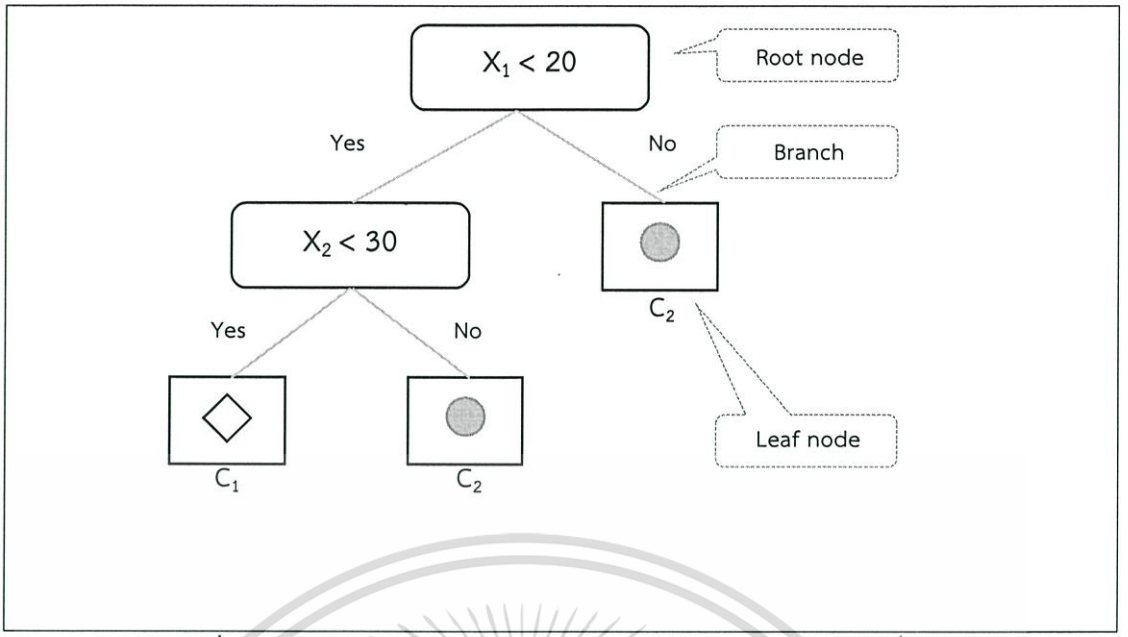
2.7.1 ขั้นตอนวิธีแบบซีโฟพอยต์ไฟว์ (C4.5)

ต้นไม้ตัดสินใจนั้นจะมีลักษณะที่เป็นโครงสร้างแบบต้นไม้กลับหัว ซึ่งจะประกอบไปด้วยโหนด (Node) ซึ่งแต่ละโหนดจะมีคุณลักษณะ (Attribute) ที่แตกต่างออกไป กล่าวคือ โหนดรากหรือโหนดเริ่มต้น (Root Node) จะเป็นส่วนของชั้นบนสุดและมีเพียงโหนดเดียว ในขณะที่กิ่งของต้นไม้หรือกิ่ง (Branch) ของโหนด จะเป็นส่วนที่แสดงค่าคุณลักษณะของโหนดด้านบน และจะมีโหนดใบ (Leaf Node) หรือโหนดสุดท้าย (Terminal Node) เป็นส่วนที่อยู่ล่างสุดของต้นไม้ตัดสินใจ โดยจะใช้เป็นส่วนของการระบุคลาส (Class) ซึ่งก็หมายถึงผลลัพธ์ที่ได้จากการทำนายดังแสดงในรูปที่ 2.8



รูปที่ 2.8 แผนภาพแสดงข้อมูลจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 แผนภาพแสดงโครงสร้างต้นไม้ตัดสินใจจากข้อมูลในรูปที่ 2.8

จากรูปที่ 2.8 จะเห็นได้ว่าข้อมูลของ x_1 ที่มีค่าที่มากกว่าหรือเท่ากับ 20 จะเป็นคลาส C_2 ทั้งหมด แต่หาก x_1 มีค่าที่น้อยกว่า 20 จะต้องทำการพิจารณาข้อมูลของ x_2 เพิ่ม ซึ่งข้อมูลของ x_2 ที่มีค่ามากกว่าหรือเท่ากับ 30 นั้นก็จะเป็นคลาส C_2 เช่นกัน ในขณะที่ข้อมูล x_2 ที่มีค่าน้อยกว่า 30 นั้นจะอยู่ในกลุ่มของ C_1 ซึ่งก็สามารถเขียนให้อยู่ในรูปแบบโครงสร้างต้นไม้ตัดสินใจได้ดังในรูปที่ 2.9 โดยมีโหนดรากเป็นคุณลักษณะ x_1 ที่มีเงื่อนไขคือค่าน้อยกว่า 20 โดยถ้าเงื่อนไขถูกต้องก็จะไปยังโหนดที่อยู่ทางซ้าย แต่ถ้าเงื่อนไขไม่ถูกต้องจะไปยังโหนดที่อยู่ทางขวา และจะคำนวณเช่นนี้ไปเรื่อย ๆ จนกว่าจะถึงโหนดสุดท้าย

อย่างไรก็ตาม ในปี 1993 Ross Quinlan ได้เสนอวิธีขั้นตอนวิธีซีโฟพอยต์ไฟว์ซึ่งเป็นขั้นตอนวิธีที่พัฒนามาจากขั้นตอนวิธีไอดีทรี (ID3) ที่ Ross Quinlan ได้เสนอไว้ในปี 1986 โดยมีสิ่งที่ได้พัฒนาขึ้นมาจากไอดีทรี คือ จะมีการเพิ่มการคำนวณค่าอัตราส่วนเกน (Gain ratio) สำหรับใช้ในการตัดสินใจเลือกคุณลักษณะเป็นโหนด ซึ่งค่าที่เพิ่มขึ้นมาคือค่าสารสนเทศของการแบ่งแยก (Split information) ของคุณลักษณะแต่ละตัวดังในสมการที่ (2.6)

$$SplitInfo_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \times log_2\left(\frac{|D_j|}{|D|}\right) \qquad (2.6)$$

โดยที่

$|D|$ คือ จำนวนของข้อมูล D

$|D_j|$ คือ จำนวนของข้อมูลหลังจากการแบ่งด้วยค่าคุณลักษณะที่ j ของคุณลักษณะ A

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

v คือ จำนวนของค่าที่เป็นไปได้ทั้งหมดในคุณลักษณะ A

โดยค่าสารสนเทศของการแบ่งแยกนี้จะถูกนำไปหารค่าเกนสารสนเทศแล้วจะได้ค่าอัตราส่วนเกน ดังในสมการที่ (2.7) ซึ่งอัตราส่วนเกนที่มีค่ามากที่สุดนั้น ก็จะถูกเลือกใช้ในการแบ่งคุณลักษณะของต้นไม้ตัดสินใจ

$$GainRatio(D) = \frac{Gain(A)}{SplitInfo_A(D)} \tag{2.7}$$

2.7.2 ขั้นตอนวิธีแบบเบย์อย่างง่าย (Naïve bayes)

ขั้นตอนวิธีแบบเบย์อย่างง่ายเป็นขั้นตอนวิธีที่ใช้หลักการของทฤษฎีความน่าจะเป็นในการจำแนกข้อมูลโดยใช้พื้นฐานมาจากหลักการของความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability) ดังสมการที่ (2.8)

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{2.8}$$

โดยที่

$P(A|B)$ คือ ความน่าจะเป็นของเหตุการณ์ A โดยที่เหตุการณ์ B เกิดขึ้นก่อน

$P(A \cap B)$ คือ ความน่าจะเป็นของเหตุการณ์ A และ เหตุการณ์ B ที่เกิดขึ้นร่วมกัน

$P(B)$ คือ ความน่าจะเป็นที่เหตุการณ์ B เกิดขึ้น

ดังนั้นในลักษณะเดียวกันสามารถเขียน $P(B|A)$ ได้ดังสมการที่ (2.9)

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{2.9}$$

จากสมการ (2.8) และสมการ (2.9) เราสามารถเขียนสมการใหม่ได้เป็นสมการที่ (2.10)

$$P(B|A) = \frac{P(A|B) \times P(B)}{P(A)} \tag{2.10}$$

โดยสมการที่ (2.10) มีชื่อเรียกว่า ทฤษฎีของเบย์ (Bayes Theorem) ซึ่งสามารถนำมาประยุกต์ใช้ในการจำแนกประเภทได้ โดยที่ความน่าจะเป็นของคลาสเมื่อกำหนดคุณลักษณะจะมีค่าเท่ากับความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

น่าจะเป็นของคุณลักษณะนั้นเมื่อกำหนดคลาสใด ๆ คูณกับความน่าจะเป็นของคลาสนั้นหารด้วยความน่าจะเป็นของคุณลักษณะดังสมการที่ (2.11)

$$P(class|arrtribute) = \frac{P(attribute|class) \times P(class)}{P(attribute)} \quad (2.11)$$

โดยที่

$P(class\|arrtribute)$	คือ ความน่าจะเป็นในการเกิดคลาส(Class) เมื่อกำหนดคุณลักษณะ(Attribute)
$P(class)$	คือ ความน่าจะเป็นในการเกิดคลาส
$P(attribute)$	คือ ความน่าจะเป็นในการเกิดคุณลักษณะ

แต่สำหรับในกรณีที่ค่าคุณลักษณะเป็นค่าที่ไม่สามารถนับความถี่ได้ (Continues number) จะต้องคำนวณด้วยสมการประมาณค่าแทนดังสมการที่ (2.12)

$$P(arrtribute|class) = \frac{1}{\sqrt{2\pi\sigma_{ij}}} e^{-\frac{1}{2}\left(\frac{arrtribute_i - \mu_{ij}}{\sigma_{ij}}\right)^2} \quad (2.12)$$

2.7.3 ขั้นตอนวิธีแบบเพื่อนบ้านใกล้สุด *k* ตัว (*k*-Nearest neighbour : *k*-NN)

ในปี 1991 Aha และ Kibler ได้เสนอขั้นตอนวิธีการจำแนกโดยใช้ผลโหวตความคล้ายคลึงกันของข้อมูลเพื่อทำการระบุคลาส ซึ่งขั้นตอนวิธีแบบเพื่อนบ้านใกล้สุด *k* ตัว ในขั้นตอนแรกนั้นจะเป็นการกำหนดค่า *k* (ค่า *k* เป็นค่าจำนวนเต็มบวกที่ผู้ใช้งานระบุ เพื่อกำหนดจำนวนผลโหวต ซึ่งค่าที่ต้องกำหนดควรเป็นค่าที่เป็นเลขคี่) ขั้นตอนต่อมาจะเป็นการหาข้อมูลเพื่อนบ้าน โดยจะเลือกข้อมูลเรียนรู้ที่อยู่ใกล้กับข้อมูลที่ต้องการระบุคลาสมากที่สุดมา *k* ตัว และทำการหาจำนวนคลาสที่มีจำนวนมากที่สุดเพื่อนำมาเป็นผลลัพธ์สุดท้าย โดยสามารถเขียนขั้นตอนวิธีแบบเพื่อนบ้านใกล้สุด *k* ตัวได้ดังรูปที่ 2.10

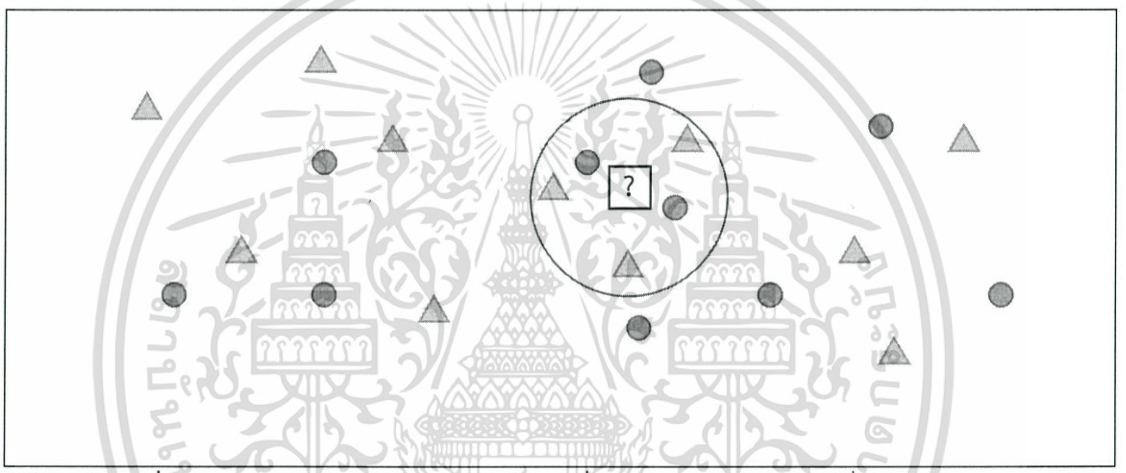
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว

1. กำหนดให้ค่า k คือ จำนวนสมาชิกเพื่อนบ้าน และ z เป็นข้อมูลที่ต้องการหาคลาส
2. คำนวณค่าระยะทางระหว่างข้อมูล ของข้อมูล z กับทุกข้อมูลในชุดข้อมูลฝึกสอน
3. เลือกข้อมูลฝึกสอนที่มีค่าระยะทางใกล้กับข้อมูล z มากที่สุด จำนวน k ตัว
4. คำนวณหาคลาสของข้อมูล z โดยหาจำนวนของข้อมูลในแต่ละคลาส จากกลุ่มข้อมูลที่เลือกมา k ตัว ที่มีจำนวนข้อมูลมากที่สุด แล้วจึงระบุคลาสของข้อมูล z เป็นคลาสนั้น

รูปที่ 2.10 ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว

ตัวอย่าง การจำแนกประเภทขั้นตอนวิธีแบบเพื่อนบ้านใกล้สุด k ตัว โดยกำหนดค่า $k = 5$



รูปที่ 2.11 แผนภาพแสดงการระบุคลาส เมื่อกำหนดให้ค่าสมาชิกเพื่อนบ้านเป็น 5

ในการระบุคลาสของขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัวดังรูปที่ 2.11 กำหนดให้มีการกระจายตัวของข้อมูล คือ คลาสสามเหลี่ยมและคลาสวงกลมดังภาพ และกำหนดค่าสมาชิกเพื่อนบ้านมีค่าเป็น 5 วิธีการนี้จะทำการเลือกข้อมูลที่มีระยะทางใกล้กับข้อมูลที่ต้องการระบุคลาสมากที่สุดมาเท่ากับค่า k ในที่นี้ k มีค่าเท่ากับ 5 ดังนั้นข้อมูลจะถูกเลือกมา 5 ตัว และจากนั้นจะทำการหาคลาสที่มีจำนวนมากที่สุดในจำนวนข้อมูลที่ถูกเลือกมาซึ่งในที่นี้จะมีคลาสสามเหลี่ยมอยู่ 3 จำนวน และมีคลาสวงกลมอยู่ 2 จำนวน ดังนั้นข้อมูลใหม่จะถูกระบุว่าเป็นคลาสสามเหลี่ยม เป็นต้น

สำหรับการหาระยะห่างของข้อมูลนั้น ใช้การหาค่าความคล้ายคลึงกันโดยการวัดระยะทางระหว่างจุด 2 จุดนั้น ซึ่งสามารถคำนวณได้หลายวิธี เช่น การวัดระยะทางแบบยูคลิด (Euclidean Distance) การวัดระยะทางแบบมาหาลาโนบิส (Mahalanobis Distance) การวัดระยะทางแบบโคซาย (Cosine Distance) การวัดระยะทางแบบเชบีเชฟ (Chebychev Distance) และการวัดระยะทางแบบสเปียแมน (Spearman Distance) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7.4 โครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial basis Function Neural Network : RBF)

ในปี 1988 Broomhead และ Lowe ได้เสนอโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial basis function neural network : RBF) ซึ่งเป็นขั้นตอนวิธีที่มีพื้นฐานโครงสร้างมาจากโครงข่ายประสาทเทียมที่เลียนแบบโครงข่ายประสาทในร่างกายมนุษย์ โดยโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันนั้นถูกคิดค้นเพื่อใช้ในการแก้ไขปัญหาการประมาณค่าในช่วง (Interpolation Problem) กับข้อมูลที่มีมิติสูง ข้อมูลนำเข้าจะถูกคำนวณด้วยฟังก์ชันเชิงรัศมีและสำหรับข้อมูลส่งออกจะถูกคำนวณด้วยฟังก์ชันเชิงเส้น ซึ่งการประมาณค่าช่วงสามารถคำนวณได้ตามสมการที่ (2.13)

$$F(x) = \sum_{i=1}^{N} w_i \cdot \varphi(\|x - x_i\|) \tag{2.13}$$

โดยที่ $\varphi(\cdot)$ คือ เบสิสฟังก์ชัน

w_i คือ ค่าน้ำหนัก (Weight)

$\|\cdot\|$ คือ นอร์มแบบยุคลิด (Euclidean Norm)

$x_i \in R^n$ คือ จุดศูนย์กลาง (Centers) ของเรเดียลเบสิสฟังก์ชัน

ตัวอย่างของเบสิสฟังก์ชันที่ใช้ในโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน ได้แก่

1) Inverse Multiquadrics

$$\varphi(r) = \frac{1}{(r^2 + c^2)^{1/2}} \tag{2.14}$$

โดย $c > 0$ และ $r \geq 0$

2) Gaussion Functions

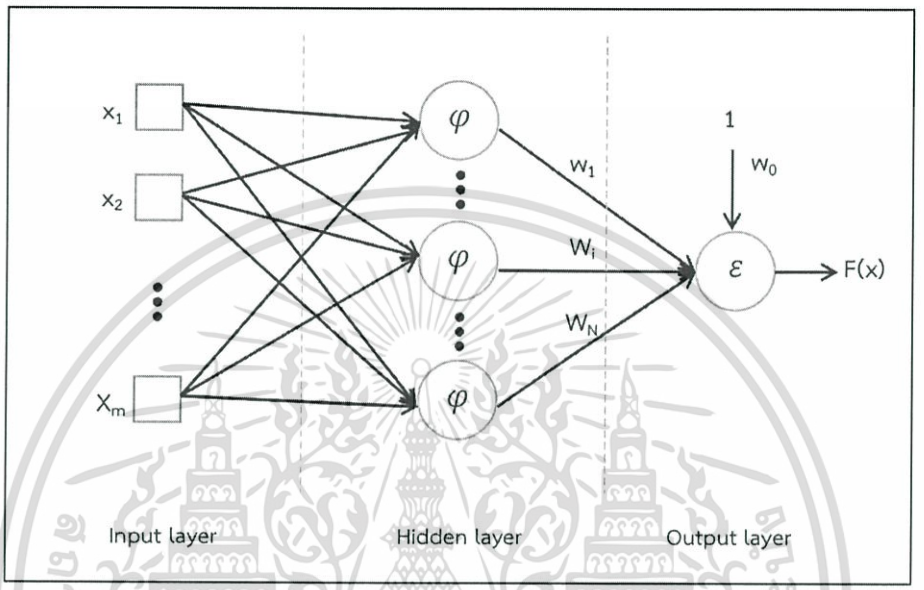
$$\varphi(r) = exp(-\frac{r^2}{2\sigma^2}) \tag{2.15}$$

โดย $\sigma > 0$ และ $r \geq 0$

จากปัญหาการประมาณค่าในช่วง (Interpolation problem) เราสามารถนำเรเดียลเบสิสฟังก์ชันมาใช้ในการสร้างโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ โดยโครงข่ายประสาทเทียมนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะทำการแมพข้อมูลจากปริภูมิข้อมูลนำเข้าไปยังปริภูมิคุณลักษณะเพื่อจำแนกประเภทบนปริภูมิคุณลักษณะ (Feature Space) ในกรณีที่ข้อมูลมีการกระจายตัวของข้อมูลต่ำ เมื่อนำฟังก์ชันเชิงรัศมีมาครอบข้อมูลจะทำให้เกิดความผิดพลาดสูงเพราะข้อมูลต่างคลาสกันอยู่ในฟังก์ชันเชิงรัศมีเดียวกัน ทำให้ข้อมูลปะปนกันและมีความผิดพลาดเกิดขึ้น รูปที่ 2.12 แสดงโครงสร้างของโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน (Radial Basis Neural Network)

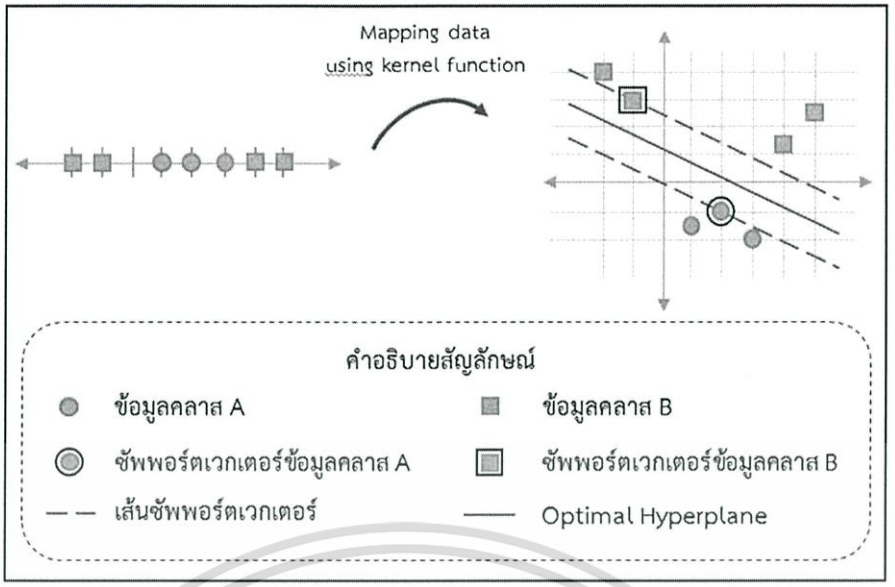


รูปที่ 2.12 โครงสร้างของเรเดียลเบสิสฟังก์ชัน (Radial Basis Neural Network)

2.7.5 อัลกอริทึมเอสเอ็มโอ (Sequential Minimal Optimizatio : SMO)

อัลกอริทึมเอสเอ็มโอพัฒนาโดยมีพื้นฐานมาจากขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน ซึ่งขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีนได้ถูกนำเสนอโดย Vapnik และ Chervonenkis ในปี 1963 เพื่อแก้ไขปัญหาที่ตัวจำแนกแบบโครงข่ายประสาทเทียมไม่สามารถหาเส้นตัดแบ่งข้อมูลที่เหมาะที่สุด (Optimal hyperplane) ได้เสมอ โดยที่หลักการทำงานของขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน คือ การแมพข้อมูลให้มีมิติของข้อมูลที่สูงขึ้น โดยอาศัยฟังก์ชันเคอร์เนล (Kernel Function) จากนั้นหาเส้นซัพพอร์ตเวกเตอร์ของแต่ละกลุ่ม แล้วจึงคำนวณหาเส้นตัดแบ่งที่เหมาะสมที่สุดโดยใช้ฟังก์ชันเชิงเส้นในการคำนวณ (Linear Function) สำหรับการตัดแบ่งข้อมูลด้วยขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีนสามารถแสดงได้ดังรูปที่ 2.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.13 ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีน

ขั้นตอนวิธีซัพพอร์ตเวกเตอร์แมชชีนใช้การหาเส้นตัดแบ่งที่เหมาะสมที่สุด โดยการหาตัวคูณลากรองจ์ (Lagrange multipliers) ที่ทำให้ฟังก์ชันวัตถุประสงค์มีค่าที่มากที่สุด ซึ่งสามารถคำนวณได้ตามสมการที่ (2.16)

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j d_i d_j k(x_i, x_j) \qquad (2.16)$$

โดยที่ Q คือ ฟังก์ชันวัตถุประสงค์

α คือ ค่าตัวคูณลากรองจ์ (Lagrange multipliers)

k คือ ฟังก์ชันเคอร์เนล (kernel function)

N คือ จำนวนข้อมูลฝึกสอนทั้งหมด

d_i คือ ผลลัพธ์ที่ต้องการของตัวที่ i

d_j คือ ผลลัพธ์ที่ต้องการของตัวที่ j

x_i คือ ข้อมูลเวกเตอร์ตัวที่ i

x_j คือ ข้อมูลเวกเตอร์ตัวที่ j

การคำนวณหาค่าตัวคูณลากรองจ์ที่เหมาะสมที่สุดหาได้จากการแก้ระบบสมการเชิงเส้นในรูปแบบเมทริกซ์ ซึ่งทำให้ความเร็วของการคำนวณด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนนั้นขึ้นอยู่กับขนาดของเมทริกซ์ ถ้าหากเมทริกซ์มีขนาดเล็กจะคำนวณได้เร็วแต่จะทำงานได้ช้าลงเมื่อขนาดของเมทริกซ์มีขนาดที่ใหญ่ขึ้น ดังนั้นในปี 1998 John Platt's [6] ได้เสนออัลกอริทึมเอสเอ็มโอซึ่งเป็นขั้นตอน

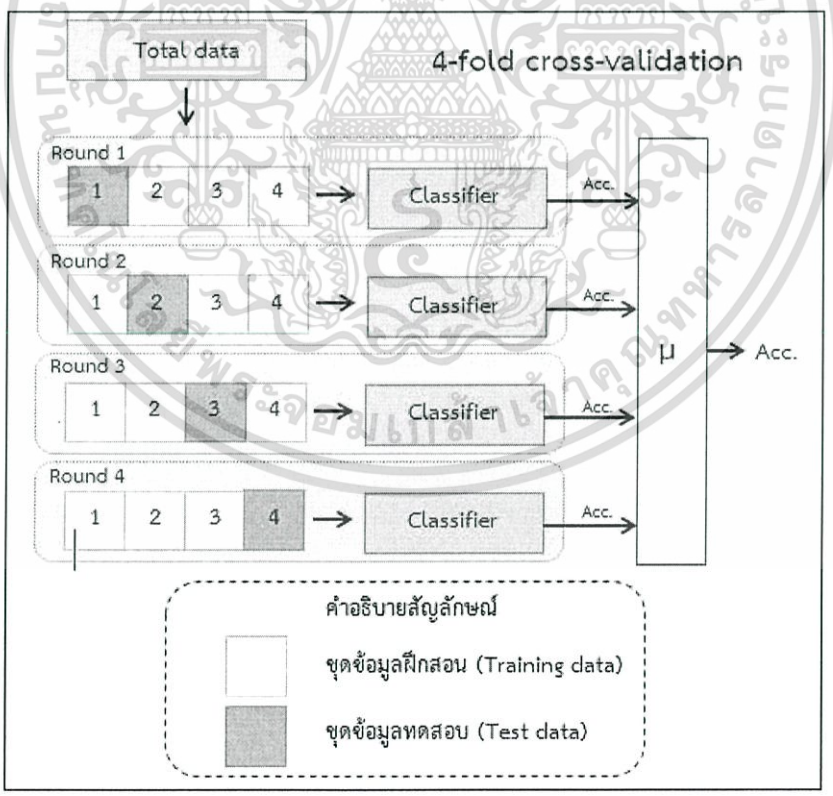
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิธีที่ปรับปรุงวิธีซัพพอร์ตเวกเตอร์แมชชีนให้ดีขึ้นโดยมี 3 ขั้นตอนหลักในอัลกอริทึมเอสเอ็มโอ ขั้นตอนแรกเป็นการเปลี่ยนวิธีคำนวณตัวคูณลากรองจ์จากวิธีการแก้ไขสมการเชิงเส้นเป็นการคำนวณด้วยวิธีการประมาณค่าตัวคูณลากรองจ์แทน ขั้นตอนที่สองใช้ฮิวริสติกในการเลือกตัวคูณที่เหมาะสมและขั้นตอนสุดท้าย คือ กระบวนการคำนวณค่าเอนเอียง (bias) ซึ่งข้อดีของอัลกอริทึมเอสเอ็มโอ คือ มีความเร็วในการคำนวณที่สูงขึ้นเพราะไม่ได้ใช้เมทริกซ์ในการคำนวณ

2.8 วิธีการวัดประสิทธิภาพของตัวจำแนกประเภท (Evaluation of Classifiers)

2.8.1 การตรวจสอบไขว้ (Cross-Validation)

การตรวจสอบไขว้ (Cross-Validation) เป็นเทคนิควิธีที่นำเสนอโดย ron ในปี 1995 มีพื้นฐานมาจากเทคนิคการสุ่มตัวอย่าง โดยการใช้งานจะเริ่มต้นจากการแบ่งข้อมูลออกเป็นส่วน ๆ ให้มีจำนวนที่เท่ากัน ซึ่งในการวัดประสิทธิภาพการทำนายตัวจำแนกประเภทนั้น จะทำการเลือกชุดข้อมูลทดสอบออกมา 1 ส่วน ในขณะที่ข้อมูลในส่วนอื่นนั้นจะถูกรวมกันเป็นชุดข้อมูลฝึกสอนแล้วจึงใช้ชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบนั้นสอนจำแนกประเภท โดยในขั้นตอนการเรียนรู้นั้น ข้อมูลทุกส่วนจะถูกสลับให้เป็นทั้งชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบจนครบตามจำนวนที่ได้แบ่งไว้ดังแสดงในรูปที่ 2.14



รูปที่ 2.14 ขั้นตอนการเลือกชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบด้วยวิธี (*k*-fold cross-validation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.14 ขั้นตอนการทำการตรวจสอบไขว้แบบ *k* ส่วน (k-fold cross-validation) จะเริ่มต้นจากกำหนดค่า *k* ซึ่งในที่นี้จะกำหนดให้มีค่าเท่ากับ 4 ดังนั้นข้อมูลทั้งหมดจะถูกแบ่งออกเป็นจำนวน 4 ส่วนเท่า ๆ กัน จากนั้นเมื่อทำการวัดประสิทธิภาพของตัวจำแนกประเภทในรอบที่ 1 ชุดข้อมูลในส่วนที่ 1 จะถูกกำหนดให้เป็นชุดข้อมูลทดสอบ ส่วนชุดข้อมูลที่เหลือ ซึ่งก็คือชุดข้อมูลที่ 2 3 และ 4 นั้นจะถูกกำหนดให้เป็นชุดชุดข้อมูลฝึกสอน ซึ่งเมื่อทำการสอนและทดสอบกับตัวจำแนกประเภทเรียบร้อยแล้วแล้ว ก็จะได้ค่าความถูกต้อง (accuracy) ของตัวจำแนกประเภทในรอบที่ 1 กลับออกมา แล้วจึงทำการวัดประสิทธิภาพตัวจำแนกประเภทในรอบที่ 2 ซึ่งชุดข้อมูลในส่วนที่ 2 ก็จะถูกกำหนดให้เป็นชุดข้อมูลสำหรับทดสอบ โดยมีชุดข้อมูลที่เหลือคือชุดข้อมูลที่ 1 3 และ 4 เป็นชุดข้อมูลฝึกสอน และเมื่อทำการสอนและทดสอบกับตัวจำแนกประเภทสำหรับรอบที่ 2 เรียบร้อยแล้ว ก็จะได้ค่าความถูกต้อง (accuracy) ของตัวจำแนกประเภทในรอบที่ 2 กลับออกมา จากนั้นจึงจะเริ่มทำการวัดประสิทธิภาพของตัวจำแนกประเภทในรอบที่ 3 ต่อไป โดยจะทำการวนสลับการเลือกชุดข้อมูลฝึกสอนและชุดข้อมูลทดสอบไปจนครบตามจำนวนที่ได้กำหนดให้กับค่า *k* และเมื่อทำการคำนวณจนรอบการทดสอบสุดท้ายเสร็จเรียบร้อย ก็ให้นำค่าความถูกต้องที่ได้จากทั้ง 4 รอบมาหาค่าเฉลี่ย ซึ่งจะได้ค่าความถูกต้องของตัวจำแนกประเภทโดยใช้ การตรวจสอบไขว้แบบ 4 ส่วน (4-fold cross-validation)

ในงานวิจัยต่าง ๆ ที่ใช้ การตรวจสอบไขว้แบบ *k* ส่วน (k-fold cross-validation) นั้น ส่วนใหญ่แล้วมักจะกำหนดให้ค่า *k* มีค่าเป็น 2 5 และ 10 แต่อย่างไรก็ตามการเลือกใช้จำนวนค่า *k* นั้นควรเลือกในจำนวนที่น้อย เพราะเป็นการจะเป็นการแสดงให้เห็นถึงความสามารถของตัวจำแนกประเภทที่ดีได้ เนื่องจากเมื่อทำการกำหนดจำนวนค่า *k* ให้มีค่าน้อยแล้วนั้น ก็จะทำให้ชุดข้อมูลถูกแบ่งไปเป็นข้อมูลทดสอบมากขึ้นและได้ชุดข้อมูลฝึกสอนที่น้อยลงตามอัตราส่วนเมื่อเปรียบเทียบกับการกำหนดจำนวน fold ที่มาก เช่น มีข้อมูลทั้งหมด 100 ข้อมูล ถ้ากำหนดค่า $k = 2$ จะเป็นการกำหนดให้ชุดข้อมูลทดสอบมีข้อมูล 50 ข้อมูลและชุดข้อมูลฝึกสอนมีจำนวนข้อมูล 50 ข้อมูลเพื่อใช้ในการทดลองกับตัวจำแนกประเภทในแต่ละรอบการทดสอบ หรือถ้ากำหนดให้ $k = 4$ จะเป็นการกำหนดให้ชุดข้อมูลทดสอบมีจำนวนข้อมูล 25 ข้อมูลและชุดข้อมูลฝึกสอนมีจำนวนข้อมูล 75 ข้อมูลสำหรับใช้ในการทดลองกับตัวจำแนกประเภทในแต่ละรอบ หรือในอีกตัวอย่างเช่นการกำหนดให้ $k = 10$ ซึ่งหมายถึงชุดข้อมูลทดสอบนั้นจะมีจำนวนข้อมูล 10 จำนวนและชุดข้อมูลฝึกสอนมีจำนวนข้อมูล 90 ข้อมูลสำหรับใช้ในการทดลองกับตัวจำแนกประเภทในแต่ละรอบ

2.8.2 Confusion Matrix

Confusion Matrix เป็นตารางเปรียบเทียบผลลัพธ์การทำนายที่ได้จากตัวจำแนกประเภทกับผลลัพธ์จริง โดยจะใช้ในการประเมินประสิทธิภาพของตัวจำแนกประเภทที่แสดงได้ดังรูปที่ 2.15 ซึ่งค่าต่าง ๆ ภายใน Confusion Matrix นั้นจะมีค่าอยู่หลากหลายค่าโดยสามารถอธิบายได้ดังต่อไปนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Positive	หมายถึง ข้อมูลที่เป็นกลุ่มบวก
Negative	หมายถึง ข้อมูลที่เป็นกลุ่มลบ
True Positive (TP)	คือ ผลรวมของการจำแนกข้อมูลที่ทำนายว่าเป็นกลุ่มบวก ซึ่งตรงกับผลลัพธ์จริงที่ระบุเป็นกลุ่มบวก
True Negative (TN)	คือ ผลรวมของการจำแนกข้อมูลที่ทำนายว่าเป็นกลุ่มลบ ซึ่งตรงกับผลลัพธ์จริงที่ระบุเป็นกลุ่มลบ
False Positive (FP)	คือ ผลรวมของการจำแนกข้อมูลที่ทำนายว่าเป็นกลุ่มบวก แต่ผลลัพธ์จริงระบุเป็นกลุ่มลบ
False Negative (FN)	คือ ผลรวมของการจำแนกข้อมูลที่ทำนายว่าเป็นกลุ่มลบ แต่ผลลัพธ์จริงระบุเป็นกลุ่มบวก

ซึ่งค่าที่ได้จาก Confusion Matrix จะถูกนำมาใช้ในการคำนวณหาค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกและค่าประสิทธิภาพโดยรวม[7] ซึ่งจะกล่าวในหัวข้อถัดไป

		Actual class	
		Positive	*Negative*
Predicted class	*Positive*	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

รูปที่ 2.15 ตารางเปรียบเทียบ (Confusion Matrix)

2.8.3 ค่าความถูกต้อง (Accuracy)

ค่าความถูกต้องนั้นสามารถคำนวณได้จากผลรวมของ true positive และ true negative หารด้วยผลรวมของ true positive, false positive, false negative และ true negative ดังสมการที่ (2.17)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.17}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.8.4 ค่าความแม่นยำ (Precision)

ค่าความแม่นยำ (Precision) คำนวณได้จาก true positive หารด้วยผลรวมของ true positive และ false positive แต่ในการคำนวณ precision ของคลาสลบ หรือคลาสอื่น ๆ ในกรณีที่มีหลายคลาสนั้น ค่า true positive จะหมายถึงคลาสที่สนใจและผลการทำนายที่ตรงกับคลาสที่สนใจ ส่วนค่า false positive จะหมายถึงผลรวมของผลการทำนายที่ไม่ตรงกับคลาสที่สนใจในแนวแถวดังสมการที่ (2.18)

$$Precision = \frac{TP}{TP + FP} \tag{2.18}$$

2.8.5 ค่าความระลึก (Recall)

ค่าความระลึก (Recall) คำนวณได้จาก true positive หารด้วยผลรวมของ true positive และ false negative แต่ในการคำนวณ precision ของคลาสลบ หรือคลาสอื่น ๆ ที่มีจำนวนคลาสมากกว่า 2 คลาส ค่า true positive จะหมายถึงคลาสที่สนใจและผลการทำนายที่ตรงกับคลาสที่สนใจ ส่วน false negative จะหมายถึงผลรวมของผลการทำนายที่ไม่ตรงกับคลาสที่สนใจในแนวหลักดังสมการที่ (2.19)

$$Recall = \frac{TP}{TP + FN} \tag{2.19}$$

2.8.6 ค่าเอฟเมเชอร์ (F-measure)

ค่าเอฟเมเชอร์ (F-measure) เป็นการวัดประสิทธิภาพ โดยใช้ค่า precision และ recall ในการคำนวณร่วมกัน สามารถคำนวณได้จากผลคูณระหว่าง precision และ recall เป็นจำนวน 2 เท่า หารด้วยผลรวมของ Precision และ Recall ตามสมการที่ (2.20)

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{2.20}$$

ตัวอย่าง การคำนวณค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกและค่าเอฟเมเชอร์จาก Confusion Matrix โดยกำหนดให้มีข้อมูล Confusion Matrix ขนาด 4x4 ซึ่งสามารถแสดงรูปแบบได้ดังตารางที่ 2.3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 2.3 ตัวอย่างข้อมูล Confusion Matrix ขนาด 4 X 4

		Actual class			
		Class 1	Class 2	Class 3	Class 4
Predicted class	Class 1	24	0	0	0
	Class 2	0	20	2	4
	Class 3	0	1	23	3
	Class 4	5	0	1	17

เราคำนวณหาค่า Accuracy, Precision, Recall และ F-measure ได้ดังนี้

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$= \frac{24 + 20 + 23 + 17}{24 + 5 + 20 + 1 + 2 + 23 + 1 + 4 + 3 + 17}$$
$$= \frac{84}{100} = 0.84$$

$$Precision = \frac{TP}{TP + FP} = \frac{17}{17 + 5 + 1} = \frac{17}{23} = \frac{84}{100} 0.7391$$

$$Recall = \frac{TP}{TP + FN} = \frac{17}{17 + 4 + 3} = \frac{17}{24} = \frac{84}{100} 0.7083$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} = \frac{2 \times 0.7391 \times 0.7083}{0.73391 + 0.7083}$$
$$= \frac{1.047}{1.442} = \frac{84}{100} 0.7260$$

2.9 งานวิจัยที่เกี่ยวข้อง

2.9.1 sEMG based classification of basic hand movements based on time-frequency features

สำหรับในงานวิจัย sEMG based classification of basic hand movements based on time-frequency features [8] ได้นำเสนอวิธีการจำแนกประเภทการเคลื่อนไหวของมือที่ใช้ในชีวิตประจำวัน โดยใช้ข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัสด้วยการใช้เทคนิควิธีการวิเคราะห์เฉพาะส่วนช่วงเวลา (Sliding Time-Window) โดยจะมีขนาดความยาวของข้อมูล 300

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

มิลลิวินาทีและอัตราการเลื่อมล้ำกันของข้อมูลทุก ๆ 30 มิลวินาทีร่วมกับการเทคนิคการแปลงคุณลักษณะข้อมูล เรียงตามลำดับขั้นตอนดังนี้

1) เทคนิควิธีอีเอ็มดี (Empirical Mode Decomposition : EMD) ซึ่งจะทำให้ได้สัญญาณที่เรียกว่า IMFs โดยใช้ IMFs ครั้งที่ 3
2) ใช้ 8 เทคนิคในการวิเคราะห์ข้อมูลเพื่อเลือกข้อมูลที่ต้องการ
3) นำข้อมูลมาลดจำนวนคุณลักษณะของข้อมูลด้วยเทคนิควิธี PCA เปรียบเทียบกับ Relief Algorithm

4. นำข้อมูลที่แปลงคุณลักษณะแล้ววัดประสิทธิภาพการจำแนกข้อมูลด้วย 2-fold cross-validation และการจำแนกแบบเชิงเส้น สมการที่ (2.21) โดยทำการจำแนกข้อมูลซ้ำ 5 ครั้งเพื่อหาค่าเฉลี่ยความถูกต้องของแต่ละข้อมูลผู้ทดลอง

$$i = \underset{i}{argmax} \{2lnP(\omega_i) - (x-\mu_i)^T C^{-1}(x-\mu_i)\} \qquad (2.21)$$

โดยที่ i คือ คลาส

$P(\omega_i)$ คือ ความน่าจะเป็นของคลาส i

C คือ ค่าความแปรปรวนร่วมของเมทริกซ์ (Covariance Matrix)

ผลลัพธ์ของงานวิจัยดังกล่าวแสดงให้เห็นว่า วิธีที่ลดจำนวนคุณลักษณะของข้อมูลด้วย PCA นั้นสามารถให้ผลการจำแนกได้ถูกต้องกว่าการใช้เทคนิค Relief Algorithm โดยผลลัพธ์ของจำแนกข้อมูลนั้นได้ค่าความถูกต้องของผู้ทดลองที่ 1 ที่ 95.16% ได้ค่าความถูกต้องของผู้ทดลองที่ 2 ที่ 90.93% ได้ค่าความถูกต้องของผู้ทดลองที่ 3 ที่ 86.74% ได้ค่าความถูกต้องของผู้ทดลองที่ 4 ที่ 88.06% ได้ค่าความถูกต้องของผู้ทดลองที่ 5 ที่ 88.86%

2.9.2 Towards sEMG Classification Based on Bayesian and *k*-NN to Control a Prosthetic Hand

Towards sEMG Classification Based on Bayesian and *k*-NN to Control a Prosthetic Hand [9] เป็นงานวิจัยที่นำเสนอการสกัดคุณลักษณะของข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อโดยใช้เทคนิควิธีการวิเคราะห์เฉพาะส่วนช่วงเวลา (Sliding Time-Window) ร่วมกับค่าเฉลี่ยกำลังสอง (Root Mean Square : RMS) ค่าความแปรปรวน (Variance : VAR) และความยาวของรูปคลื่น (Waveform Length : WL) ข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อที่ใช้ในงานวิจัยนี้มีจำนวน 4 ช่องสัญญาณ (Channel) มีการเก็บข้อมูลดังนี้ เก็บคลื่นสัญญาณไฟฟ้าจากอุปกรณ์ที่ชื่อว่า NI USB-6009 ข้อมูลที่เก็บมี Sampling rate ที่ 1000 เฮิร์ต (H_Z) โดยที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลในช่องสัญญาณที่ 1 จากกล้ามเนื้อ Flexor Pollicis Longus

ข้อมูลในช่องสัญญาณที่ 2 จากกล้ามเนื้อ Flexor Digitorum Superficialis

ข้อมูลในช่องสัญญาณที่ 3 จากกล้ามเนื้อ Flexor Carpi Radial and Ulnaris

ข้อมูลในช่องสัญญาณที่ 4 จากกล้ามเนื้อ Extensor Carpi radial and Ulnaris

โดยที่งานวิจัยดังกล่าวนั้นมีจุดมุ่งหมาย คือ การวิเคราะห์การเคลื่อนไหวของนิ้วมือ 15 รูปแบบ เช่น การผ่อนคลาย การเคลื่อนไหวของนิ้วทั้ง 5 นิ้ว โดยเคลื่อนไหวทีละนิ้ว ทีละ 2 นิ้วที่อยู่ติดกัน การแบมือโดยหัดฝ่ามือเข้าหาหรือออกจากลำตัว การกำมือ การจับของ การบีบนิ้ว เป็นต้น

ทั้งนี้ ในการทดลองจะทำการวิเคราะห์ข้อมูลโดยให้เลื่อมล้ำกันทุก ๆ 5 วินาที โดยมีอัตราการเลื่อมล้ำที่ 50% ของข้อมูลก่อนหน้า และทุกข้อมูลที่นำมาวิเคราะห์จะทำการสกัดคุณลักษณะข้อมูลด้วยค่าเฉลี่ยกำลังสอง ค่าความแปรปรวนและความยาวของรูปคลื่น สำหรับขั้นตอนการจำแนกข้อมูลได้ใช้ 10-fold cross-validation ร่วมกับตัวจำแนกแบบเบย์ (Bayesian Classifier) และขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว (k-NN) ซึ่งผลการจำแนกที่ได้นั้น งานวิจัยดังกล่าวได้รายงานว่าตัวจำแนกแบบเบย์ให้ผลความถูกต้องที่ 88% ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัวให้ผลความถูกต้องที่ 95%

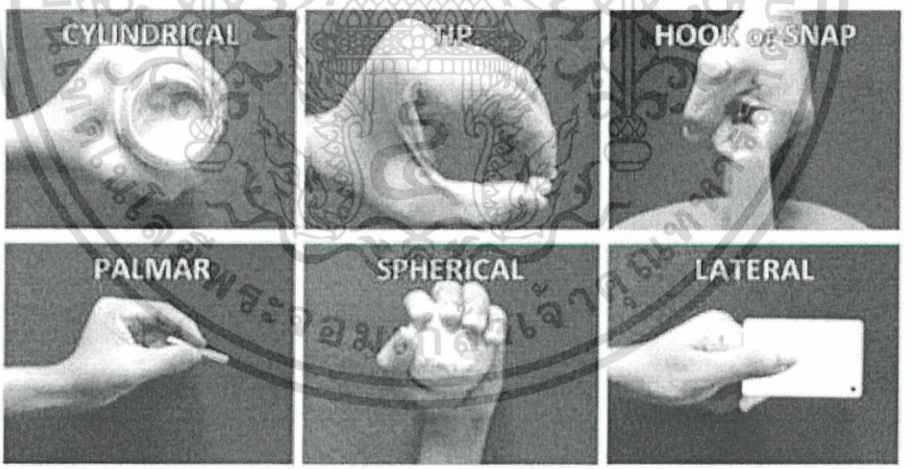
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3
การสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

งานวิจัยนี้เป็นการศึกษาเกี่ยวกับการลดจำนวนมิติของข้อมูลโดยมุ่งเน้นที่การลดจำนวนมิติข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ ซึ่งเป็นข้อมูลการเคลื่อนไหวของมือจำนวน 6 รูปแบบ โดยในส่วนนี้จะกล่าวถึงขั้นตอนการเตรียมข้อมูล ขั้นตอนการสกัดคุณลักษณะของข้อมูล และการวัดประสิทธิภาพการจำแนก

3.1 ขั้นตอนการเตรียมข้อมูล

ข้อมูลที่ใช้ในงานวิจัยนี้ ผู้วิจัยได้ใช้ข้อมูลจากฐานข้อมูลออนไลน์ UCI Machine Learning Repository [10] โดยชุดข้อมูลที่ใช้คือชุดข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อ [11] ซึ่งจัดเก็บข้อมูลขณะทำการเคลื่อนไหวมือเพื่อจับวัตถุ โดยมีลักษณะท่าทางการเคลื่นไหวทั้งสิ้น 6 รูปแบบ ได้แก่ การจับวัตถุทรงกลม (Spherical) การจับวัตถุขนาดเล็ก (Tip) การจับวัตถุที่บาง (Lateral) การจับวัตถุทรงกระบอก (Cylindrical) การจับวัตถุที่มีน้ำหนักมาก (Hook or Snap) และการจับวัตถุโดยหันวัตถุเข้าหาตัวเอง (Palmar) สามารถแสดงได้ดังรูปที่ 3.1



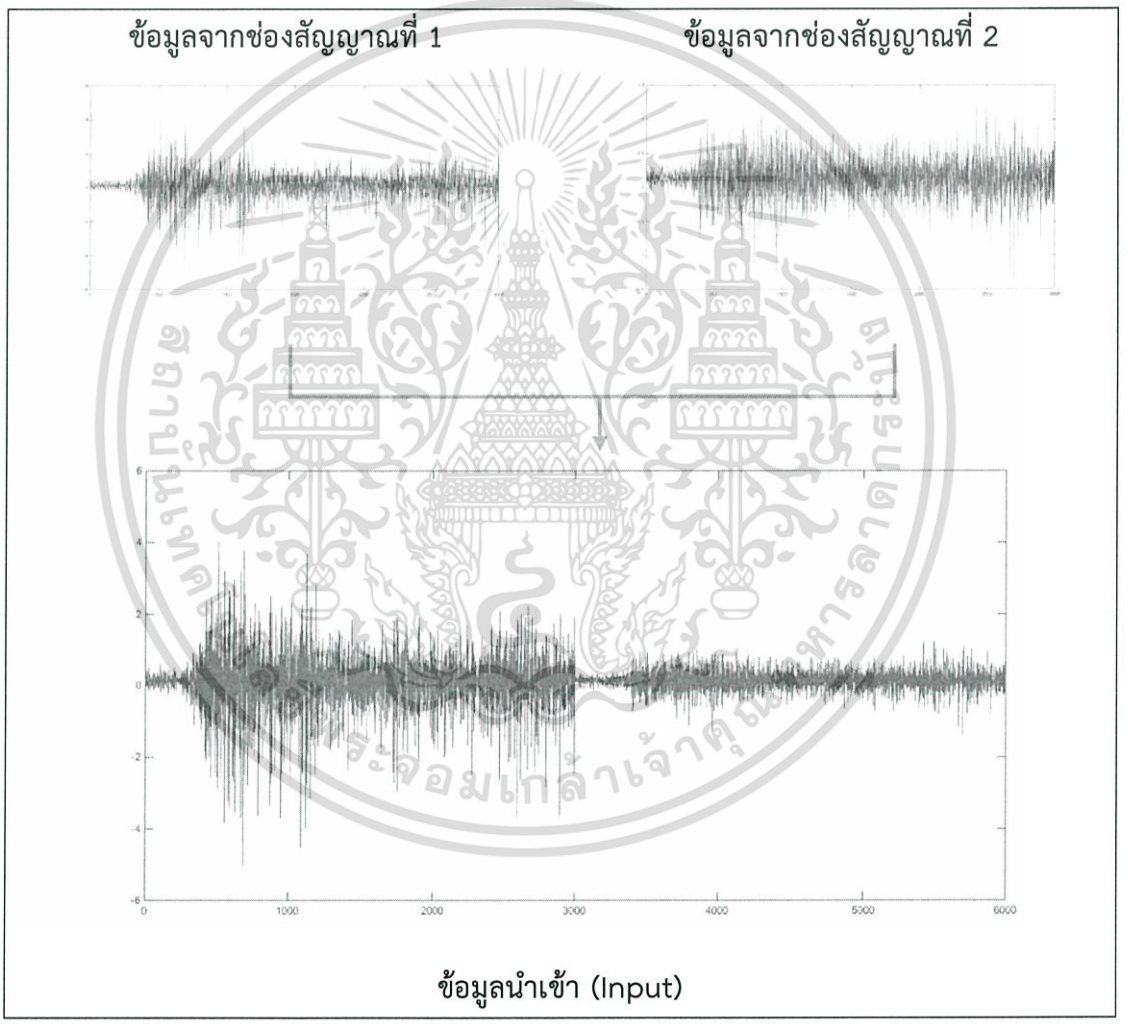
รูปที่ 3.1 การเคลื่อนไหวพื้นฐานของการจับใน 6 รูปแบบ

สำหรับการวัดและเก็บบันทึกข้อมูลสัญญาณไฟฟ้ากล้ามเนื้อนั้นจะใช้ Sampling rate ที่ 500 เฮิร์ต (H_Z) ซึ่งเป็นการวัดโดยการใช้ขั้วไฟฟ้าแบบผิวสัมผัสบริเวณกล้ามเนื้อส่วน Flexor Capri Ulnaris, Extensor Capri Radialis Longus, และ Extensor Capri Radialis Brevis ที่มีจำนวน 2 ช่องสัญญาณ โดยชุดข้อมูลนี้ถูกเก็บจากอาสาสมัครทั้งหมด 5 คน แบ่งเป็นผู้ชาย 2 คนและผู้หญิง 3 คน ซึ่งในแต่ละรูปแบบการเคลื่อนไหวของมือนั้น ผู้เข้ารับการทดสอบจะต้องทำท่าทางต่าง ๆ ท่าละ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

30 ครั้ง โดยในแต่ละครั้งจะมีระยะเวลาในการเคลื่อนไหว 6 วินาที (sec) ดังนั้นข้อมูลที่ได้ในการเคลื่อนไหวมือ 1 ครั้งจะมีจำนวนคุณลักษณะของข้อมูลอยู่ที่ 3,000 คุณลักษณะ (500 เฮิร์ต x 6 วินาที) ต่อ 1 ช่องสัญญาณ

ข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อของผู้เข้ารับการทดลองแต่ละคนนั้น มีจำนวนช่องสัญญาณอยู่ 2 ช่องสัญญาณ โดยข้อมูลในช่องสัญญาณที่ 2 จะถูกนำมารวมกับข้อมูลในช่องสัญญาณที่ 1 ในลักษณะการต่อท้ายข้อมูลให้เป็นข้อมูลเดียวกัน ซึ่งจะทำให้ได้ข้อมูลใหม่ที่มีจำนวนคุณลักษณะทั้งหมด 6,000 คุณลักษณะต่อการเคลื่อนไหวมือใน 1 ครั้ง โดยที่วิธีนี้ถือเป็นเพียงการเตรียมข้อมูลแบบปกติเท่านั้น ดังแสดงในรูปที่ 3.2



รูปที่ 3.2 ตัวอย่างคลื่นสัญญาณไฟฟ้ากล้ามเนื้อก่อนและหลังรวมช่องสัญญาณ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 วิธีการแปลงข้อมูลแบบเวกเตอร์เป็นข้อมูลแบบเมทริกซ์

ในวิทยานิพนธ์นี้มุ่งเน้นไปที่การลดจำนวนมิติของข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อโดยจะทำการแปลงข้อมูลซึ่งอยู่ในรูปแบบเวกเตอร์ให้อยู่ในรูปแบบเมทริกซ์ โดยที่เราสามารถคำนวณจำนวนแถวของเมทริกซ์ได้จากสมการที่ (3.1)

$$\text{จำนวนแถวของเมทริกซ์} = \left\lceil \frac{\text{ความยาวของเวกเตอร์}}{\text{จำนวนคอลัมน์ที่ต้องการ}} \right\rceil \tag{3.1}$$

โดยที่

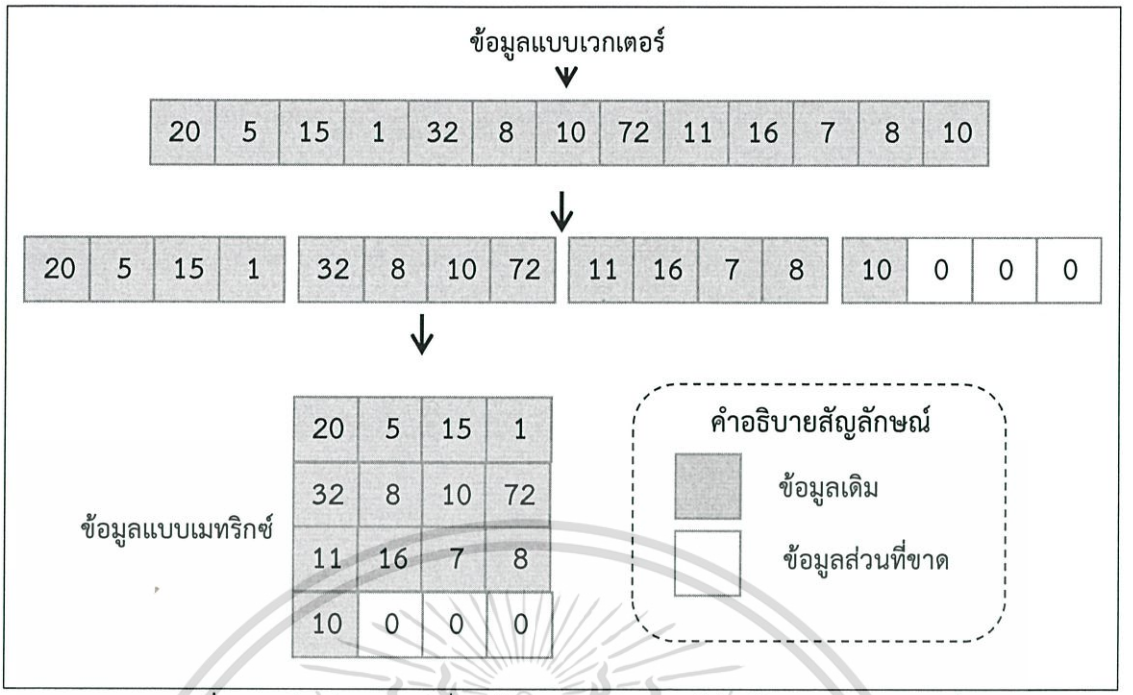
$\lceil \ \rceil$ คือ ฟังก์ชันสำหรับปัดเศษขึ้น

ตัวอย่าง สมมุติว่าข้อมูล 1 แถวใน 1 ช่องสัญญาณมี 13 คุณลักษณะ ถ้าต้องการเปลี่ยนข้อมูลให้เป็นแบบเมทริกซ์ขนาด 4 คอลัมน์ จะสามารถคำนวณหาจำนวนแถวของเมทริกซ์ได้ดังนี้

$$\text{จำนวนแถวของเมทริกซ์} = \left\lceil \frac{\text{ความยาวของเวกเตอร์}}{\text{จำนวนคอลัมน์ที่ต้องการ}} \right\rceil = \left\lceil \frac{13}{4} \right\rceil = \lceil 3.25 \rceil = 4$$

หมายความว่า ถ้ามีความยาวของเวกเตอร์ (ซึ่งในที่นี้คือจำนวนคุณลักษณะของข้อมูล) เท่ากับ 13 คุณลักษณะ และกำหนดจำนวนคอลัมน์ของเมทริกซ์ให้มีค่าเป็น 4 คอลัมน์ เราจะสามารถหาจำนวนแถวของเมทริกซ์ได้เท่ากับ 3.25 แถว เมื่อทำการปัดเศษขึ้นจะได้เป็น 4 แถว เพราะหากทำการปัดเศษลงจะทำให้เกิดการสูญหายของข้อมูลหลังจากที่ทำการแปลงข้อมูลให้อยู่ในรูปแบบเมทริกซ์ เนื่องจากข้อมูลในแถวสุดท้ายที่มีข้อมูลไม่เต็มจำนวนคอลัมน์จะไม่ถูกนำมารวมกับข้อมูลในรูปแบบเมทริกซ์ จึงจำเป็นต้องปัดเศษขึ้นเพื่อรองรับข้อมูลในแถวสุดท้าย ซึ่งข้อมูลในแถวสุดท้ายส่วนที่ไม่มีข้อมูลจะถูกแทนค่าข้อมูลด้วยเลข 0 โดยสามารถแสดงตัวอย่างได้ดังรูปที่ 3.3 อย่างไรก็ตามการแปลงข้อมูลแบบเวกเตอร์เป็นข้อมูลแบบเมทริกซ์นี้จะกระทำกับข้อมูลเวกเตอร์ทีละแถวซึ่งแยกช่องสัญญาณกัน โดยเมื่อได้ข้อมูลที่อยู่ในรูปแบบเมทริกซ์แล้วจะทำการสกัดคุณลักษณะด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน ซึ่งจะกล่าวในหัวข้อถัดไป

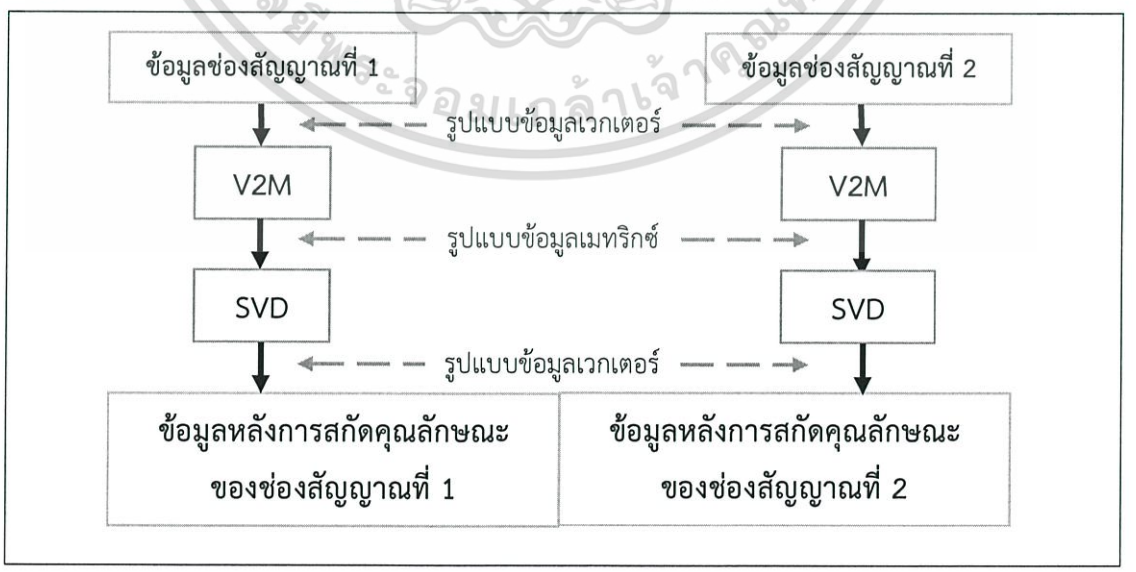
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.3 ตัวอย่างการเปลี่ยนข้อมูลแบบเวกเตอร์เป็นข้อมูลแบบเมทริกซ์

3.3 การสกัดคุณลักษณะของข้อมูลด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน

การสกัดคุณลักษณะของข้อมูลที่นำเสนอในงานวิจัยนี้ได้ประยุกต์ใช้หลักการแปลงข้อมูลที่อยู่ในรูปแบบเวกเตอร์ให้เป็นข้อมูลที่อยู่ในรูปแบบเมทริกซ์โดยจะกระทำกับข้อมูลในแต่ละช่องสัญญาณ ซึ่งเมื่อได้ข้อมูลที่เป็นรูปแบบเมทริกซ์ในแต่ละช่องสัญญาณแล้ว ก็จะใช้วิธีการวิเคราะห์องค์ประกอบของข้อมูลด้วยวิธีซิงกูร์แวลูดีคอมโพซิชัน จากนั้นจึงนำข้อมูลหลังการสกัดคุณลักษณะแล้วของแต่ละช่องสัญญาณมาเชื่อมต่อกัน ซึ่งสามารถแสดงขั้นตอนการสกัดคุณลักษณะของข้อมูลได้ดังรูปที่ 3.4 โดยที่ V2M คือการแปลงเวกเตอร์เป็นเมทริกซ์



รูปที่ 3.4 ขั้นตอนการสกัดคุณลักษณะด้วยวิธีซิงกูลาร์แวลูดีคอมโพซิชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการเก็บข้อมูลสัญญาณที่ได้จากเซ็นเซอร์นั้น หากมีจำนวนของเซ็นเซอร์ที่เกี่ยวข้องมากกว่า 1 ตัว จะสามารถทำการเก็บชุดข้อมูลสัญญาณเซ็นเซอร์แยกออกเป็นของเซ็นเซอร์แต่ละตัวได้ ซึ่งโดยปกติแล้วเซ็นเซอร์ 1 ตัว ก็จะให้สัญญาณเป็น 1 ช่องสัญญาณ ทั้งนี้ ชุดข้อมูลที่นำมาใช้ในการวิจัยนี้จะมีจำนวนของช่องสัญญาณอยู่ 2 ช่องสัญญาณ ดังในรูปที่ 3.4 ซึ่งเมื่อนำชุดข้อมูลดังกล่าวมาเข้าสู่กระบวนการสกัดข้อมูล ผู้วิจัยจะใช้วิธีให้ชุดข้อมูลของแต่ละช่องสัญญาณผ่านขั้นตอนวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันก่อน แล้วจึงนำชุดข้อมูลที่ได้จากการผ่านขั้นตอนวิธีดังกล่าวมารวมกันโดยนำชุดข้อมูลช่องสัญญาณที่ 2 มาต่อท้ายชุดข้อมูลช่องสัญญาณที่ 1 โดยมีรายละเอียดขั้นตอนวิธีดังต่อไปนี้

ขั้นตอนการทำงานของการสกัดคุณลักษณะข้อมูลโดยการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน (Vector to Matrix and SVD algorithm)

ข้อมูลนำเข้า :

ข้อมูลจากช่องสัญญาณที่ 1 และช่องสัญญาณที่ 2

กำหนดค่า K เป็นจำนวนคุณลักษณะที่ต้องการ

โดยที่ค่าต่ำสุดของ K จะต้องมีค่าไม่น้อยกว่าจำนวนของช่องสัญญาณทั้งหมด และหากต้องการค่า K ที่สูงกว่าค่าต่ำสุด ให้กำหนดค่าในลักษณะเพิ่มขึ้นทีละ 1 เท่าของจำนวนช่องสัญญาณทั้งหมด

ขั้นตอนวิธี :

1. คำนวณหาค่าจำนวนคอลัมน์ในแต่ละช่องสัญญาณจาก

$$\text{จำนวนคอลัมน์ในแต่ละช่องสัญญาณ} = \frac{\text{จำนวนคุณลักษณะที่ต้องการ}(K)}{\text{จำนวนช่องสัญญาณทั้งหมด}} \qquad (3.2)$$

2. คำนวณหาแถวของเมทริกซ์ในแต่ละช่องสัญญาณจาก

$$\text{แถวของเมทริกซ์ใหม่} = \left\lceil \frac{\text{ความยาวของเวกเตอร์ในช่องสัญญาณนั้น}}{\text{จำนวนคอลัมน์ในแต่ละช่องสัญญาณ}} \right\rceil \qquad (3.3)$$

กรณีที่คำนวณแล้วมีค่าเป็นเศษ ค่าเศษจะถูกปัดขึ้นเป็นจำนวนเต็ม

3. ข้อมูลเวกเตอร์ในช่องสัญญาณที่ 1 จะถูกตัดแบ่งออกเป็นส่วน ๆ แล้วนำมาเรียงต่อกันเป็นรูปแบบเมทริกซ์ โดยที่ส่วนแรกของข้อมูลที่ตัดแบ่งจะกลายเป็นแถวที่ 1 ของเมทริกซ์ ส่วนที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สองของข้อมูลที่ตัดแบ่งจะกลายเป็นแถวที่ 2 ซึ่งตำแหน่งของคอลัมน์แรกของส่วนที่ 1 และคอลัมน์แรกของส่วนที่ 2 จะต้องตรงกัน โดยให้วนซ้ำจนครบตามจำนวนแถวของเมทริกซ์ใหม่ กรณีในแถวสุดท้ายของเมทริกซ์ หากมีข้อมูลไม่ครบก็ให้เติมด้วยค่า 0 จนจนครบหลักในแถวสุดท้าย

4. ข้อมูลเวกเตอร์ในช่องสัญญาณที่ 2 จะถูกทำเหมือนกับในขั้นตอนที่ 3
5. นำข้อมูลเมทริกซ์ที่ได้จากขั้นตอนที่ 3 และ 4 มาทำการสกัดข้อมูลโดยใช้ซิงกูลาร์แวลูดีคอมโพซิชัน (SVD) ซึ่งค่าที่ได้คือ เมทริกซ์ยูนิแทรี เมทริกซ์เฉียง และเมทริกซ์สลับเปลี่ยน
6. นำค่าในแนวทแยงของเมทริกซ์เฉียงของข้อมูลแต่ละช่องสัญญาณ มาเรียงใหม่ให้อยู่ในรูปแบบเวกเตอร์ ซึ่งจะได้ข้อมูลเวกเตอร์ของช่องสัญญาณที่ 1 และข้อมูลเวกเตอร์ของช่องสัญญาณที่ 2
7. นำเวกเตอร์ของช่องสัญญาณที่ 1 และ 2 มารวมกัน โดยนำข้อมูลเวกเตอร์ของช่องสัญญาณที่ 2 มาต่อท้ายของข้อมูลเวกเตอร์ของช่องสัญญาณที่ 1

ข้อมูลส่งออก :

การสกัดคุณลักษณะของข้อมูลนี้ ข้อมูลนำเข้า 1 ตัว เมื่อสกัดคุณลักษณะของข้อมูลแล้วจะได้ชุดข้อมูลใหม่จำนวน 1 ชุด โดยจะเป็นข้อมูลที่จะอยู่ในมิติใหม่ซึ่งมีจำนวนคุณลักษณะน้อยกว่าชุดข้อมูลเดิมตามจำนวนที่ได้กำหนดไว้ ซึ่งจำนวนคุณลักษณะที่เป็นไปได้จะมีจำนวนไม่เกินรากที่ 2 ของจำนวนคุณลักษณะเดิม

ในขั้นตอนของการสกัดคุณลักษณะของข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัสได้กำหนดให้ข้อมูลนำเข้า ได้แก่ ข้อมูลจากช่องสัญญาณที่ 1 ข้อมูลจากช่องสัญญาณที่ 2 และค่าจำนวนคุณลักษณะที่ต้องการหรือค่า K จากนั้นการเริ่มต้นสกัดคุณลักษณะของข้อมูล จะเริ่มจากการคำนวณหาค่าจำนวนคอลัมน์ในแต่ละช่องสัญญาณ โดยหาจากจำนวนคุณลักษณะที่ต้องการหารด้วยจำนวนช่องสัญญาณทั้งหมด จากนั้นคำนวณหาจำนวนแถวของเมทริกซ์ใหม่โดยคำนวณจากความยาวเวกเตอร์ของช่องสัญญาณหนึ่งหารด้วยจำนวนคอลัมน์ในแต่ละช่องสัญญาณ จากนั้นทำการตัดแบ่งข้อมูลเวกเตอร์ของช่องสัญญาณที่ 1 โดยให้มีขนาดเท่ากับจำนวนคอลัมน์ในแต่ละช่องสัญญาณ แล้วนำมาเรียงต่อกันให้อยู่ในรูปแบบเมทริกซ์ ส่วนข้อมูลในช่องสัญญาณที่ 2 ก็ทำการตัดแบ่งแล้วจัดเรียงให้อยู่ในรูปเมทริกซ์เช่นเดียวกัน จากนั้นนำเมทริกซ์ที่ได้ไปสกัดคุณลักษณะของข้อมูลทีละเมทริกซ์ของแต่ละช่องสัญญาณ โดยผลลัพธ์ที่จะได้จะเป็นเมทริกซ์ 3 รูปแบบ คือ เมทริกซ์ยูนิแทรี เมทริกซ์เฉียง และเมทริกซ์สลับเปลี่ยน ให้เลือกใช้เฉพาะในเมทริกซ์เฉียง โดยนำข้อมูลที่อยู่ในแนวทแยงซึ่งเป็นค่าที่ไม่ใช่ศูนย์มาจัดให้อยู่ในรูปแบบของเวกเตอร์ ซึ่งข้อมูลที่ได้หลังการสกัดคุณลักษณะจะมีข้อมูลแบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เวกเตอร์ 2 ข้อมูล คือ ข้อมูลเวกเตอร์จากช่องสัญญาณที่ 1 และข้อมูลเวกเตอร์จากช่องสัญญาณที่ 2 ขั้นตอนสุดท้ายให้นำข้อมูลทั้ง 2 เวกเตอร์มาเรียงต่อกันโดยนำข้อมูลเวกเตอร์จากช่องสัญญาณที่ 2 มาต่อท้ายข้อมูลเวกเตอร์จากช่องสัญญาณที่ 1 เป็นข้อมูลเวกเตอร์ตัวใหม่ จึงเป็นอันแล้วเสร็จขั้นตอนการสกัดคุณลักษณะข้อมูล 1 ตัว โดยการสกัดคุณลักษณะข้อมูลจะทำจนครบจำนวนของข้อมูลที่มีในชุดข้อมูลจึงถือว่าสิ้นสุดการสกัดคุณลักษณะ จากนั้นจะทำการนำข้อมูลที่ได้ไปจำแนกและวัดประสิทธิภาพในขั้นตอนต่อไป

3.4 ขั้นตอนการวัดประสิทธิภาพ

ในงานวิจัยนี้ได้ใช้ข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัสจากฐานข้อมูลออนไลน์ที่มีชื่อว่า UCI มาใช้ในการทดสอบเพื่อวัดประสิทธิภาพของข้อมูล ซึ่งผู้วิจัยได้นำข้อมูลดิบมาผ่านขั้นตอนของการเตรียมข้อมูลและการสกัดคุณลักษณะของข้อมูลโดยใช้โปรแกรม MATLAB เวอร์ชัน R2014b เมื่อได้ข้อมูลที่ผ่านการเตรียมและการสกัดคุณลักษณะของข้อมูลแล้วนำมาทำการวัดประสิทธิภาพด้วยการวัดค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกและค่าเอฟเมเชอร์ ด้วยขั้นตอนวิธีการจำแนกต่าง ๆ ด้วยโปรแกรม WEKA เวอชัน 3.6

สำหรับการวัดประสิทธิภาพของการทดลอง ข้อมูลที่ใช้ในการทดลองเป็นข้อมูลที่ได้มาจากผู้ทดลอง 5 คน ซึ่งการวัดประสิทธิภาพของข้อมูลจะแบ่งได้การทดลองดังนี้ คือ การทดลองกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและการทดลองกับข้อมูลหลังจากสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันใช้การวัดค่าความถูกต้องเพื่อดูประสิทธิภาพโดยรวมของการจำแนกข้อมูล สำหรับค่าความแม่นยำ ค่าความระลึกและค่าเอฟเมเชอร์ใช้วัดเพื่อดูประสิทธิภาพของวิธีการจำแนกข้อมูลในแต่ละคลาส เนื่องจากงานวิจัยนี้ได้เสนอรูปแบบวิธีการสกัดคุณลักษณะ ดังนั้นได้ใช้วิธีสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ในการเปรียบเทียบผลลัพธ์ของการจำแนก โดยใช้รูปแบบการสกัดคุณลักษณะทีละช่องสัญญาณเพื่อให้สอดคล้องกับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน แต่ในการวัดประสิทธิภาพจะวัดเฉพาะค่าความถูกต้องเท่านั้น สำหรับการทดลองแต่ละวิธีจะใช้เทคนิคการตรวจสอบไขว้แบบ 2 ส่วน (2-fold cross validation) ซึ่งเป็นเทคนิคการแบ่งข้อมูลออกเป็นชุด ๆ ดังได้กล่าวไว้ในหัวข้อที่ 2.8.1 ร่วมกับวิธีการจำแนกประเภทอีก 5 วิธี คือ ขั้นตอนวิธีแบบเบย์อย่างง่าย ขั้นตอนวิธีโครงข่ายประสาทเทียมเรดียลเบสิสฟังก์ชัน ขั้นตอนวิธีซีโฟพอยต์ไฟว์ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว และอัลกอริทึมเอสเอ็มโอ นอกจากนี้ยังวัดประสิทธิภาพจากข้อมูลผู้ทดลองทั้ง 5 คนรวมกันอีกด้วย ในการกำหนดค่าพารามิเตอร์ต่าง ๆ และผลการทดลองจะกล่าวไว้ในหัวข้อถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4
การทดลองและผลการทดลอง

ในบทนี้จะกล่าวถึงการทดลองและผลการทดลองในรูปแบบต่าง ๆ ซึ่งได้ทำวัดประสิทธิภาพและการนำผลลัพธ์มาเปรียบเทียบกัน มีดังนี้ การทดลองข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ การทดลองและเปรียบเทียบผลการทดลองระหว่างข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับการทดลองข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ การทดลองและเปรียบเทียบผลการทดลองระหว่างข้อมูลที่สกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ที่ละช่องสัญญาณ โดยจะใช้ชุดข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัส ดังที่ได้กล่าวมาแล้วในบทที่ผ่านมา

4.1 การทดลอง

4.1.1 โปรแกรมที่ใช้ในงานวิจัยและการกำหนดค่าพารามิเตอร์

ในงานวิจัยนี้ได้ใช้โปรแกรม Matlab ในส่วนของขั้นตอนการเตรียมข้อมูลและการสกัดคุณลักษณะกับข้อมูลคลื่นสัญญาณไฟฟ้า ซึ่งข้อมูลที่ได้จะถูกนำมาทำการจำแนกประเภทด้วยวิธีต่าง ๆ ด้วยการใช้โปรแกรม weka ซึ่งในการทดลองแต่ละการทดลองที่ได้กล่าวมาข้างต้นจะใช้กับชุดข้อมูลผู้ทดลองจำนวน 5 คนแบ่งเป็นการทดลองแต่ละคนและข้อมูลรวมกันทั้งหมด โดยที่มีค่าจำนวนคุณลักษณะของข้อมูลและพารามิเตอร์สำหรับใช้ในขั้นตอนวิธีการจำแนกต่าง ๆ ได้ดังนี้

1) การกำหนดค่าจำนวนคุณลักษณะของข้อมูล

ในการเปรียบเทียบข้อมูลก่อนและหลังการสกัดคุณลักษณะด้วยแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันในแต่ละช่องสัญญาณ กำหนดให้ข้อมูลหลังการสกัดคุณลักษณะ มีจำนวนคุณลักษณะ 10 คุณลักษณะถึง 110 คุณลักษณะ โดยมีการเพิ่มขึ้นทีละ 10 คุณลักษณะเปรียบเทียบกับข้อมูลก่อนการสกัดคุณลักษณะที่มีจำนวนคุณลักษณะ 6,000 คุณลักษณะ ในการเปรียบเทียบข้อมูลหลังการสกัดคุณลักษณะด้วยแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันในแต่ละช่องสัญญาณกับข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ กำหนดจำนวนคุณลักษณะหลังการสกัดให้มีจำนวนคุณลักษณะเป็น 20 40 60 80 และ 100 คุณลักษณะ

2) การกำหนดค่าพารามิเตอร์ในขั้นตอนวิธีการจำแนกประเภท

การกำหนดค่าพารามิเตอร์ในขั้นตอนวิธีแบบเบย์อย่างง่าย ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และขั้นตอนวิธีซีโฟพอยต์ไฟว์ ใช้ค่าพารามิเตอร์ที่เป็นค่าพารามิเตอร์พื้นฐานของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม การกำหนดค่าพารามิเตอร์ในขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัว ใช้ค่า *k* เป็น 3 และการวัดระยะทางด้วยวิธียูคลิด การกำหนดค่าพารามิเตอร์ในขั้นตอนวิธีอัลกอริทึมเอสเอ็มโอ ใช้ค่าความซับซ้อนของข้อมูล (C) ที่ 1000 และค่า Kernel เป็น RBF Kernel ที่ใช้ค่าพารามิเตอร์ที่เป็นค่าพื้นฐานของโปรแกรม

4.2 ผลการทดลอง

4.2.1 การทดลองกับข้อมูลผู้ทดลองที่ 1

4.2.1.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกค่าความถูกต้องของข้อมูลผู้ทดลองที่ 1 ที่ผ่านการเตรียมข้อมูลแบบปกติมีดังต่อไปนี้ เริ่มจากการจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้อง 75.55% การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้อง 16.11% การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้อง 52.77% การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้อง 30.00% และการจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้อง 48.88% ค่าความถูกต้องของการจำแนกข้อมูลแต่ละวิธีสามารถแสดงได้ดังตารางที่ 4.1

ตารางที่ 4.1 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 1

	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
6000	75.55	16.11	52.77	30.00	48.88

ACC. = Accuracy.

4.2.1.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 1 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันแล้วมีดังต่อไปนี้ การจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้องสูงสุด 88.88% ที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้องสูงสุด 94.44% ที่ 20 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้องสูงสุด 91.11% ที่

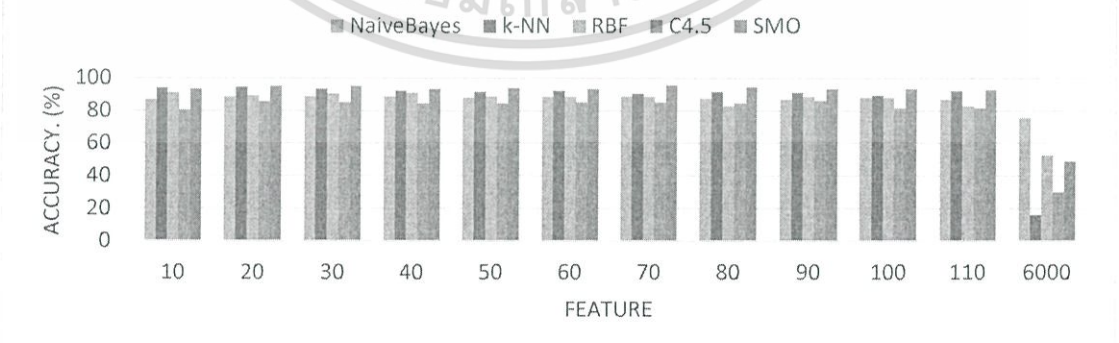
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

10 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้องสูงสุด 86.11% ที่ 90 คุณลักษณะ การจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องสูงสุด 95.55% ที่จำนวน 70 คุณลักษณะ ผลการจำแนกกับข้อมูลผู้ทดลองที่ 1 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันแสดงได้ดังตารางที่ 4.2 ซึ่งเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติได้ดังรูปที่ 4.1

ตารางที่ 4.2 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 1

V2M-SVD	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
10	86.66	93.88	91.11	80.55	93.33
20	88.33	94.44	89.44	85.55	95.00
30	88.88	93.33	90.55	85.00	95.00
40	88.88	92.22	91.11	84.44	93.33
50	87.77	91.66	88.88	84.44	93.88
60	88.33	92.22	88.33	85.00	93.33
70	88.88	90.55	88.33	85.00	95.55
80	87.22	91.66	82.77	84.44	94.44
90	86.66	91.11	88.33	86.11	93.33
100	87.77	89.44	87.77	81.66	93.33
110	86.66	92.22	82.77	81.66	92.77

ACC. = Accuracy.



รูปที่ 4.1 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดผลลัพธ์การจำแนกข้อมูลเชิงลึกสามารถแสดงข้อมูลได้ดังตารางที่ 4.3

ตารางที่ 4.3 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 1

วิธีการจำแนกข้อมูล	คลาส	ข้อมูลก่อนสกัดคุณลักษณะ			ข้อมูลหลังจากสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน					
		6000 คุณลักษณะ			20 คุณลักษณะ			30 คุณลักษณะ		
		P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)
Naïve Bayes	Cylindrical	100	46.7	63.6	93.1	90.0	91.5	90.0	90.0	90.0
	Hook	89.3	83.3	86.2	85.3	96.7	90.6	84.8	93.3	88.9
	Lateral	100	60.0	75.0	91.7	73.3	81.5	92.3	80.0	85.7
	Palm	71.9	76.7	74.2	79.4	90.0	84.4	81.8	90.0	85.7
	Spherical	58.3	93.3	71.8	96.4	90.0	93.1	96.4	90.0	93.1
	Tip	70.0	93.3	80.0	87.1	90.0	88.5	90.0	90.0	90.0
k-NN	Cylindrical	0.0	0.0	0.0	90.3	93.3	91.8	87.5	93.3	90.3
	Hook	0.0	0.0	0.0	96.7	96.7	96.7	93.5	96.7	95.1
	Lateral	16.6	96.7	28.3	96.7	96.7	96.7	96.7	96.7	96.7
	Palm	0.0	0.0	0.0	93.1	90.0	91.5	93.1	90.0	91.5
	Spherical	0.0	0.0	0.0	96.6	93.3	94.9	96.3	86.7	91.2
	Tip	0.0	0.0	0.0	93.5	96.7	95.1	93.5	96.7	95.1
RBF	Cylindrical	24.1	23.3	23.7	90.0	90.0	90.0	87.5	93.3	90.3
	Hook	86.2	83.3	84.7	96.7	96.7	96.7	93.5	96.7	95.1
	Lateral	83.3	16.7	27.8	92.6	83.3	87.7	92.3	80.0	85.7
	Palm	51.2	70.0	59.2	77.1	90.0	83.1	84.4	90.0	87.1
	Spherical	50.9	93.3	65.9	90.3	93.3	91.8	100	90.0	94.7
	Tip	45.0	30.0	36.0	92.6	83.3	87.7	87.5	93.3	90.3
C4.5	Cylindrical	28.6	33.3	30.8	86.7	86.7	86.7	82.8	80.0	81.4
	Hook	56.3	30.0	39.1	83.3	83.3	83.3	84.8	93.3	88.9
	Lateral	34.4	36.7	35.5	80.0	80.0	80.0	82.1	76.7	79.3
	Palm	18.5	16.7	17.5	83.3	83.3	83.3	84.4	90.0	87.1
	Spherical	34.5	33.3	33.9	87.1	90.0	88.5	89.7	86.7	88.1
	Tip	22.0	30.0	25.4	93.1	90.0	91.5	86.2	83.3	84.7
SMO	Cylindrical	30.9	100	47.2	90.3	93.3	91.8	90.3	93.3	91.8
	Hook	0.0	0.0	0.0	93.5	96.7	95.1	96.7	96.7	96.7
	Lateral	65.6	70.0	67.7	96.7	96.7	96.7	96.7	96.7	96.7
	Palm	75.0	60.0	66.7	93.3	93.3	93.3	96.4	90.0	93.1
	Spherical	0.0	0.0	0.0	100	93.3	96.6	96.6	93.3	94.9
	Tip	70.4	63.3	66.7	96.7	96.7	96.7	93.8	100	96.8

P.=Precision. R.=Recall. F.=F-Measure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 1 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้ว่ามีประสิทธิภาพในการจำแนกข้อมูลที่ต่ำ โดยจำนวนคุณลักษณะที่ 20 ได้ค่าความถูกต้องสูงสุดเพียง 35.56% ด้วยขั้นตอนการจำแนกแบบเบย์อย่างง่าย ในขณะข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้นได้ค่าความถูกต้องสูงสุดถึง 88.89% ด้วยอัลกอริทึมเอสเอ็มโอ สำหรับผลลัพธ์ของการจำแนกข้อมูลระหว่างการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์และวิธีซิงกูลาร์แวลูดีคอมโพซิชันสามารถแสดงได้ดังตารางที่ 4.4 และตารางที่ 4.5 นอกจากนี้สามารถดูรูปแสดงผลลัพธ์เปรียบเทียบการสกัดคุณลักษณะและการจำแนกได้ดังรูปที่ 4.2

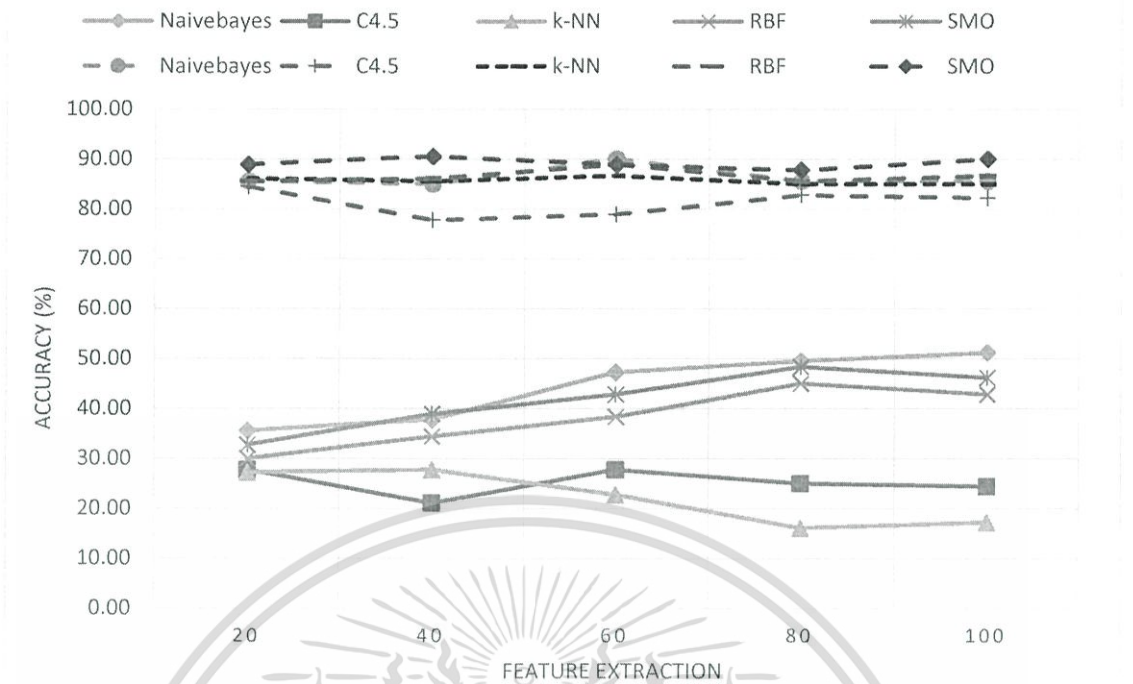
ตารางที่ 4.4 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 1 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	35.56	27.78	27.22	30.00	32.78
40	37.78	21.11	27.78	34.44	38.89
60	47.22	27.78	22.78	38.33	42.78
80	49.44	25.00	16.11	45.00	48.33
100	51.11	24.44	17.22	42.78	46.11

ตารางที่ 4.5 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 1 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	85.56	84.44	86.11	85.56	88.89
40	85.00	77.78	85.56	86.11	90.56
60	90.00	78.89	86.67	88.89	88.89
80	85.56	82.78	85.00	85.56	87.78
100	85.56	82.22	85.00	86.67	90.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 1

จากรูปที่ 4.2 วิธีการจำแนกข้อมูลที่เป็นเส้นทึบหมายถึงข้อมูลที่ใช้วิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ซึ่งวิธีการจำแนกกับข้อมูลชุดนี้ให้ผลการจำแนกที่ต่ำมาก โดยมีเพียงขั้นตอนวิธีแบบเบย์อย่างง่าย ที่จำนวนคุณลักษณะที่ 100 คุณลักษณะเท่านั้นที่ให้ผลการจำแนกได้เกิน 50% สำหรับวิธีการจำแนกข้อมูลที่เป็นเส้นปะหมายถึงข้อมูลที่ใช้วิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับการซิงกูลาร์แวลูดีคอมโพซิชัน วิธีการจำแนกกับข้อมูลชุดนี้ที่ให้ผลได้ต่ำที่สุดคือ ขั้นตอนวิธีซีโฟพอยต์ไฟว์ ที่จำนวนคุณลักษณะ 40 และ 60 คุณลักษณะให้ผลการจำแนกไม่ถึง 80% ซึ่งวิธีการจำแนกอื่น ๆ ให้ค่าความถูกต้องที่เกินกว่า 85% ในทุกจำนวนคุณลักษณะที่ได้ทำการทดสอบ และวิธีการจำแนกที่ให้ค่าได้สูงที่สุด คือ อัลกอริทึมเอสเอ็มโอ ซึ่งสามารถจำแนกข้อมูลให้ค่าความถูกต้องสูงถึง 90.56%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 ผลการทดลองกับข้อมูลผู้ทดลองที่ 2

4.2.2.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกค่าความถูกต้องของข้อมูลผู้ทดลองที่ 2 ที่ผ่านการเตรียมข้อมูลแบบปกติมีดังต่อไปนี้ เริ่มจากการจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้อง 75.00% การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้อง 16.66% การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้อง 52.77% การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้อง 31.11% และการจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้อง 41.11% ค่าความถูกต้องของการจำแนกข้อมูลแต่ละวิธีสามารถแสดงได้ดังตารางที่ 4.6

ตารางที่ 4.6 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 2

	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
6000	75.00	16.66	52.77	31.11	41.11

ACC. = Accuracy.

4.2.2.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 2 ที่ได้รับการแปลงคุณลักษณะของข้อมูลแล้วมีดังต่อไปนี้ การจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้องสูงสุด 92.77% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 60 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้องสูงสุดได้ 93.88% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 10 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้องสูงสุดได้ 93.33% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้องสูงสุดได้ 93.88% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 50 คุณลักษณะ การจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องสูงสุดได้ 98.33% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 10 คุณลักษณะ ผลการจำแนกกับข้อมูลผู้ทดลองที่ 2 ที่ได้รับการแปลงคุณลักษณะของ

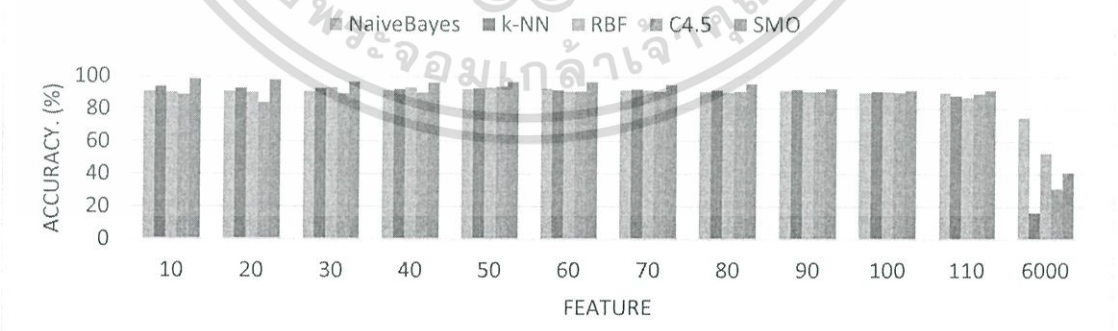
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ข้อมูลแล้วแสดงได้ดังตารางที่ 4.7 ซึ่งเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติได้ดังรูปที่ 4.3

ตารางที่ 4.7 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 2

V2M-SVD	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
10	91.11	93.88	90.55	88.88	98.33
20	91.11	92.77	90.00	83.88	97.77
30	91.11	92.77	93.33	89.44	96.66
40	91.66	92.22	93.33	90.00	96.11
50	92.22	92.77	93.33	93.88	96.66
60	92.77	91.66	91.11	91.11	96.66
70	91.66	92.22	91.66	91.11	95.00
80	91.11	91.66	90.55	91.11	95.55
90	91.66	92.22	91.11	91.11	92.77
100	90.00	91.11	91.11	90.55	91.66
110	90.00	88.33	87.22	89.44	91.66

ACC. = Accuracy.



รูปที่ 4.3 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 2 ในแต่ละวิธีการจำแนกประเภท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายละเอียดผลลัพธ์การจำแนกข้อมูลเชิงลึกสามารถแสดงได้ดังตารางที่ 4.8

ตารางที่ 4.8 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 2

วิธีการจำแนกข้อมูล	คลาส	ข้อมูลก่อนสกัดคุณลักษณะ			ข้อมูลหลังจากสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน					
		6000 คุณลักษณะ			50 คุณลักษณะ			30 คุณลักษณะ		
		P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)
Naïve Bayes	Cylindrical	67.6	83.3	74.6	90.9	100	95.2	88.2	100	93.8
	Hook	74.2	76.7	75.4	100	100	100	100	100	100
	Lateral	100	70.0	82.4	87.9	96.7	92.1	87.5	93.3	90.3
	Palm	65.2	50.0	56.6	87.5	93.3	90.3	90.0	90.0	90.0
	Spherical	69.8	100	82.2	100	100	100	100	100	100
	Tip	84.0	70.0	76.4	86.4	63.3	73.1	79.2	63.3	70.4
k-NN	Cylindrical	0.0	0.0	0.0	93.8	100	96.8	93.8	100	96.8
	Hook	0.0	0.0	0.0	100	100	100	100	100	100
	Lateral	16.7	100	28.6	90.6	96.7	93.5	93.5	96.7	95.1
	Palm	0.0	0.0	0.0	87.1	90.0	88.5	84.4	90.0	87.1
	Spherical	0.0	0.0	0.0	100	100	100	100	100	100
	Tip	0.0	0.0	0.0	84.0	70.0	76.4	84.0	70.0	76.4
RBF	Cylindrical	46.9	50.0	48.4	90.9	100	95.2	90.9	100	95.2
	Hook	73.7	46.7	57.1	100	100	100	100	100	100
	Lateral	100	36.7	53.7	85.7	100	92.3	85.7	100	92.3
	Palm	36.7	36.7	36.7	93.3	93.3	93.3	93.3	100	93.3
	Spherical	46.2	100	63.2	100	100	100	100	100	100
	Tip	60.9	46.7	52.8	90.9	66.7	76.9	90.9	66.7	76.9
C4.5	Cylindrical	21.4	30.0	25.0	90.9	100	95.2	96.6	93.3	94.9
	Hook	30.3	33.3	31.7	96.6	93.3	94.9	90.3	93.3	91.8
	Lateral	40.0	46.7	43.1	93.5	96.7	95.1	96.4	90.0	93.1
	Palm	17.9	16.7	17.2	90.6	96.7	93.5	75.0	90.0	81.8
	Spherical	46.2	20.0	27.9	96.7	96.7	96.7	96.6	93.3	94.9
	Tip	41.4	40.0	40.7	96.0	80.0	87.3	85.2	76.7	80.7
SMO	Cylindrical	0.0	0.0	0.0	90.9	100	95.2	93.8	100	96.8
	Hook	30.6	50.0	38.0	100	100	100	100	100	100
	Lateral	92.3	80.0	85.7	96.8	100	98.4	96.8	100	98.4
	Palm	26.3	70.0	38.2	93.8	100	96.8	93.5	96.7	95.1
	Spherical	0.0	0.0	0.0	100	100	100	100	100	100
	Tip	56.0	46.7	50.9	100	80.0	88.9	96.2	83.3	89.3

P.=Precision. R.=Recall. F.=F-Measure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2.1 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 2 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้ว่ามีประสิทธิภาพในการจำแนกข้อมูลที่ต่ำ ซึ่งที่จำนวนคุณลักษณะ 20 คุณลักษณะมีค่าความถูกต้องสูงสุดและต่ำสุดที่ 35.56% ด้วยขั้นตอนการจำแนกแบบเบย์อย่างง่ายและ 24.44% ด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว ตามลำดับ ในขณะข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้นให้มีค่าความถูกต้องสูงสุดถึง 91.67% ด้วยวิธีอัลกอริทึมเอสเอ็มโอและมีค่าความถูกต้องต่ำสุดที่ 83.33% ด้วยขั้นตอนวิธีซีโฟพอยไฟต์ สำหรับผลลัพธ์ของการจำแนกข้อมูลระหว่างการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณและวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันสามารถแสดงได้ดังตารางที่ 4.9 และตารางที่ 4.10 นอกจากนี้สามารถดูรูปแสดงผลลัพธ์เปรียบเทียบการสกัดคุณลักษณะและการจำแนกได้ดังรูปที่ 4.4

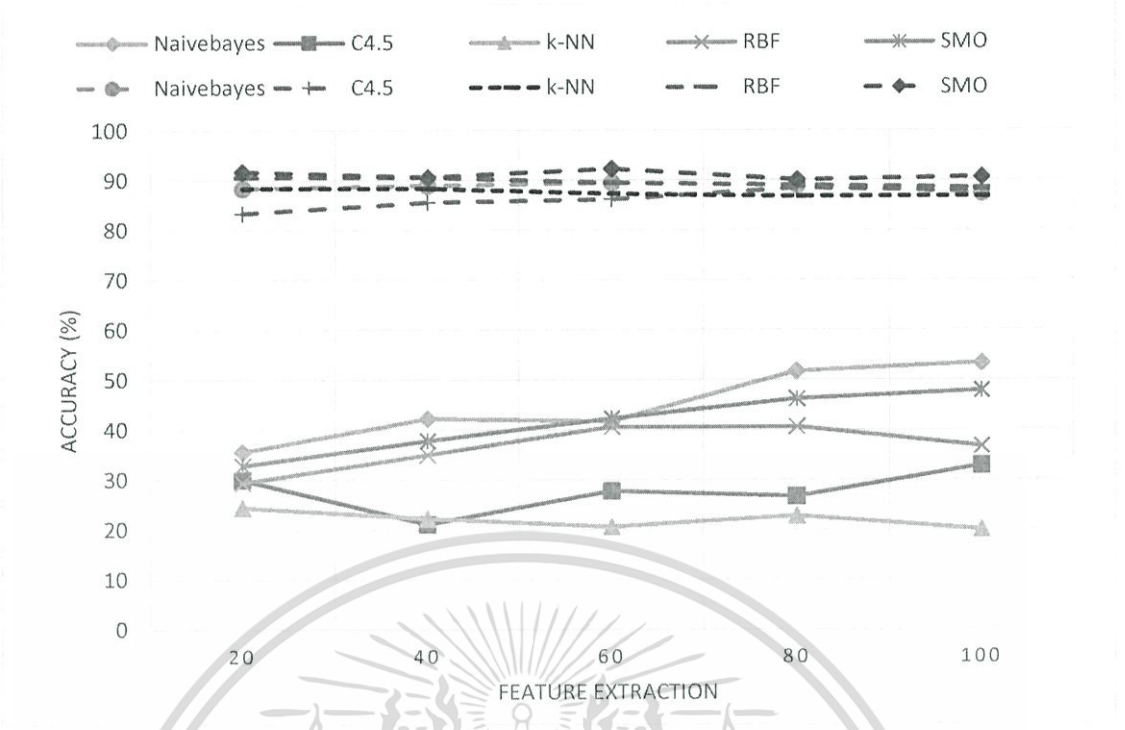
ตารางที่ 4.9 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 2 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	35.56	30.00	24.44	29.44	32.78
40	42.22	21.11	22.22	35.00	37.78
60	41.67	27.78	20.56	40.56	42.22
80	51.67	26.67	22.78	40.56	46.11
100	53.33	32.78	20.00	36.67	47.78

ตารางที่ 4.10 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 2 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	88.33	83.33	88.33	90.56	91.67
40	88.89	85.56	88.33	90.56	90.56
60	89.44	86.11	87.22	89.44	92.22
80	88.89	88.33	86.67	88.89	90.00
100	87.22	87.78	86.67	88.33	90.56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.4 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 2

จากรูปที่ 4.4 เส้นทึบหมายถึงข้อมูลที่ใช้วิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ซึ่งวิธีการจำแนกกับข้อมูลชุดนี้จะเห็นได้ว่าทิศทางของค่าความถูกต้องจะสูงขึ้นเมื่อจำนวนคุณลักษณะที่เพิ่มมากขึ้น โดยที่ขั้นตอนวิธีแบบเบย์อย่างง่าย ให้ค่าความถูกต้องของการจำแนกที่เกิน 50% ที่จำนวนคุณลักษณะที่ 80 และ 100 คุณลักษณะ วิธีการจำแนกรองลงมาคืออัลกอริทึมเอสเอ็มโอ, ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน, ขั้นตอนวิธีซีโฟพอยต์ไฟว์และขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวตามลำดับ สำหรับในส่วนของเส้นปะหมายถึงข้อมูลที่ใช้วิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับการซิงกูลาร์แวลูดีคอมโพซิชัน โดยข้อมูลชุดนี้ทุกวิธีการจำแนกได้ค่าความถูกต้องที่เกินกว่า 80% ในแต่ละจำนวนคุณลักษณะที่นำมาทดสอบ ซึ่งค่าความถูกต้องที่ได้นั้นมีค่าที่ใกล้เคียงกันมากโดยที่มีค่าความถูกต้องห่างกันมากสุดประมาณ 7% และน้อยสุด 3.4%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.3 ผลการทดลองกับข้อมูลผู้ทดลองที่ 3

4.2.3.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกค่าความถูกต้องของข้อมูลผู้ทดลองที่ 3 ที่ผ่านการเตรียมข้อมูลแบบปกติ มีดังต่อไปนี้ เริ่มจากการจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้อง 72.77% การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้อง 22.77% การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้อง 60.55% การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้อง 35.00% และการจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้อง 53.33% ค่าความถูกต้องของการจำแนกข้อมูลแต่ละวิธีสามารถแสดงได้ดังตารางที่ 4.11

ตารางที่ 4.11 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 3

	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
6000	72.77	22.77	60.55	35.00	53.33

ACC. = Accuracy.

4.2.3.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 3 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันแล้วมีดังต่อไปนี้ การจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้องสูงสุดได้ 88.33% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้องสูงสุดได้ 95.55% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้องสูงสุดได้ 95.55% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้องสูงสุดได้ 95.00% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 60 คุณลักษณะ การจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องสูงสุดได้ 96.66% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 50 คุณลักษณะ ผลการจำแนกกับข้อมูลผู้ทดลองที่ 3 ที่ได้รับการแปลงคุณลักษณะของข้อมูลแล้วแสดงได้ดังตารางที่ 4.12 ซึ่งเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติได้ดังรูปที่ 4.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.12 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 3

V2M-SVD	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
10	87.77	93.33	94.44	84.44	89.44
20	87.77	93.88	94.44	93.33	95.00
30	88.33	95.55	95.55	90.00	95.55
40	87.22	92.22	93.33	93.88	93.33
50	87.77	95.00	91.66	90.00	96.66
60	85.55	93.88	87.22	95.00	92.77
70	87.22	92.22	91.66	94.44	94.44
80	85.55	90.55	87.77	94.44	93.88
90	86.66	91.11	87.22	94.44	96.11
100	85.00	88.88	87.22	92.22	92.22
110	85.55	87.77	87.77	92.77	90.00

ACC. = Accuracy.

รูปที่ 4.5 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 3 ในแต่ละวิธีการจำแนกประเภท

รายละเอียดผลลัพธ์การจำแนกข้อมูลเชิงลึกสามารถแสดงได้ดังตารางที่ 4.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.13 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 3

วิธีการจำแนกข้อมูล	คลาส	ข้อมูลก่อนสกัดคุณลักษณะ			ข้อมูลหลังจากสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน					
		6000 คุณลักษณะ			30 คุณลักษณะ			20 คุณลักษณะ		
		P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)
Naïve Bayes	Cylindrical	95.5	93.3	88.9	87.1	90.0	88.5	87.1	90.0	88.5
	Hook	82.9	96.7	89.2	100	96.7	98.3	100	93.3	96.6
	Lateral	47.1	26.7	34.0	84.0	70.0	76.4	81.5	73.3	77.2
	Palm	74.1	66.7	70.2	89.3	83.3	86.2	82.3	80.0	85.7
	Spherical	85.7	80.0	82.8	90.9	100	95.2	88.2	100	93.8
	Tip	55.0	73.3	62.9	79.4	90.0	84.4	79.4	90.0	84.4
k-NN	Cylindrical	0.0	0.0	0.0	90.6	96.7	93.5	93.5	96.7	95.1
	Hook	0.0	0.0	0.0	100	100	100	100	100	100
	Lateral	32.7	53.3	40.5	96.4	90.0	93.1	86.7	86.7	86.7
	Palm	0.0	0.0	0.0	96.4	90.0	93.1	96.3	86.7	91.2
	Spherical	0.0	0.0	0.0	100	96.7	98.3	100	100	100
	Tip	19.1	83.3	31.1	90.9	100	95.2	87.5	93.3	90.3
RBF	Cylindrical	42.2	63.3	50.7	100	96.7	98.3	100	96.7	98.3
	Hook	93.3	93.3	93.3	100	100	100	96.8	100	98.4
	Lateral	64.3	30.0	40.9	82.9	96.7	89.2	95.8	76.7	85.2
	Palm	26.7	26.7	26.7	96.7	96.7	96.7	96.7	96.7	96.7
	Spherical	69.8	100	82.2	100	100	100	100	100	100
	Tip	83.3	50.0	62.5	96.2	83.3	89.3	96.2	83.3	89.3
C4.5	Cylindrical	21.4	30.0	25.0	86.2	83.3	84.7	87.5	93.3	90.3
	Hook	50.0	40.0	44.4	88.2	100	93.8	96.8	100	98.4
	Lateral	35.5	36.7	36.1	86.7	86.7	86.7	100	83.3	90.9
	Palm	20.0	16.7	18.2	96.3	86.7	91.2	89.3	83.3	86.2
	Spherical	50.0	43.3	46.4	100	96.7	98.3	100	100	100
	Tip	40.6	43.3	41.9	83.9	86.7	85.2	88.2	100	93.8
SMO	Cylindrical	33.9	66.7	44.9	100	96.7	98.3	100	96.7	98.3
	Hook	38.5	50.0	43.5	100	100	100	100	100	100
	Lateral	75.0	80.0	77.4	87.9	96.7	92.1	89.3	83.3	86.2
	Palm	80.0	40.0	53.3	96.7	96.7	96.7	96.7	96.7	96.7
	Spherical	0.0	0.0	0.0	100	96.7	98.3	100	100	100
	Tip	71.4	83.3	76.9	92.9	86.7	89.7	84.8	93.3	88.9

P.=Precision. R.=Recall. F.=F-Measure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.3.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 3 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้ว่ามีประสิทธิภาพในการจำแนกข้อมูลที่ต่ำ โดยจำนวนคุณลักษณะที่ 20 ให้ค่าจำแนกสูงสุดได้เพียง 49.44% ด้วยขั้นตอนแบบเบย์อย่างง่าย ในขณะข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้นได้ค่าความถูกต้องสูงสุดถึง 90.00% ด้วยอัลกอริทึมเอสเอ็มโอ สำหรับผลลัพธ์ของการจำแนกข้อมูลระหว่างการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณและวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันสามารถแสดงได้ดังตารางที่ 4.14 และตารางที่ 4.15 นอกจากนี้สามารถดูรูปแสดงผลลัพธ์เปรียบเทียบการสกัดคุณลักษณะและการจำแนกได้ดังรูปที่ 4.6

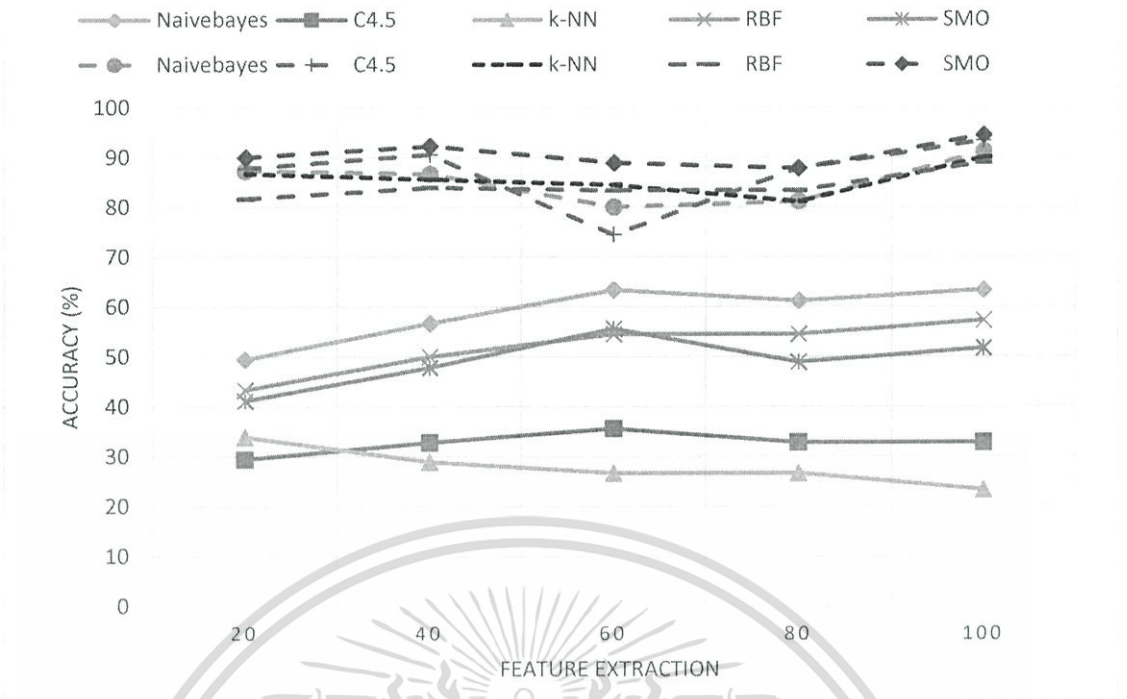
ตารางที่ 4.14 ผลลัพธ์การจำแนกข้อมูลผู้ทดลอง 3 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	49.44	29.44	33.89	43.33	41.11
40	56.67	32.78	28.89	50.00	47.78
60	63.33	35.56	26.67	54.44	55.56
80	61.11	32.78	26.67	54.44	48.89
100	63.33	32.78	23.33	57.22	51.67

ตารางที่ 4.15 ผลลัพธ์การจำแนกข้อมูลผู้ทดลอง 3 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	87.22	87.78	86.67	81.67	90.00
40	86.67	90.56	85.56	83.89	92.22
60	80.00	74.44	84.44	83.33	88.89
80	81.11	87.77	81.11	83.33	87.78
100	91.11	93.33	90.00	88.89	94.44

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.6 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 3

จากรูปที่ 4.6 เส้นทึบ คือ ข้อมูลที่ใช้วิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ซึ่งวิธีการจำแนกกับข้อมูลชุดนี้จะเห็นได้ว่า ขั้นตอนวิธีแบบเบย์อย่างง่าย, ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และอัลกอริทึมเอสเอ็มโอ ให้ค่าความถูกต้องของการจำแนกที่เกิน 50% ที่จำนวนคุณลักษณะที่ 60 80 และ 100 คุณลักษณะและมีแนวโน้มที่ค่าความถูกต้องจะสูงขึ้นตามจำนวนคุณลักษณะที่เพิ่มมากขึ้น สำหรับในส่วนของเส้นปะ คือ ข้อมูลที่ใช้วิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับการซิงกูลาร์แวลูดีคอมโพซิชัน โดยข้อมูลชุดนี้ทุกวิธีการจำแนกให้ผลการได้ค่าความถูกต้องที่เกินกว่า 80% ในแต่ละจำนวนคุณลักษณะที่นำมาทดสอบ ยกเว้นแต่เพียง ขั้นตอนวิธีซีโฟพอยต์ไฟว์ ในการจำแนกข้อมูลที่มีจำนวนคุณลักษณะที่ 60 คุณลักษณะซึ่งให้ค่าความถูกต้องที่ 74.44%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.4 ผลการทดลองกับข้อมูลผู้ทดลองที่ 4

4.2.4.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกค่าความถูกต้องของข้อมูลผู้ทดลองที่ 4 ที่ผ่านการเตรียมข้อมูลแบบปกติมีดังต่อไปนี้ เริ่มจากการจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้อง 73.88% การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้อง 20.00% การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้อง 37.77% การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้อง 33.33% และการจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้อง 50.55% ค่าความถูกต้องของการจำแนกข้อมูลแต่ละวิธีสามารถแสดงได้ดังตารางที่ 4.16

ตารางที่ 4.16 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 4

	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
6000	73.88	20.00	37.77	33.33	50.55

ACC. = Accuracy.

4.2.4.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

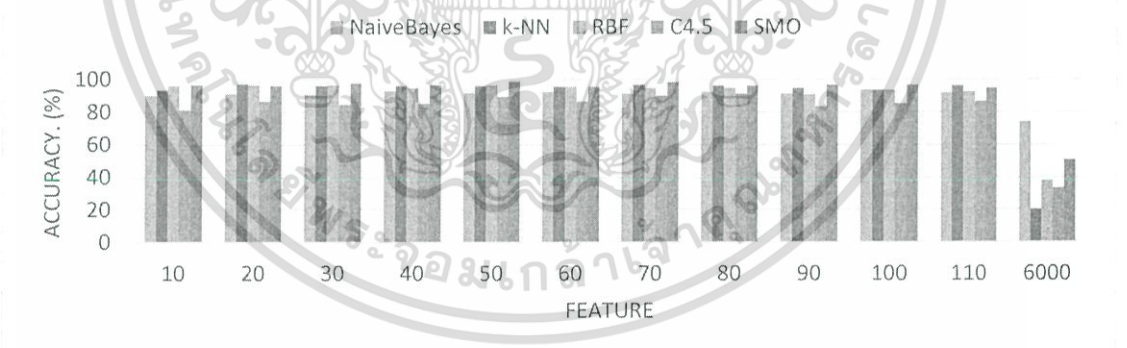
ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 4 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันแล้วมีดังต่อไปนี้ การจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้องสูงสุดได้ 92.77% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 40 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้องสูงสุดได้ 97.22% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 20 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้องสูงสุดได้ 96.66% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 20 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้องสูงสุดได้ 91.11% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 80 คุณลักษณะ การจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องสูงสุดได้ 98.88% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 50 คุณลักษณะ ผลการจำแนกกับข้อมูลผู้ทดลองที่ 4 ที่ได้รับการแปลงคุณลักษณะของข้อมูลแล้วแสดงได้ดังตารางที่ 4.17 ซึ่งเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติได้ดังรูปที่ 4.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.17 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 4

V2M-SVD	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
10	90.00	93.33	96.11	81.11	96.66
20	91.11	97.22	96.66	86.66	96.11
30	90.55	96.11	96.66	84.44	97.77
40	92.77	96.11	95.00	85.00	96.66
50	91.66	96.11	96.66	89.44	98.88
60	92.22	95.55	95.55	86.66	95.55
70	91.11	96.66	94.44	90.00	98.33
80	92.22	96.11	94.44	91.11	96.11
90	91.11	94.44	90.55	83.33	96.66
100	93.33	93.33	93.33	85.00	96.66
110	91.66	96.11	92.22	86.66	94.44

ACC. = Accuracy.



รูปที่ 4.7 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 4 ในแต่ละวิธีการจำแนกประเภท

รายละเอียดผลลัพธ์การจำแนกข้อมูลเชิงลึกสามารถแสดงข้อมูลได้ดังตารางที่ 4.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.18 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 4

วิธีการจำแนกข้อมูล	คลาส	ข้อมูลก่อนสกัดคุณลักษณะ			ข้อมูลหลังจากสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน					
		6000 คุณลักษณะ			50 คุณลักษณะ			70 คุณลักษณะ		
		P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)
Naïve Bayes	Cylindrical	70.3	86.7	77.6	100	96.7	98.3	100	100	100
	Hook	90.6	96.7	93.5	96.8	100	98.4	100	100	100
	Lateral	54.3	63.3	58.5	84.8	93.3	88.9	83.3	83.3	83.3
	Palm	84.2	53.3	65.3	86.2	83.3	84.7	86.7	86.7	86.7
	Spherical	78.8	86.7	82.5	93.5	96.7	95.1	96.8	100	98.4
	Tip	70.8	56.7	63.0	88.9	80.0	84.2	79.3	76.7	78.0
k-NN	Cylindrical	0.0	0.0	0.0	100	100	100	100	100	100
	Hook	0.0	0.0	0.0	100	100	100	100	100	100
	Lateral	50.0	3.3	6.3	93.3	93.3	93.3	96.7	96.7	96.7
	Palm	19.5	96.7	32.4	92.9	86.7	89.7	96.3	86.7	91.2
	Spherical	0.0	0.0	0.0	100	100	100	100	100	100
	Tip	20.7	20.0	20.3	90.6	96.7	93.5	87.9	96.7	92.1
RBF	Cylindrical	0.0	0.0	0.0	100	100	100	100	100	100
	Hook	75.0	100	85.7	100	100	100	100	100	100
	Lateral	25.0	6.7	10.5	100	93.3	93.6	96.4	90.0	93.1
	Palm	63.6	23.3	34.1	87.5	93.3	90.3	83.3	83.3	83.3
	Spherical	22.0	66.7	33.1	100	100	100	100	100	100
	Tip	47.4	30.0	36.7	93.3	93.3	93.3	87.5	93.3	90.3
C4.5	Cylindrical	34.6	30.0	32.1	93.3	93.3	93.3	100	96.7	98.3
	Hook	66.7	40.0	50.0	93.8	100	96.8	100	100	100
	Lateral	20.0	23.3	21.5	92.6	83.3	87.7	82.8	80.0	81.4
	Palm	40.0	46.7	43.1	83.3	83.3	83.3	85.7	80.0	82.8
	Spherical	27.8	33.3	30.3	100	100	100	90.9	100	95.2
	Tip	26.7	26.7	26.7	74.2	76.7	75.4	80.6	83.3	82.0
SMO	Cylindrical	31.6	100	48.0	100	100	100	100	100	100
	Hook	0.0	0.0	0.0	100	100	100	100	100	100
	Lateral	72.7	80.0	76.2	96.8	100	98.4	100	100	100
	Palm	76.7	76.7	76.7	96.7	96.7	96.7	96.6	93.3	94.9
	Spherical	0.0	0.0	0.0	100	100	100	100	100	100
	Tip	63.6	46.7	53.8	100	96.7	98.3	93.5	96.7	95.1

P.=Precision. R.=Recall. F.=F-Measure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.4.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 4 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้ว่ามีประสิทธิภาพในการจำแนกข้อมูลที่ต่ำ โดยจำนวนคุณลักษณะที่ 20 ให้ค่าจำแนกสูงสุดได้เพียง 42.22% ด้วยอัลกอริทึมเอสเอ็มโอ ในขณะข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้นได้ค่าความถูกต้องสูงสุดถึง 96.67% ด้วยอัลกอริทึมเอสเอ็มโอ สำหรับผลลัพธ์ของการจำแนกข้อมูลระหว่างการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณและวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันสามารถแสดงได้ดังตารางที่ 4.19 และตารางที่ 4.20 นอกจากนี้สามารถดูรูปแสดงผลลัพธ์เปรียบเทียบการสกัดคุณลักษณะและการจำแนกได้ดังรูปที่ 4.8

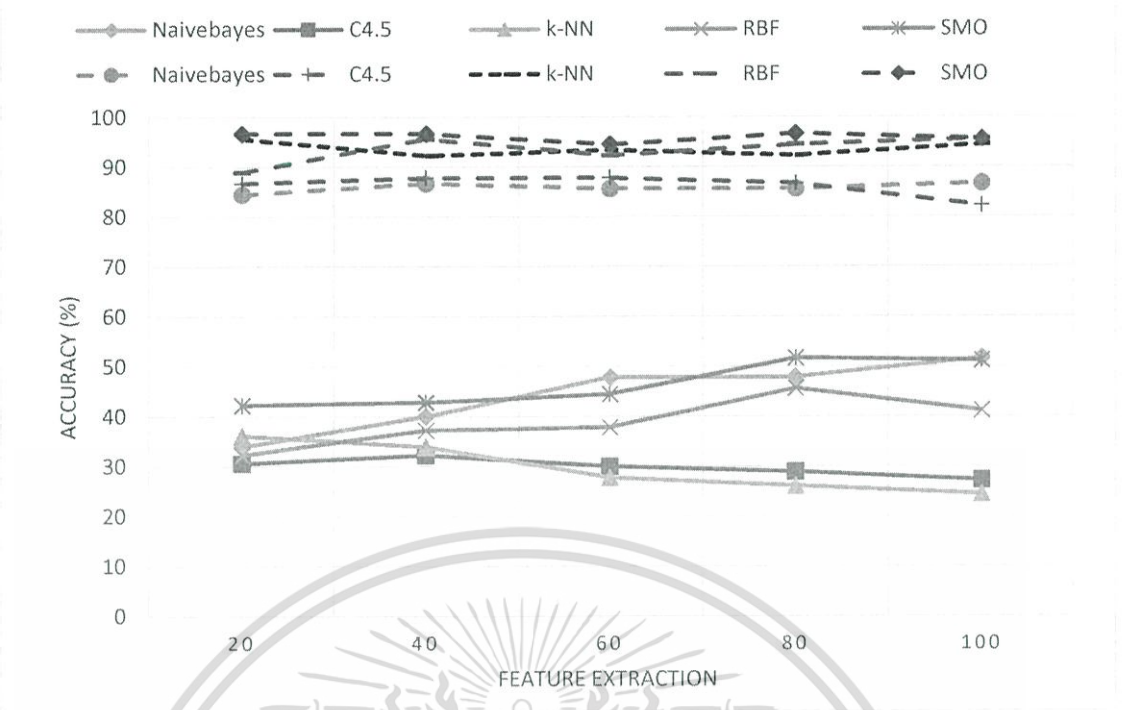
ตารางที่ 4.19 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 4 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	33.89	30.56	36.11	32.22	42.22
40	40.00	32.22	33.89	37.22	42.78
60	47.78	30.00	27.78	37.78	44.44
80	47.78	28.89	26.11	45.56	51.67
100	51.67	27.22	24.44	41.11	51.11

ตารางที่ 4.20 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 4 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	84.44	86.67	95.56	88.89	96.67
40	86.67	87.78	92.22	95.56	96.67
60	85.56	87.78	93.33	92.22	94.44
80	85.56	86.67	92.22	94.44	96.67
100	86.67	82.22	94.44	95.56	95.56

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 4

จากรูปที่ 4.8 เส้นทึบหมายถึงข้อมูลที่ใช้วิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ในช่วงจำนวนคุณลักษณะที่ 20 และ 40 คุณลักษณะ ทุกขั้นตอนวิธีการจำแนกจะอยู่ในเกณฑ์ที่ให้ค่าความถูกต้องของการจำแนกที่ไม่น้อยกว่า 30% และไม่มากกว่า 45% แต่เมื่อจำนวนคุณลักษณะที่เพิ่มขึ้นจะเห็นได้อย่างชัดเจนว่า ขั้นตอนวิธีแบบเบย์อย่างง่าย, ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และอัลกอริทึมเอสเอ็มโอ มีค่าความถูกต้องเพิ่มสูงขึ้นแตกต่างกับขั้นตอนวิธีซีโฟพอยต์ไฟว์และขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัว ที่ได้ค่าความถูกต้องของการจำแนกที่น้อยสูงเมื่อจำนวนคุณลักษณะที่เพิ่มมากขึ้น สำหรับในส่วนของเส้นปะหมายถึงข้อมูลที่ใช้วิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับการซิงกูลาร์แวลูดีคอมโพซิชัน โดยข้อมูลชุดนี้ทุกวิธีการจำแนกได้ค่าความถูกต้องที่เกินกว่า 80% ในแต่ละจำนวนคุณลักษณะที่นำมาทดสอบ และขั้นตอนวิธีที่ให้ค่าความถูกต้องของการจำแนกได้สูงสุดเท่ากันที่ 96.56% คือ อัลกอริทึมเอสเอ็มโอ และขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน ที่จำนวนคุณลักษณะที่ 100 คุณลักษณะ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.5 ผลการทดลองกับข้อมูลผู้ทดลองที่ 5

4.2.5.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกค่าความถูกต้องของข้อมูลผู้ทดลองที่ 5 ที่ผ่านการเตรียมข้อมูลแบบปกติมีดังต่อไปนี้ เริ่มจากการจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้อง 61.66% การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้อง 33.88% การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้อง 58.33% การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้อง 33.33% และการจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้อง 43.33% ค่าความถูกต้องของการจำแนกข้อมูลแต่ละวิธีสามารถแสดงได้ดังตารางที่ 4.21

ตารางที่ 4.21 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของผู้ทดลองที่ 5

	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
6000	61.66	33.88	58.33	33.33	43.33

ACC. = Accuracy.

4.2.5.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 5 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันแล้วมีดังต่อไปนี้ การจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้องสูงสุดได้ 90.00% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้องสูงสุดได้ 89.44% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 20 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้องสูงสุดได้ 93.88% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 50 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้องสูงสุดได้ 93.33% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 40 คุณลักษณะ การจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องสูงสุดได้ 100% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 50

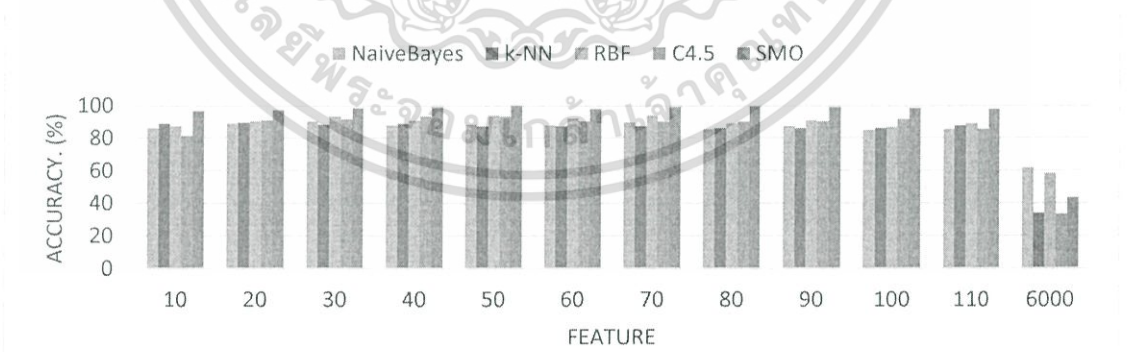
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คุณลักษณะ ผลการจำแนกกับข้อมูลผู้ทดลองที่ 5 ที่ได้รับการแปลงคุณลักษณะของข้อมูลแล้วแสดงได้ดังตารางที่ 4.22 ซึ่งเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติได้ดังรูปที่ 4.9

ตารางที่ 4.22 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 5

V2M-SVD	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
10	86.11	88.88	87.22	81.66	96.66
20	88.88	89.44	90.55	91.11	97.22
30	90.00	88.33	93.33	91.66	98.33
40	87.77	88.88	91.11	93.33	98.88
50	88.33	87.22	93.88	93.33	100
60	87.77	87.22	92.22	90.55	97.77
70	89.44	87.22	93.88	90.00	98.88
80	85.55	86.11	88.88	90.00	99.44
90	87.22	86.11	91.11	90.55	98.88
100	85.00	86.11	86.66	91.66	98.33
110	85.55	87.77	88.88	85.55	97.77

ACC. = Accuracy.



รูปที่ 4.9 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 5 ในแต่ละวิธีการจำแนกประเภท

รายละเอียดผลลัพธ์การจำแนกข้อมูลเชิงลึกสามารถแสดงได้ดังตารางที่ 4.23

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.23 ผลการวัดค่าความถูกต้องเชิงลึกของการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองที่ 5

วิธีการจำแนกข้อมูล	คลาส	ข้อมูลก่อนสกัดคุณลักษณะ			ข้อมูลหลังจากสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน					
		6000 คุณลักษณะ			50 คุณลักษณะ			40 คุณลักษณะ		
		P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)	P.(%)	R.(%)	F.(%)
Naïve Bayes	Cylindrical	60.0	20.0	30.0	100	43.3	60.5	87.5	46.7	60.9
	Hook	40.0	60.0	48.0	63.8	100	77.9	65.1	93.3	76.7
	Lateral	46.5	66.7	54.8	100	96.7	98.3	100	96.7	98.3
	Palm	81.8	90.0	85.7	90.9	100	95.2	90.9	100	95.2
	Spherical	76.9	100	87.0	96.8	100	98.4	93.8	100	96.8
	Tip	100	33.3	50.0	100	90.0	94.7	100	90.0	94.7
k-NN	Cylindrical	0.0	0.0	0.0	71.4	66.7	69.0	80.0	66.7	72.7
	Hook	100	10	18.2	68.8	73.3	71.0	71.4	83.3	76.9
	Lateral	62.9	73.3	67.7	93.5	96.7	95.1	93.5	96.7	95.1
	Palm	55.6	16.7	25.6	90.6	96.7	93.5	90.6	96.7	93.5
	Spherical	100	3.3	6.5	100	100	100	100	100	100
	Tip	22.7	100	30.7	100	90.0	94.7	100	90.0	94.7
RBF	Cylindrical	38.6	56.7	45.9	92.6	83.3	87.7	90.9	66.7	76.9
	Hook	58.6	56.7	57.6	87.5	93.3	90.3	71.8	93.3	81.2
	Lateral	43.3	43.3	43.3	93.5	96.7	95.1	96.7	96.7	96.7
	Palm	81.8	60.0	69.2	93.8	100	96.8	93.8	100	96.8
	Spherical	66.7	100	80.0	96.8	100	98.4	100	100	100
	Tip	100	33.3	50.0	100	90.0	94.7	100	90.0	94.7
C4.5	Cylindrical	28.0	23.3	25.5	93.3	93.3	93.3	93.3	93.3	93.3
	Hook	17.6	20.0	18.8	92.6	83.3	87.7	92.6	83.3	87.7
	Lateral	31.0	43.3	36.1	96.7	96.7	96.7	96.7	96.7	96.7
	Palm	34.5	33.3	33.9	83.3	100	90.9	83.3	100	90.0
	Spherical	40.0	26.7	32.0	100	93.3	96.6	100	93.3	96.6
	Tip	53.3	53.3	53.3	96.6	93.3	94.9	96.6	93.3	94.9
SMO	Cylindrical	22.1	50.0	30.6	100	100	100	100	96.7	98.3
	Hook	62.5	33.3	43.5	100	100	100	96.8	100	98.4
	Lateral	91.3	70.0	79.2	100	100	100	100	96.7	98.3
	Palm	26.8	50.0	34.9	100	100	100	96.8	100	98.4
	Spherical	0.0	0.0	0.0	100	100	100	100	100	100
	Tip	100	56.7	72.3	100	100	100	100	100	100

P.=Precision. R.=Recall. F.=F-Measure

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.5.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองที่ 5 ที่ได้รับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้ว่ามีประสิทธิภาพในการจำแนกข้อมูลที่ต่ำ โดยจำนวนคุณลักษณะที่ 20 ได้ค่าความถูกต้องต่ำสุดที่ 30.00% ด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว และค่าจำแนกสูงสุดได้เพียง 45.56% ด้วยขั้นตอนการจำแนกแบบเบย์อย่างง่าย ในขณะข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้น ในระดับเดียวกันคือ จำนวนคุณลักษณะที่ 20 ได้ค่าความถูกต้องสูงสุดถึง 92.22% ด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ และขั้นตอนการจำแนกที่ได้ค่าความถูกต้องต่ำสุดคือ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัวได้ค่าความถูกต้องต่ำสุดที่ 76.67% สำหรับผลลัพธ์ของการจำแนกข้อมูลระหว่างการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์และวิธีซิงกูลาร์แวลูดีคอมโพซิชันสามารถแสดงได้ดังตารางที่ 4.24 และตารางที่ 4.25 นอกจากนี้สามารถดูรูปแสดงผลลัพธ์เปรียบเทียบการสกัดคุณลักษณะและการจำแนกได้ดังรูปที่ 4.10

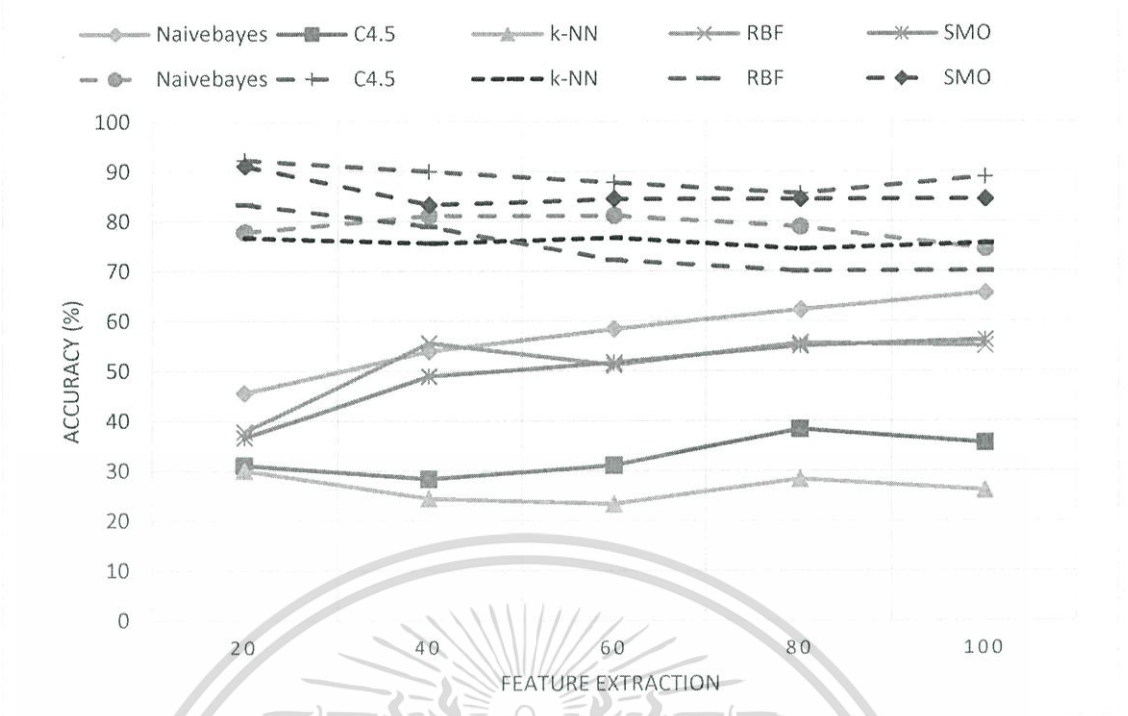
ตารางที่ 4.24 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 5 ที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	45.56	31.11	30.00	37.78	36.67
40	53.89	28.33	24.44	55.56	48.89
60	58.33	31.11	23.33	51.11	51.67
80	62.22	38.33	28.33	55.56	55.00
100	65.56	35.56	26.11	55.00	56.11

ตารางที่ 4.25 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองที่ 5 ที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	77.78	92.22	76.67	83.33	91.11
40	81.11	90.00	75.56	78.89	83.33
60	81.11	87.78	76.67	72.22	84.44
80	78.89	85.56	74.44	70.00	84.44
100	74.44	88.89	75.56	70.00	84.44

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.10 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองที่ 5

จากรูปที่ 4.10 เส้นทึบหมายถึงข้อมูลที่ใช้วิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ในช่วงจำนวนคุณลักษณะที่ 60 80 และ 100 คุณลักษณะ ขั้นตอนวิธี ขั้นตอนวิธีแบบเบย์อย่างง่าย, ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน และอัลกอริทึมเอสเอ็มโอ ค่าความถูกต้องของการจำแนกที่ไม่น้อยกว่า 50% สำหรับในส่วนของเส้นปะหมายถึงข้อมูลที่ใช้วิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับการซิงกูลาร์แวลูดีคอมโพซิชัน โดยข้อมูลชุดนี้ทุกวิธีการจำแนกให้ผลการได้ค่าความถูกต้องที่ต่ำสุดที่ 70% โดยเป็นค่าความถูกต้องของการจำแนกของ ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.6 ผลการทดลองโดยรวมของข้อมูลผู้ทดลองทั้งหมด

4.2.6.1 ผลลัพธ์การจำแนกของข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

ผลลัพธ์การจำแนกค่าความถูกต้องของข้อมูลโดยรวมของข้อมูลผู้ทดลองทั้งหมดที่ผ่านการเตรียมข้อมูลแบบปกติมีดังต่อไปนี้ เริ่มจากการจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้อง 71.77% การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้อง 21.88% การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้อง 52.43% การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้อง 32.55% และการจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้อง 47.44% ค่าความถูกต้องของการจำแนกข้อมูลแต่ละวิธีสามารถแสดงได้ดังตารางที่ 4.26

ตารางที่ 4.26 ผลการวัดค่าความถูกต้องของการจำแนกข้อมูล 5 วิธีกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติของข้อมูลโดยรวม

	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
6000	71.77	21.88	52.43	32.55	47.44

ACC. = Accuracy.

4.2.6.2 ผลลัพธ์การจำแนกข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ

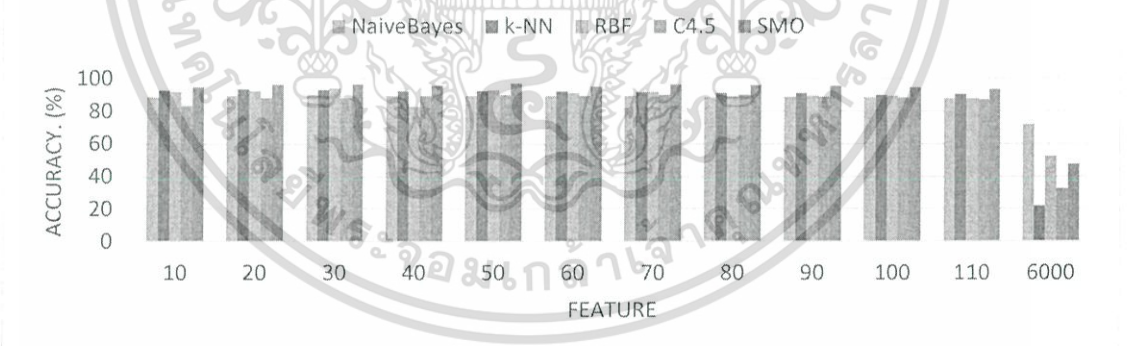
ผลลัพธ์การจำแนกโดยเฉลี่ยของข้อมูลผู้ทดลองทั้ง 5 คนที่ได้รับการแปลงคุณลักษณะของข้อมูลแล้วมีดังต่อไปนี้ การจำแนกข้อมูลด้วยขั้นตอนวิธีแบบเบย์อย่างง่ายได้ค่าความถูกต้องสูงสุดได้ 89.77% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีเพื่อนบ้านใกล้สุด *k* ตัวได้ค่าความถูกต้องสูงสุดได้ 93.55% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 20 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันได้ค่าความถูกต้องสูงสุดได้ 93.88% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 30 คุณลักษณะ การจำแนกข้อมูลด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์ได้ค่าความถูกต้องสูงสุดได้ 90.22% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 80 คุณลักษณะ การจำแนกข้อมูลด้วยอัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องสูงสุดได้ 97.21% ซึ่งชุดข้อมูลที่ให้ค่าความถูกต้องนี้มีจำนวนคุณลักษณะที่ 50 คุณลักษณะ ผลการจำแนกกับข้อมูลผู้ทดลองทั้งหมดที่ได้รับการแปลงคุณลักษณะของข้อมูลแล้วแสดงได้ดังตารางที่ 4.27

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.27 ผลการวัดค่าความถูกต้องของวิธีการจำแนกข้อมูล 5 วิธีกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองทั้งหมด

V2M-SVD	ขั้นตอนวิธีการจำแนกข้อมูล				
จำนวนคุณลักษณะ	NaiveBayes ACC. (%)	*k*-NN ACC. (%)	RBF ACC. (%)	C4.5 ACC. (%)	SMO ACC. (%)
10	88.33	92.66	91.88	83.32	94.88
20	89.44	93.55	92.21	88.10	96.22
30	89.77	93.21	93.88	88.10	96.66
40	89.66	92.33	82.77	89.33	95.66
50	89.55	92.55	92.88	90.21	97.21
60	89.32	92.10	90.88	89.66	95.21
70	89.66	91.77	91.99	90.11	96.44
80	88.33	91.21	88.88	90.22	95.88
90	88.66	90.99	89.66	89.1	95.55
100	88.22	89.77	89.21	88.21	94.44
110	87.88	90.44	87.77	87.21	93.32

ACC. = Accuracy.



รูปที่ 4.11 การเปรียบเทียบค่าความถูกต้องของการจำแนกข้อมูลระหว่างข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติและข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองทั้งหมด ในแต่ละวิธีการจำแนกประเภท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.6.3 ผลลัพธ์การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

ผลลัพธ์การจำแนกกับข้อมูลผู้ทดลองทั้ง 5 คนที่ได้รับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้ว่ามีประสิทธิภาพในการจำแนกข้อมูลที่ต่ำ โดยจำนวนคุณลักษณะที่ 20 ได้ค่าความถูกต้องต่ำสุดที่ 29.78% ด้วยขั้นตอนวิธีซีโฟพอยต์ไฟว์และค่าจำแนกสูงสุดได้เพียง 40.00% ด้วยขั้นตอนการจำแนกแบบเบย์อย่างง่าย ในขณะข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้นได้ค่าความถูกต้องต่ำสุดที่ 84.67% ด้วยขั้นตอนการจำแนกแบบเบย์อย่างง่ายและขั้นตอนการจำแนกที่ได้ค่าความถูกต้องสูงสุด คือ อัลกอริทึมเอสเอ็มโอได้ค่าความถูกต้องที่ 91.67% สำหรับผลลัพธ์ของการจำแนกข้อมูลระหว่างการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์และวิธีซิงกูลาร์แวลูดีคอมโพซิชันสามารถแสดงได้ดังตารางที่ 4.28 และตารางที่ 4.29 นอกจากนี้สามารถดูรูปแสดงผลลัพธ์เปรียบเทียบการสกัดคุณลักษณะและการจำแนกได้ดังรูปที่ 4.12

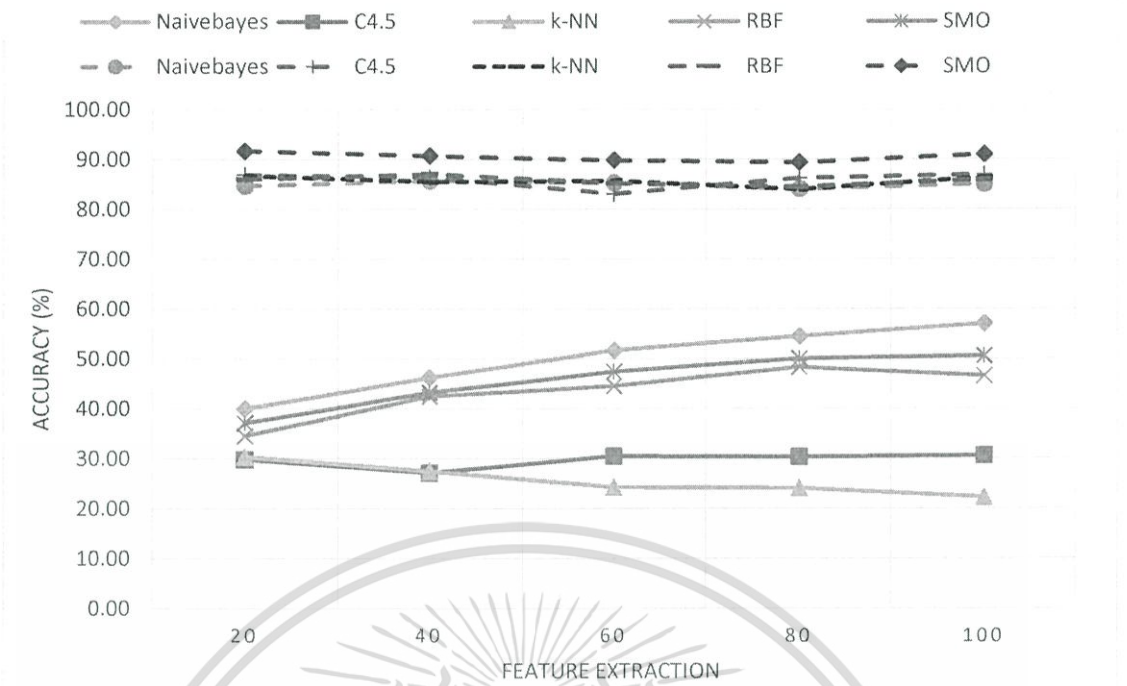
ตารางที่ 4.28 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองทั้งหมดที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	40.00	29.78	30.33	34.55	37.11
40	46.11	27.11	27.44	42.44	43.22
60	51.67	30.45	24.22	44.44	47.33
80	54.44	30.33	24.00	48.22	50.00
100	57.00	30.56	22.22	46.56	50.56

ตารางที่ 4.29 ผลลัพธ์การจำแนกข้อมูลผู้ทดลองทั้งหมดที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน

จำนวนคุณลักษณะ	NaiveBayes ACC.(%)	C4.5 ACC.(%)	*k*-NN ACC.(%)	RBF ACC.(%)	SMO ACC.(%)
20	84.67	86.89	86.67	86.00	91.67
40	85.67	86.34	85.45	87.00	90.67
60	85.22	83.00	85.67	85.22	89.78
80	84.00	86.22	83.89	84.44	89.33
100	85.00	86.89	86.33	85.89	91.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 การเปรียบเทียบวิธีสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันและวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับข้อมูลผู้ทดลองทั้งหมด

จากรูปที่ 4.12 เส้นทึบหมายถึงข้อมูลที่ใช้วิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ จะเห็นได้อย่างชัดเจนตั้งแต่จำนวนคุณลักษณะที่ 20 ถึง 100 ว่าขั้นตอนวิธีที่จำแนกกับข้อมูลชุดนี้ดีที่เรียงลำดับจากมากไปหาน้อย คือ ขั้นตอนวิธีแบบเบย์อย่างง่าย อัลกอริทึมเอสเอ็มโอ ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน ขั้นตอนวิธีซีโฟพอยต์ไฟว์ และ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว โดยที่ขั้นตอนวิธีแบบเบย์อย่างง่าย ทำการจำแนกได้ค่าความถูกต้องสูงสุดไม่เกิน 60% ที่จำนวนคุณลักษณะที่ 100 คุณลักษณะ สำหรับในส่วนของเส้นปะหมายถึงข้อมูลที่ใช้วิธีแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับการซิงกูลาร์แวลูดีคอมโพซิชันเห็นได้อย่างชัดเจนในส่วนของขั้นตอนวิธีที่จำแนกข้อมูลได้ดีที่สุดลำดับที่ 1 คืออัลกอริทึมเอสเอ็มโอ โดยเฉลี่ยค่าความถูกต้องของการจำแนกนั้นอยู่ประมาณ 90%

4.2.6.4 ผลลัพธ์ของการวัดค่าความระลึกของข้อมูลที่สกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันกับข้อมูลผู้ทดลองทั้ง 5 คน

ในการทดลองกับข้อมูลผู้ทดลองทั้ง 5 คนมีผลการจำแนกเชิงลึกของข้อมูลที่สกัดคุณลักษณะด้วยวิธีวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันในแต่ละคนที่ดีที่สุด 2 อันดับแรก ในที่นี้จะดูเฉพาะค่าความระลึกที่ได้ 100% ทั้ง 2 ชุดข้อมูลที่ดีที่สุด ซึ่งได้ผลดังนี้ ชุดข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผู้ทดลองที่ 1 ไม่มีคลาสใดที่ได้ค่าความระลึก 100% ทั้ง 2 ชุดข้อมูล ชุดข้อมูลผู้ทดลองที่ 2 คลาสที่ได้ค่าความระลึก 100% ทั้ง 2 ชุดข้อมูลคือ การจับวัตถุทรงกระบอก การจับวัตถุที่มีน้ำหนักมาก การจับวัตถุที่แบนบางและการจับวัตถุทรงกลม ชุดข้อมูลผู้ทดลองที่ 3 คลาสที่ได้ค่าความระลึก 100 % มีเพียงคลาสเดียวคือการจับวัตถุที่มีน้ำหนักมาก ชุดข้อมูลผู้ทดลองที่ 4 คลาสที่ได้ค่าความระลึก 100 % เหมือนกับชุดข้อมูลผู้ทดลองที่ 3 และชุดข้อมูลผู้ทดลองที่ 5 คลาสที่ได้ค่าความระลึก 100 % คือการจับวัตถุที่มีน้ำหนักมาก การจับวัตถุโดยหันฝ่ามือเข้าหาตัวเอง การจับวัตถุทรงกลมและการจับวัตถุขนาดเล็ก ซึ่งผลลัพธ์ทั้งหมดสามารถแสดงได้ดังตารางที่ 4.30

ตารางที่ 4.30 ผลลัพท์ค่าความระลึกของคลาสที่ได้ 100% ในชุดข้อมูลที่ให้ผลการจำแนกดีที่สุด 2 อันดับของแต่ละข้อมูลผู้ทดลอง

คลาส	ผู้ทดลองที่ 1	ผู้ทดลองที่ 2	ผู้ทดลองที่ 3	ผู้ทดลองที่ 4	ผู้ทดลองที่ 5
Cylindrical	-	✓	-	✓	-
Hook	-	✓	✓	✓	✓
Lateral	-	✓	-	✓	-
Palm	-	-	-	-	✓
Spherical	-	✓	-	✓	✓
Tip	-	-	-	-	✓

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

งานวิจัยนี้ได้ทำการแปลงคุณลักษณะของข้อมูลด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันโดยใช้งานกับข้อมูลคลื่นสัญญาณไฟฟ้ากล้ามเนื้อแบบผิวสัมผัสเพื่อลดจำนวนของคุณลักษณะลงและทำการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติเปรียบเทียบกับชุดข้อมูลที่มีการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันด้วยวิธีการจำแนก 5 วิธี คือ ขั้นตอนวิธีซีโฟพอยต์ไฟว์ ขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว ขั้นตอนวิธีแบบเบย์อย่างง่าย ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชันและขั้นตอนวิธีเอสเอ็มโออัลกอริทึม และทำการจำแนกชุดข้อมูลที่มีการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเปรียบเทียบกับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ในการวัดประสิทธิภาพ

ข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติมีจำนวนคุณลักษณะของข้อมูลอยู่ที่ 6,000 คุณลักษณะหลังผ่านการจำแนกประเภทด้วยวิธีต่าง ๆ ได้ค่าความถูกต้องของการจำแนกในระดับที่ต่ำซึ่งส่วนใหญ่ให้ค่าความถูกต้องได้ไม่เกิน 50% เรียงลำดับวิธีการจำแนกประเภทจากค่าความถูกต้องสูงสุดไปหาน้อยสุด คือ ขั้นตอนวิธีแบบเบย์อย่างง่าย ขั้นตอนวิธีโครงข่ายประสาทเทียมเรเดียลเบสิสฟังก์ชัน ขั้นตอนวิธีเอสเอ็มโออัลกอริทึม ขั้นตอนวิธีซีโฟพอยต์ไฟว์และขั้นตอนวิธีเพื่อนบ้านใกล้สุด k ตัว แต่เมื่อทำการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันแล้วจำนวนคุณลักษณะที่ได้ลดลงเป็นอย่างมาก โดยเหลือเพียงไม่เกิน 110 คุณลักษณะและยังช่วยเพิ่มประสิทธิภาพของข้อมูลให้ง่ายต่อการจำแนกประเภท โดยทุกวิธีการจำแนกประภทสามารถให้ค่าความถูกต้องได้เกินกว่า 80% ขึ้นไป ทั้งนี้วิธีการจำแนกที่ดีที่สุดจากวิธีการจำแนกประเภททั้ง 5 วิธี คือ อัลกอริทึมเอสเอ็มโอ ที่สามารถให้ค่าความถูกต้องได้โดยเฉลี่ยแล้วอยู่ที่ 90% ซึ่งเป็นไปตามสมมติฐานที่ได้กล่าวไว้ในบทที่ 1 นอกจากนี้สามารถยืนยันประสิทธิภาพของการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน โดยผู้วิจัยได้ทำการเปรียบเทียบกับการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ผลลัพธ์ที่ได้แสดงให้เห็นว่าการสกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณให้ค่าความถูกต้องได้ดีเมื่อมีจำนวนคุณลักษณะที่เพิ่มมากขึ้น แต่หากใช้จำนวนคุณลักษณะในปริมาณที่น้อยจะมีประสิทธิภาพในการจำแนกที่ต่ำ โดยให้ค่าความถูกต้องได้ไม่เกิน 60% และมีข้อสังเกตได้อยู่อย่างหนึ่ง คือ การจำแนกข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณกับการจำแนกข้อมูลที่ผ่านการเตรียมข้อมูลแบบปกติ วิธีการจำแนกที่ให้ค่าความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถูกต้องได้ดีสำหรับข้อมูลทั้ง 2 ชุด คือ ขั้นตอนวิธีแบบเบย์อย่างง่าย สำหรับข้อมูลที่ได้รับการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันเมื่อเปรียบเทียบกับข้อมูลที่สกัดคุณลักษณะด้วยวิธีการวิเคราะห์เชิงเส้นแบบดิสคริมิแนนท์ทีละช่องสัญญาณ ในจำนวนคุณลักษณะที่เท่ากัน การสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันสามารถทำการจำแนกข้อมูลได้ดีกว่าในทุกจำนวนคุณลักษณะ และทุกวิธีการจำแนกประเภทที่ใช้ในการทดลอง

ผลการทดลองในการวัดค่าความถูกต้องเชิงลึก โดยใช้ข้อมูลหลังการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชันนั้น คลาสที่อัลกอริทึมเอสเอ็มโอสามารถจำแนกได้ค่าความระลึกสูงถึง 100% คือ คลาสการจับวัตถุที่มีน้ำหนักมาก โดยได้ค่าความระลึก 100% จากข้อมูลผู้ทดลอง 4 คนในจำนวนข้อมูลผู้ทดลอง 5 คนและคลาสการจับวัตถุทรงกลมได้ค่าความระลึก 100% จากข้อมูลผู้ทดลอง 3 คนในจำนวนข้อมูลผู้ทดลอง 5 คน

5.2 ข้อเสนอแนะ

งานวิจัยนี้ได้ใช้เทคนิคการสกัดคุณลักษณะด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน โดยที่ข้อมูลที่ใช้ในงานวิจัยเป็นข้อมูลที่ไม่ได้ผ่านการลดคลื่นสัญญาณรบกวนเมื่อทำการจำแนกด้วยวิธีต่าง ๆ จึงให้ผลลัพธ์ได้ตามที่กล่าวมาในข้างต้น ผู้วิจัยมีข้อเสนอแนะ ควรใช้เทคนิคการลดคลื่นสัญญาณรบกวนก่อนใช้เทคทิคการสกัดคุณลักษณะของข้อมูลด้วยวิธีการแปลงเวกเตอร์เป็นเมทริกซ์ร่วมกับซิงกูลาร์แวลูดีคอมโพซิชัน น่าจะเป็นการเพิ่มประสิทธิภาพของข้อมูลให้มีค่าความถูกต้องของการจำแนกได้ดียิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง

[1] รศ.ดร.สมชาย รัตนทองคำ. 2015. **การตรวจประสาท-กล้ามเนื้อด้วยไฟฟ้า**. [Online]. Available : http://ams.kku.ac.th/aalearn/resource/edoc/es54/emgdoc54.pdf.

[2] จิตฤดี โพธิศิริ. 2549. **กายวิภาคศาสตร์**. โครงการสวัสดิการวิชาการ สถาบันพระบรมราชชนก.

[3] วัดผล. 2015. **เมตริกซ์ในการวิเคราะห์การถดถอย**. [Online]. Available : http://www.watpon.com/regression/chap4.pdf.

[4] Massachusetts Institute of Technology. 2015. **Singular Value Decomposition (SVD) tutorial**. [Online]. Available : http://web.mit.edu/be.400/www/SVD/Singular_Value_Decomposition.htm.

[5] Aly A. Farag, and Shireen Y. Elhabian. 2008. **A Tutorial on Data Reduction Linear Discriminant Analysis (LDA)**. [Online]. Available : http://www.di.univr.it/documenti/Occorrenzalns/matdid/matdid437773.pdf.

[6] J. C. Platt. 1998. **Fast training of support vector machines using sequential minimal optimization.** [Online]. Available : http://research.microsoft.com/pubs/69644/tr-98-14.pdf.

[7] พยุง มีสัจ, วาทินี นุ้ยเพียร, และผุสดี บุญรอด. 2556. **ระบบการค้นคืนเชิงความหมายจากข้อมูลบรรณานุกรม**. JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY. |

[8] C. Sapsanis, G. Georgoulas and A. Tzes, "**EMG based classification of basic hand movements based on time-frequency features**", 21th Mediterranean Conference on Control & Automation (MED).

[9] R.M.G. Tello, T. Bastos-Filho, R.M. Costa†, A.Frizera-Neto, S. Arjunan, D. Kumar, "**Towards sEMG Classification Based on Bayesian and k-NN to Control a Prosthetic Hand**".

[10] Lichman, M. 2013. **UCI Machine Learning Repository**. [Online]. Available : http://archive.ics.uci.edu/ml Irvine, CA: University of California, School of Information and Computer Science.

[11] C. Sapsanis and A. Tzes. 2014. **sEMG for Basic Hand movements Data Set**. [Online]. Available : https://archive.ics.uci.edu/ml/datasets/sEMG+ for+Basic+ Hand+movements.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก
งานวิจัยที่ตีพิมพ์

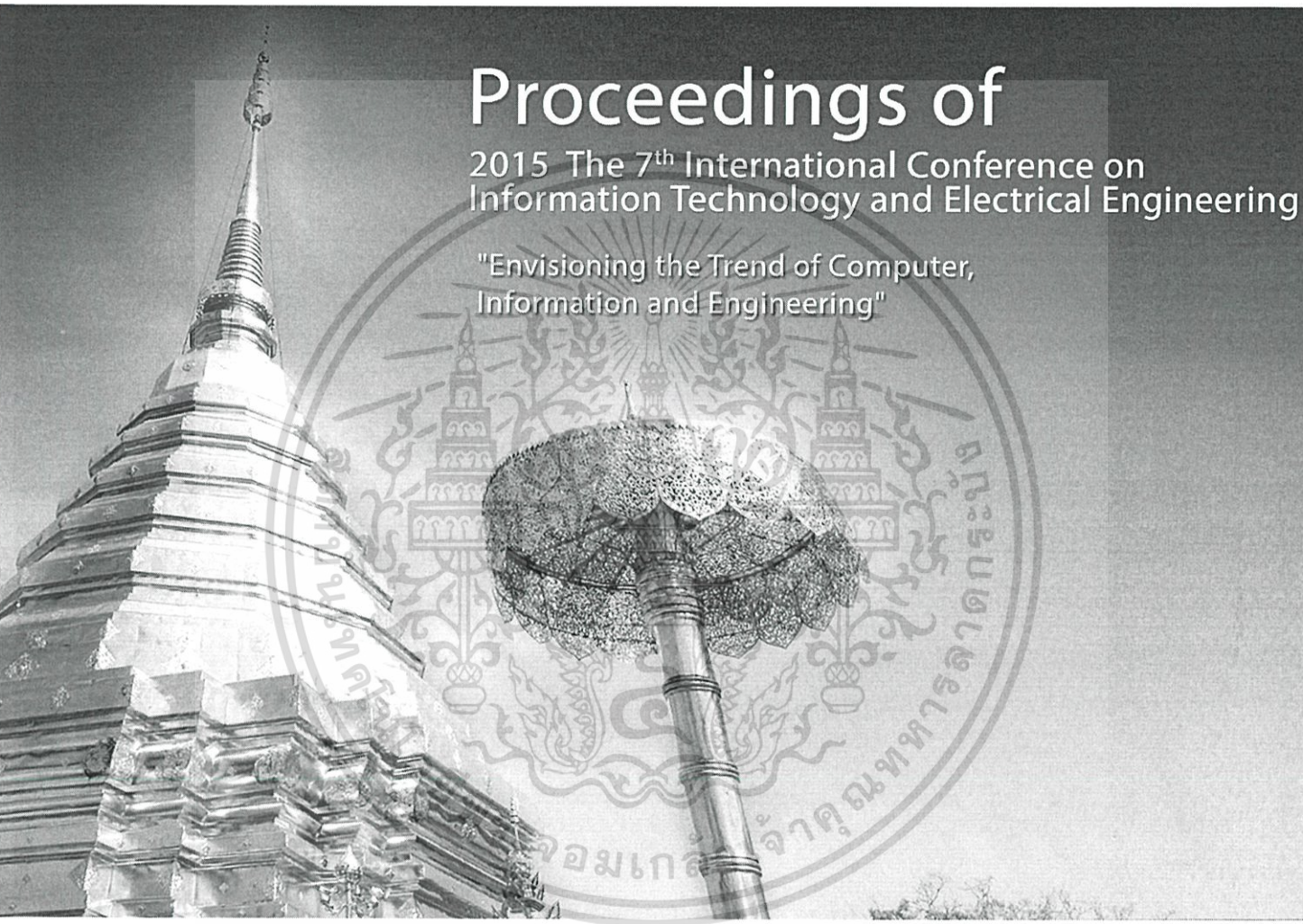
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ISBN: 978-1-4673-7862-8

IEEE
Advancing Technology for Humanity

ICITEE2015

2015 The 7th International Conference on Information Technology and Electrical Engineering



Proceedings of

2015 The 7th International Conference on Information Technology and Electrical Engineering

"Envisioning the Trend of Computer, Information and Engineering"

Le Méridien Chiang Mai Hotel, Thailand

29-30 October 2015

Organized by
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang, THAILAND

Co-organized by
Department of Electrical Engineering and Information Technology
Universitas Gadjah Mada, INDONESIA

เอกสารนี้ใช้เอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่อนุญาตให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

sEMG Signal Classification Using SMO Algorithm and Singular Value Decomposition

Yotsapat Ruangpaisarn
Department of Computer Science, Faculty of Science,
King Mongkut's Institute of Technology Ladkrabang,
Thailand
ru.yotsapat@gmail.com

Saichon Jaiyen
Department of Computer Science, Faculty of Science,
King Mongkut's Institute of Technology Ladkrabang,
Thailand
kjsaicho@kmitl.ac.th

Abstract— **Surface Electromyography (sEMG) signal analysis is a challenging task in neuroscience. The signal is associated with an activity of muscles in Human body. It is a part of how human can control the robotic arm for helping people with disabilities. In this paper, we propose a new method based on Singular Value Decomposition (SVD) and SMO algorithm for classifying sEMG signals into six basic hand movements. By this proposed method, SVD is adopted for feature extraction and SMO classifier is used for classifying sEMG signals into six classes of basic hand movements in five subjects. In preliminary experiment, we investigates the number of features that can yield the best performance in the classification and it is found that the optimal number of features is 50. For performance evaluation, five classifiers including Decision Tree, K-nearest neighbor, Naïve Bayes, RBF, and SMO, with 10 fold cross-validation technique are adopted. The experimental results have shown that SMO algorithm with V2M-SVD feature extraction can achieve the best performance for the classification of basic hand movements.**

Keywords: Feature Extraction, Classification, sEMG, V2M, SVD, SMO.

I. INTRODUCTION

In medical science, brain, nerve, and muscle tissues are more sensitive to stimuli. When the tissues are stimulated by the stimuli, they will generate electrical signals and transmit them. Therefore, measurement and analysis of the electrical signals are beneficial to several kinds of medical treatment and diagnosis. The electrical signals of our body are as follows. Electrooculogram (EOG) is the electric signals originating in the eye, Electrocardiogram (ECG or EKG) is the electrical activity of heart, Electroencephalogram (EEG) is the electrical activity of brain, and Electromyography (EMG) is the electrical activity of muscle.

Since we are interested in the classification of basic movements which is associated with the activity of muscle, EMG signal is focused on this experiment. For related work, Gaoxiang Ouyang [1] presented an adaptive neuro-fuzzy inference system (ANFIS) model for classification 4 hand actions. The model yielded 99.1% of accuracy. Then, Fuzzy Gaussian Mixture Models (FGMMs) were proposed by Zhaojie Ju [2]. The model, which adopts multifeature combining Willison Amplitude and Determinism, has classification accuracy of 10 hand motions at 96.7%. Afarin Nazemi [3] presented windowing method and Multi-Layer Perceptron (MLP) classifiers using different combinations of features in classification 52 movements of wrist, hand and finger, and the model got accuracy rate of 96.34%.

In addition, Richard M. G. Tello [4] proposed bayesian and k-NN for sEMG classification of 15 classes with 94%-95% of accuracy. Sapsanis [5] presented machine learning method for classification the EMG signal. They used Empirical Mode Decomposition (EMD) into Intrinsic Mode Functions (IMFs) in their data set and subsequently used a feature extraction to select 40 features. The method gained an average of accuracy at 89.21%. Li Zhang [6] offered Noise Reduction of sEMG by SVD based on Neural Network. Besides, Nayan M. Kakoty [7] proposed radial basis function (RBF) based on SVM model for classification 6 hand movements of sEMG signals with 97.5% of average recognition rate. M. Khezri1 [8] used wavelet thresholding technique for classifying 3 hand movements and the technique gained an average accuracy at 94%.

II. PROPOSED METHOD

A. sEMG Data set

This research uses sEMG signal data set [9] of basic hand movements from UCI Repository of Machine Learning Databases. The signals were collected at sampling rate of 500 Hz. That was band-pass filtered using a Butterworth Band Pass filter with low and high cutoff at 15Hz and 500Hz respectively and a notch filter at 50Hz to eliminate line interference artifacts. The subjects were asked to perform repeatedly the following six movements, which can be considered as daily hand grasps as follows.

a) Spherical: for holding spherical tools

b) Tip: for holding small tools

c) Palmar: for grasping with palm facing the object

d) Lateral: for holding thin, flat objects

e) Cylindrical: for holding cylindrical tools

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

978-1-4673-7863-5/15/$31.00 ©2015 IEEE

f) Hook: for supporting a heavy load

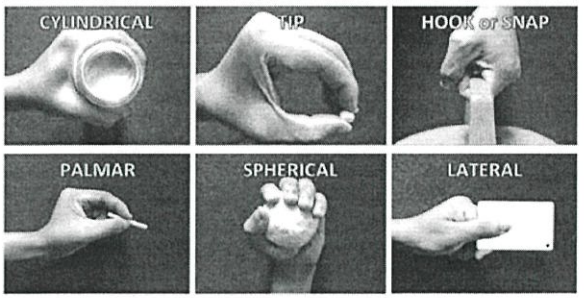


Fig 1. The 6 activities of hand movements.

B. Vector to Martrix (V2M) Transformation

In this paper, we proposed the feature extraction method to reduce redundant features in sEMG signals using SVD. To apply SVD to perform feature extraction, the data must be in the form of matrix. Therefore, sEMG signals must be transformed to matrix form. Firstly, each sEMG signal is transformed to a vector. Secondly, each vector is reshaped to matrix with n columns where n is the positive integer. In other word, each row in the matrix contains n elements cut from the vector. If the number of elements in the last row is less than n, the remaining elements is set to 0.

C. Singular value decomposition (SVD)

The singular value decomposition of a rectangular matrix A is the factorization of A into the product of three matrices which are an orthogonal matrix U, a diagonal matrix S, and the transpose of an orthogonal matrix V. The singular value decomposition of the rectangular matrix A with m rows and n columns can be written as

$$A_{m\times n} = U_{m\times m} S_{m\times n} V^T_{n\times n} \qquad (1)$$

where $U^T U = I$ and $V^T V = I$.

D. Data Perprocessing

sEMG data set used in the experiments contains signals of 2 channels with 5 subjects. Firstly, the raw sEMG data set were preprocessed by concatenating the signal from the first channel and the second channel to generate a single input vector as shown in Fig 2.

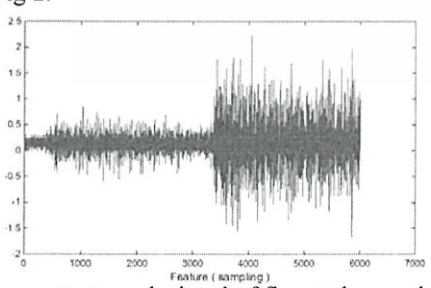


Fig 2. The concatenated signal of first and second channels.

In this paper, we propose the feature extraction method called *V2M-SVD feature extraction* which is used for reducing the redundant features in sEMG signals as shown in Fig 3. Firstly, each signal from the first channel (Ch1) is preprocessed by using SVD and each signal from the second channel (2) is also preprocessed the same way. Finally, the diagonal elements from SVD of first and second channels are concatenated together to form input vectors.

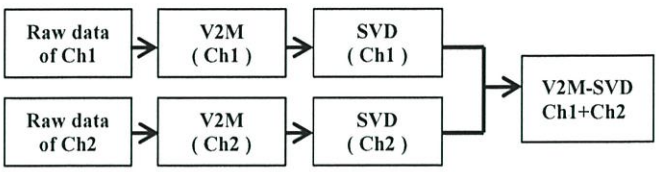


Fig 3. V2M-SVD feature extraction.

E. SMO algorithm

In this paper, we adopt SMO algorithm [10] for training support vector machines in order to classify sEMG signals into 6 classes of basic hand movements in five subjects. The performance of SMO algorithm is evaluated and compared to 4 classifiers including Decision Tree, K-nearest neighbor, Naïve Bayes, and Radial Basis Function Network.

III. Criteria Used for Classification Evaluation

A. Confusion Matrix

A confusion matrix is a table that contains information about actual and predicted classifications done by a classifier. It is used to describe the performance of the classifier on a set of testing data. The performance of the classifier is evaluated using the information in this table. In this paper, the confusion matrix is used to evaluate the performance of the proposed method compared to 4 classifiers. The confusion matrix is demonstrated in Fig 4.

Predicted class / Actual class	Class 1	Class 2
Class 1	True Positive (TP)	False Negative (FN)
Class 2	False Positive (FP)	True Negative (TN)

Fig. 4. The confusion matrix.

From the confusion matrix shown in Fig. 4. The true positive (TP) is the number of data in testing set that are classified as Class 1 and the answer is Class 1. The false negative (FN) is the number of data in testing set that are classified as Class 2 but the answer is Class 1. The false positive (FP) is the number of data in testing set that are

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

classified as Class 1 but the answer is Class 2. The true negative (TN) is the number of data in testing set that are classified as Class 2 and the answer is Class 2.

In the experiments, the values from the confusion matrix is used for evaluating the performance of all classifiers. The performance of all classifiers is evaluated by accuracy, precision, recall, and F-measure described as follows.

B. Accuracy

Accuracy is a measure of the precision of classification models. The accuracy can be calculated by using equation 2.

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \quad (2)$$

C. Precision

Precision is a performance metric used to measure the exactness. The precision can be computed by using equation 3.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

D. Recall

Recall is a performance matric used to measure the completeness. The recall can be calculated by using equation 4.

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

E. F-measure

F-measure is a measurement that combines precision and recall by using the harmonic mean of precision and recall. The traditional F-measure or balanced F-score can be computed by using equation 5.

$$F - measure = 2\,x\,\frac{Precision\,.\,Recall}{Precision + Recall} \quad (5)$$

IV. Experiment Results And Discussion

A. Experimental Result

In this experiment, the performance of SMO classifier is evaluated and compared to 4 classifiers including Decision Tree (C4.5), k-Nearest Neighbor (KNN), Naïve Bayes, and Radial Basis Function Network (RBF). For SMO algorithm, the kernel function is RBF kernel, the gamma values is set as 0.1, and the complexity parameter is set to 1000. For *KNN*, the parameter *k* is set to 3 and the distance function used to finding the neighbors is Euclidean distance. For RBF, the minimum standard deviation for the clusters is set to 10 for data without preprocessing and 100 for data with feature extraction. All classifiers used the settings parameters for the best results.

After classification with SMO, Figure 5 illustrates accuracy values in different feature numbers obtained from V2M-SVD feature extraction. From this figure, the maximum accuracy (100%) of subject 5 is achieved when the numbers of features are 40 and 50. While comparing the accuracy values in 40 and 50 features in subject 1 and subject 3, it is found that the overall performance in 50 features is better than that of 40 features. Therefore, the number of features used in this experiment is 50.

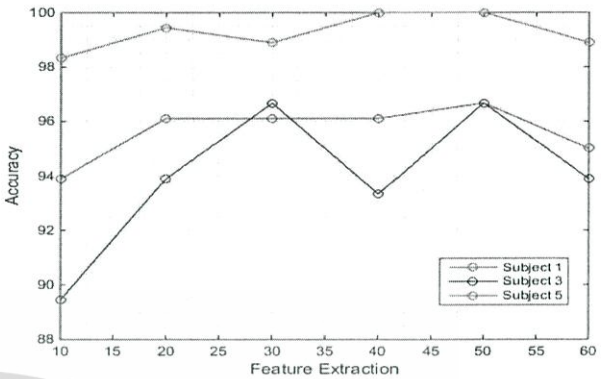


Fig 5. The accuracy of different feature numbers after V2M-SVD feature extraction on SMO classification.

B. Performance

Figure 6 demonstrates comparative results of raw data and V2M-SVD data in 5 subjects after 10-fold cross-validations and classifying with 5 different classifiers. From this figure, when classified by Naïve Bayes and RBF network, raw data obtains over 60% of average accuracy in all subjects, and over 80% of average accuracy in subject 2. Besides, V2M-SVD data with 50 features can achieve 85% of average accuracy in all classifiers, and it can also reach the highest accuracy of classification at 100% in subject 5 by SMO classifier.

The accuracy values obtained from these five classifiers in all subjects are revealed in Table I. The results show that the minimum accuracy (89.44%) is gained by RBF classifier in subject 3, while the maximum accuracy (100%) is achieved by SMO classifier in subject 5.

Additionally, Table II shows the performance evaluation; including precision, recall, and F-measure, of V2M-SVD feature extraction after classifying by Decision Tree, K-nearest neighbor, Naïve Bayes, RBF, and SMO in 5 subjects under 6 basic movements. For Decision Tree, it obtains 100% of all measures, including precision, recall, and F-measures, in tip class of subject 1 and in spherical class of subject 3, while this classifier gains 92.67% of average accuracy. The K-nearest neighbor technique can gain the maximum values of all measures in cylindrical class of subject 4; in hook class of subject 2 and 4; in spherical class of subject 2, 3, 4, and 5, while it gains 94.78% of average accuracy. For Naïve Bayes, it obtains the maximum values of all measures in cylindrical class of subject 4; in hook and spherical class of subject 2, while it obtains 91.67% of average accuracy. In addition, the RBF network gains the maximum values of all measures in hook and spherical classes of subject 2, while this network gains 94.00% of average accuracy. Especially, the results have shown that SMO classifiers can achieve the best performance in classification of basic movement in this experiment. To

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

illustrate this point, SMO can obtains the most maximum values of all measures, i.e. in cylindrical class of subject 3-5; in hook class of subject 2-5; in lateral class of subject 5; in palm class of subject 2 and 5; in spherical class of subject 2-5; and in tip class of subject 1 and 5. Furthermore, it can yield the maximum average accuracy at 98.22%. After consideration of the average accuracy, the classifiers are sorted from high to low performance as follows: SMO; K-nearest neighbor technique; RBF network; Decision Tree; and Naïve Bayes, respectively.

TABLE I. THE ACCURACY OF V2M-SVD FEATURE EXTRACTION FOR BASIC MOVEMENT CLASSIFICATION.

Classifier	Accuracy of Subject				
	S1	S2	S3	S4	S5
Decision Tree	90.00%	91.67%	94.44%	90.56%	96.67%
KNN	93.33%	95.56%	95.00%	97.22%	92.78%
Naïve Bayes	90.00%	93.89%	90.56%	91.67%	92.22%
RBF	90.56%	96.67%	89.44%	96.67%	96.67%
SMO	**96.67%**	**98.89%**	**96.67**	**98.89%**	**100%**

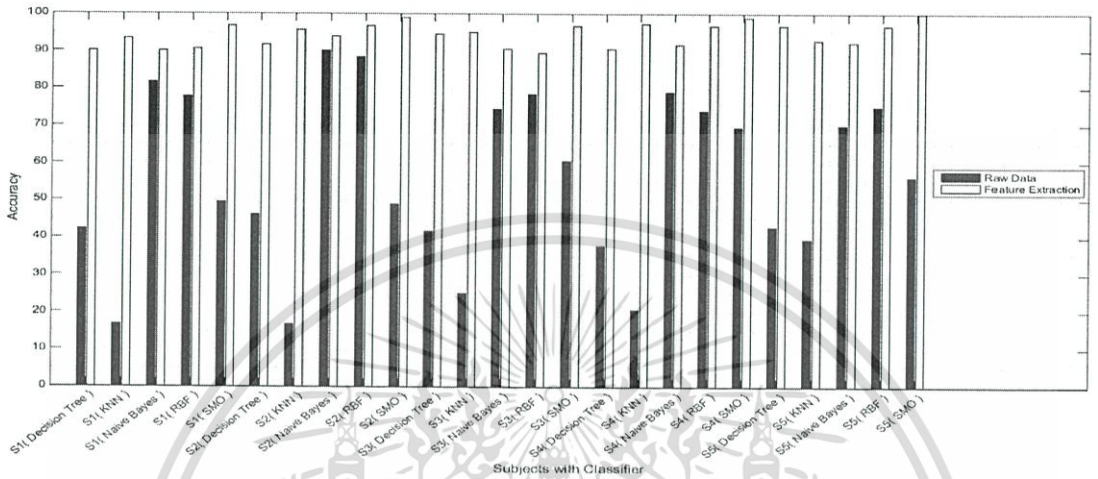


FIG 6. THE COMPARATIVE RESULTS BETWEEN RAW DATA AND V2M-SVD DATA WITH 50 FEATURES AFTER APPLYING TO 5 CLASSIFIERS.

TABLE II. PERFORMANCE EVALUATION OF V2M-SVD FEATURE EXTRACTION FOR BASIC MOVEMENT CLASSIFICATION.

Classifier	Class	Subject														
		S1			S2			S3			S4			S5		
		P. (%)	R. (%)	F. (%)	P. (%)	R. (%)	F. (%)	P. (%)	R. (%)	F. (%)	P. (%)	R. (%)	F. (%)	P. (%)	R. (%)	F. (%)
Decision Tree	Cylindrical	87.1	90.0	88.5	93.3	93.3	93.3	96.7	96.7	96.7	96.6	93.3	94.9	96.7	96.7	96.7
	Hook	96.3	86.7	91.2	96.4	90.0	93.1	96.7	96.7	96.7	96.8	100	98.4	96.7	96.7	96.7
	Lateral	89.3	83.3	86.2	93.1	90.0	91.5	96.2	83.3	89.3	93.3	93.3	93.3	96.7	96.7	96.7
	Palm	83.9	86.7	85.2	87.5	93.3	90.3	93.3	93.3	93.3	82.1	76.7	79.3	93.8	100	96.8
	Spherical	84.8	93.3	88.9	90.6	96.7	93.5	100	100	100	96.8	100	98.4	100	96.7	98.3
	Tip	100	100	100	89.7	86.7	88.1	85.3	96.7	90.6	77.4	80.0	78.7	96.6	93.3	94.9
KNN	Cylindrical	90.3	93.3	91.8	96.8	100	98.4	96.7	96.7	96.7	100	100	100	84.4	90.0	87.1
	Hook	93.5	96.7	95.1	100	100	100	100	96.7	98.3	100	100	100	86.7	86.7	86.7
	Lateral	96.6	93.3	94.9	93.3	93.3	93.3	87.1	90.0	88.5	93.5	96.7	95.1	100	93.3	96.6
	Palm	93.1	90.0	91.5	93.5	96.7	95.1	96.7	96.7	96.7	96.4	90.0	93.1	88.2	100	93.8
	Spherical	96.4	90.0	93.1	100	100	100	100	100	100	100	100	100	100	100	100
	Tip	90.6	96.7	93.5	89.3	83.3	86.2	90.0	90.0	90.0	93.5	96.7	95.1	100	86.7	92.9
Naïve Bayes	Cylindrical	96.6	93.3	94.9	93.8	100	96.8	96.6	93.3	94.9	100	100	100	100	66.7	80.0
	Hook	88.2	100	93.8	100	100	100	100	96.7	98.3	96.8	100	98.4	75.0	100	85.7
	Lateral	88.9	80.0	84.2	87.9	96.7	92.1	84.0	70.0	76.4	84.8	93.3	88.9	100	96.7	98.3
	Palm	84.4	90.0	87.1	87.9	96.7	92.1	92.9	86.7	89.7	83.3	83.3	83.3	90.9	100	95.2
	Spherical	96.4	90.0	93.1	100	100	100	93.8	100	96.8	96.7	96.7	96.7	96.8	100	98.4
	Tip	86.7	86.7	86.7	95.5	70.0	80.8	78.4	96.7	86.6	88.5	76.7	82.1	100	90.0	94.7
RBF	Cylindrical	93.5	96.7	95.1	96.7	96.7	96.7	100	93.3	96.6	93.3	93.3	93.3	93.3	93.3	93.3
	Hook	93.5	96.7	95.1	100	100	100	93.5	96.7	95.1	100	93.3	96.6	100	93.3	96.6
	Lateral	89.3	83.3	86.2	96.6	93.3	94.9	77.8	70.0	73.7	96.8	100	98.4	96.8	100	98.4
	Palm	81.8	90.0	85.7	96.8	100	98.4	96.6	93.3	94.9	100	96.7	98.3	100	96.7	98.3
	Spherical	96.6	93.3	94.9	100	100	100	96.8	100	98.4	90.9	100	95.2	90.9	100	95.2
	Tip	89.3	83.3	86.2	90.0	90.0	90.0	73.5	83.3	78.1	100	96.7	98.3	100	96.7	98.3
SMO	Cylindrical	96.7	96.7	96.7	96.8	100	98.4	100	100	100	100	100	100	100	100	100
	Hook	96.7	96.7	96.7	100	100	100	100	100	100	100	100	100	100	100	100
	Lateral	96.7	96.7	96.7	96.8	100	98.4	90.6	96.7	93.5	100	96.7	98.3	100	100	100
	Palm	96.7	96.7	96.7	100	100	100	96.7	96.7	96.7	96.8	100	98.4	100	100	100
	Spherical	93.3	93.3	93.3	100	100	100	100	100	100	100	100	100	100	100	100
	Tip	100	100	100	100	93.3	96.6	92.9	86.7	89.7	96.7	96.7	96.7	100	100	100

P. = Precision, R. = Recall, F. = F-measure.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

V. Conclusion

In this research, we present the new feature extraction technique, which is called Vector to Matrix - Singular value decomposition (V2M-SVD), for reducing redundant features. In addition, we present the method for classifying six basic movements from Surface Electromyography using SMO classifier. In preliminary results, V2M-SVD with 50 features yielded the best results, so this set of features is subsequently used for further experiment. This feature selection can reduce the original features from 6000 features to the efficient set of 50 features, then, it results in 120% of data reduction and also relieves a time complexity. For performance evaluation, we used 10 fold cross-validation with 5 classifiers, which are Decision Tree, K-nearest neighbor technique, Naïve Bayes, RBF neural network, and SMO. The experimental results show that all classifiers can yield greater than 90% of average accuracy and SMO can achieve the best performance in classifying the basic movements.

References

[1] G. Ouyang, X. Zhu, and Z. Ju, "Dynamical Characteristics of Surface EMG Signals of Hand Grasps via Recurrence Plot", IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, 2014.

[2] Z. Ju, and H. Liu, "Human Hand Motion Analysis With Multisensory Information", IEEE/ASME TRANSACTIONS ON MECHATRONICS, 2014.

[3] A. Nazemi, and A. Maleki, "Artificial Neural Network Classifier in Comparison with LDA and LS-SVM Classifiers to Recognize 52 Hand Postures and Movements", 4th International Conference on Computer and Knowledge Engineering (ICCKE), 2014.

[4] R.M. G. Tello, T. Bastos-Filho, R. M. Costa, and A. Frizera-Neto, "Towards sEMG Classification Based on Bayesian and k-NN to Control a Prosthetic Hand", Biosignals and Biorobotics Conference (BRC), 2013.

[5] C. Sapsanis, G. Georgoulas, A. Tzes, and D. Lymberopoulos, "Improving EMG based Classification of basic hand movements using EMD," IEEE Engineering in Medicine and Biology Society, 2013.

[6] L. Zhang, Y. Li1, Z. Xu1, and Y. Tian, "Noise Reduction of sEMG by SVD Based on Neural Network", 10th World Congress on Intelligent Control and Automation, 2012.

[7] N. M. Kakoty, and S. M. Hazarika, "Recognition of Grasp Types through Principal Components of DWT based EMG Features", IEEE International Conference on Rehabilitation Robotics, 2011.

[8] M. Khezri, and M. Jahed, "Surface Electromyogram Signal Estimation Based on Wavelet Thresholding Technique", 30th Annual International IEEE EMBS Conference, 2008.

[9] C. Sapsanis, A. Tzes and G. Georgoulas, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml] , 2013.

[10] John C. Platt, "Fast Training of Support Vector Machines using Sequential Minimal Optimization", Microsoft Research , 1998.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประวัติผู้เขียน

ชื่อ	นายยศภัทร เรืองไพศาล
วัน เดือน ปีเกิด	10 มีนาคม 2531
ที่อยู่ปัจจุบัน	148/62 หมู่ 5 ตำบลห้วยกะปิ อำเภอชลบุรี จังหวัดชลบุรี 20130
ประวัติการศึกษา	2552 วิทยาศาสตรบัณฑิต สาขาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ย 3.53 มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก วิทยาเขตบางพระ
	2558 วิทยาศาสตรมหาบัณฑิต สาขาวิทยาการคอมพิวเตอร์ เกรดเฉลี่ย 3.56 สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ผลงานทางวิชาการ	1.“ sEMG Signal Classification Using SMO Algorithm and Singular Value Decomposition ”, 2015 7th International Conference on Information Technology and Electrical Engineering: ICITEE.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้