

การเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ความถดถอย
ร่วมกับแนวคิดของตัวอธิบายรูปร่าง

น.ส.ศิลัมพา อานุภาพแสนยากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the Graduate School.

AN ENHANCED TIME SERIES CLASSIFICATION USING LINEAR REGRESSION BASED
SHAPE DESCRIPTOR

Miss Silumpa Arnupapsanyakorn

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2018
Copyright of Chulalongkorn University



425352952

CU ThesIs 6070323921 thesis / recv: 06082562 14:42:31 / seq: 83

หัวข้อวิทยานิพนธ์	การเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง
โดย	น.ส.ศิลัมพา อานุกาพแสนยากร
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	รองศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้หัวข้อวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(อาจารย์ ดร.ดวงดาว วิชาตากุล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(รองศาสตราจารย์ ดร.โชติรัตน์ รัตนามหัทธนะ)

..... กรรมการภายนอกมหาวิทยาลัย
(ดร.เหมวรรณ ศิวรักษ์)

ศิลัมพา อานุภาพแสนยากร : การเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง. (AN ENHANCED TIME SERIES CLASSIFICATION USING LINEAR REGRESSION BASED SHAPE DESCRIPTOR) อ.ที่ปรึกษาหลัก : รศ. ดร.โชติรัตน์ รัตนามัทธนะ

การจำแนกประเภทอนุกรมเวลา (Time Series Classification-TSC) เป็นวิธีที่ใช้กันอย่างแพร่หลาย และเป็นที่น่าสนใจในหลาย ๆ ด้าน ซึ่งการจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด (Nearest Neighbor-NN) ร่วมกับการวัดระยะทางแบบไดนามิกไทม์วอร์ปิง (Dynamic Time Warping-DTW) ถือเป็นหนึ่งในวิธีที่มีประสิทธิภาพและได้รับความนิยมสำหรับจัดการกับปัญหาการจำแนกประเภทอนุกรมเวลา ด้วยการวัดระยะทางที่สามารถจับคู่จุดที่ลำดับต่างกันได้ตามเงื่อนไขที่กำหนด ทำให้มีประสิทธิภาพในการจับคู่อนุกรมเวลาสองชุดร่วมกับตัวจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด อย่างไรก็ตาม แม้ไดนามิกไทม์วอร์ปิงจะสามารถค้นหาการปรับแนวที่เหมาะสมที่สุดระหว่างสองข้อมูลอนุกรมเวลาได้ แต่วิธีการดังกล่าวอาจไม่สามารถใช้ในการปรับแนวเฉพาะส่วนได้อย่างสมเหตุสมผล ในงานวิจัยนี้จึงได้นำเอากรอบการทำงานของไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shape Dynamic Time Warping-shapeDTW) มาใช้เพื่อจำแนกข้อมูลอนุกรมเวลาภายใต้การทำงานของตัวอธิบายรูปร่างแบบฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์แบบ 1 มิติ (Histogram of Oriented Gradients 1D-HOG-1D) โดยพิจารณาโครงสร้างเฉพาะส่วนมาใช้ในขั้นตอนการปรับแนว โดยโครงสร้างที่มีรูปร่างคล้ายคลึงกันจะจับคู่กันโดยอิงตามระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่างที่ถูกสร้างจากตัวอธิบายรูปร่าง ซึ่งผู้ใช้สามารถออกแบบตัวอธิบายรูปร่างโดยอิงตามลักษณะและขอบเขตของข้อมูล ในงานวิจัยนี้จึงได้นำเสนอตัวอธิบายรูปร่างที่เรียกว่า HOG1D-L ซึ่งอิงจากแนวคิดสำคัญจาก HOG-1D และการวิเคราะห์ถดถอยเชิงเส้นมาใช้ในการพัฒนาตัวอธิบายรูปร่าง โดยงานวิจัยนี้ได้ทำการทดลองจากชุดข้อมูลอนุกรมเวลา UCR Time Series Classification Archive จำนวน 78 ชุดข้อมูลและผลที่ได้แสดงให้เห็นว่าความแม่นยำหรือประสิทธิภาพในการจำแนกประเภทข้อมูลมีความแม่นยำมากขึ้นจากการทดลองกับชุดข้อมูลที่หลากหลาย

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ลายมือชื่อนิสิต

ปีการศึกษา 2561

ลายมือชื่อ อ.ที่ปรึกษาหลัก

6070323921 : MAJOR COMPUTER SCIENCE

KEYWORD: Time series classification shape descriptor Dynamic Time Warping
Histogram of Oriented Gradient

Silumpa Arnupapsanyakorn : AN ENHANCED TIME SERIES CLASSIFICATION
USING LINEAR REGRESSION BASED SHAPE DESCRIPTOR. Advisor: Assoc. Prof.
Dr. CHOTIRAT RATANAMAHATANA

Time series classification (TSC) has been used extensively with a wide range in many fields of real-world applications. Nearest Neighbor (NN) with Dynamic Time Warping (DTW) is regarded as one of the most effective and popular method used for a time series classification problem. Being a non-linear-alignment distance measure, DTW with NN classifier is an effective tool for matching two time series sequences. However, the fact that DTW is specifically designed to discover an optimal alignment, it may not be able to achieve sensible local alignment. In this work, the shapeDTW framework was adopted in order to classify time series data based on the state-of-the-art HOG-1D shape descriptor. Point-wise local structures was utilized in the alignment process, similarly-shaped structures are matched based on their euclidean distance between descriptor subsequences which were built by a shape descriptor. Which allows users to design the shape descriptors based on the data characteristics and domains. An enhanced shape descriptor representation called HOG1D-L was proposed based on two key concepts of the state-of-the-art HOG-1D descriptor and linear regression. The proposed work on 78 UCR Time Series Classification Archive datasets was extensively tested and the results demonstrate that our approach can maintain or achieve better classification accuracy in most of the datasets.

Field of Study: Computer Science

Student's Signature

Academic Year: 2018

Advisor's Signature

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงด้วยดีด้วยความกรุณาและความช่วยเหลืออย่างสูงยิ่งจากผู้เกี่ยวข้องทุกท่านที่ให้คำแนะนำช่วยเหลืออย่างดียิ่งในการดำเนินการวิจัยมาโดยตลอด ด้วยการสนับสนุนและกำลังใจจากทุกๆฝ่ายทำให้สามารถก้าวผ่านช่วงเวลาอันสำคัญแห่งชีวิตมาได้ ผู้วิจัยรู้สึกซาบซึ้งในความกรุณาและขอกราบขอบพระคุณอย่างสูงไว้ ณ โอกาสนี้

ขอขอบพระคุณ รองศาสตราจารย์ ดร. โชติรัตน์ รัตนมัทธนะ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้อุทิศเวลาอันมีค่า ให้คำปรึกษาแนะนำรวมทั้งแนวคิดในการทำวิทยานิพนธ์ ตลอดจนเอาใจใส่ตรวจแก้ไขข้อบกพร่องต่างๆ ทำให้วิทยานิพนธ์มีความสมบูรณ์และคุณค่ายิ่งขึ้น อีกทั้งนอกจากจะคอยให้คำปรึกษาทั้งด้านวิชาการ งานวิจัยและตรวจทานวิทยานิพนธ์แล้วยังคอยให้คำปรึกษาด้านทักษะการใช้ชีวิต และยังเป็นตัวอย่างที่ดีให้แก่ผู้วิจัยในด้านการทำงานและการแก้ไขปัญหาโดยตลอด

ขอขอบพระคุณคณะกรรมการสอบวิทยานิพนธ์ ได้แก่ ดร.ดวงดาว วิชิตากุล ที่ให้ความกรุณาเป็นประธานกรรมการ และ ดร.เหมวรรณ ศิวรักษ์ กรรมการสอบวิทยานิพนธ์ ที่ท่านให้ความเมตตาพร้อมทั้งชี้แนะแนวทางในการปรับปรุงวิทยานิพนธ์ให้มีประสิทธิภาพมากขึ้น

ขอขอบคุณพี่ และเพื่อนๆทุกคน ที่คอยให้ความช่วยเหลือและให้กำลังใจตลอดจนคอยอำนวยความสะดวก ให้คำแนะนำด้วยอภัยยศยไมตรี เป็นกัลยามิตรที่ดียิ่งตลอดระยะเวลาในการศึกษาและการจัดทำวิทยานิพนธ์

ขอขอบคุณครอบครัวที่คอยสนับสนุนผลักดันและเป็นแรงใจสำคัญที่ทำให้ผ่านปัญหาอุปสรรคมาสู่ความสำเร็จมีชีวิตที่งดงามได้ด้วยใจที่เข้มแข็ง

คุณค่าและประโยชน์อันพึงมีจากวิทยานิพนธ์ฉบับนี้ ผู้วิจัยขอมอบเป็นเครื่องสักการะบูชา พระคุณบิดามารดา ครูอาจารย์ ตลอดจนผู้มีพระคุณทุกท่านที่ให้การอบรมสั่งสอนประสิทธิ์ประสาทวิชา ส่งเสริมสนับสนุนและชี้แนะแนวทางการศึกษาจนทำให้ผู้วิจัยก้าวมาสู่ความสำเร็จในครั้งนี้

ศิลัมพา อานุภาพแสนยากร

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญรูป.....	ฅ
สารบัญตาราง.....	ฒ
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 ขอบเขตของงานวิจัย.....	3
1.4 ประโยชน์ที่ได้รับจากงานวิจัย.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	5
2.1 ทฤษฎีที่เกี่ยวข้อง.....	5
2.1.1 ข้อมูลอนุกรมเวลา (Time Series Data).....	5
2.1.2 ลำดับย่อยของอนุกรมเวลา (Time series subsequence).....	6
2.1.3 การจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด k ตัว (k -Nearest Neighbor Classification).....	7
2.1.4 ระยะทางยูคลิด (Euclidean Distance).....	7
2.1.5 ไดนามิกไทม์วอร์ปิง (Dynamic Time Warping-DTW).....	8
2.1.6 ฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradients-HOG).....	10
2.1.7 การวิเคราะห์ความถดถอยเชิงเส้น (Linear Regression Analysis).....	13

2.2 งานวิจัยที่เกี่ยวข้อง.....	14
2.2.1 ไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shape Dynamic Time Warping-shapeDTW) [4].....	14
2.2.2 งานวิจัยที่เกี่ยวข้องกับตัวอธิบายรูปร่าง	23
2.2.2.1 การแบ่งลำดับย่อยของอนุกรมเวลาแบบ Raw-subsequence	23
2.2.2.2 การแปลงเวฟเล็ตแบบไม่ต่อเนื่อง (Discrete Wavelet Transform-DWT)..	24
2.2.2.3 Piecewise Aggregate Approximation (PAA)	24
2.2.2.4 ฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์แบบ 1 มิติ (Histogram of Oriented Gradients 1D-HOG-1D).....	26
บทที่ 3 แนวคิดและวิธีการดำเนินงาน.....	28
3.1 การเตรียมข้อมูล	28
3.2 การเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง	29
3.2.1 การแบ่งลำดับย่อยออกเป็นช่วง	31
3.2.2 การแบ่งช่วงของลำดับย่อยออกเป็น <i>nLevel</i>	31
3.2.3 การหาแนวเส้นความถดถอยเชิงเส้น	32
3.2.4 การหาขนาดและทิศทางของเกรเดียนท์	33
3.2.5 เปรียบเทียบแนวเส้นความถดถอยของการวิเคราะห์ความถดถอย	34
3.2.6 เกรเดียนท์โทวัต	35
3.2.7 การนำค่าในฮิสโตแกรมของเกรเดียนท์มาเรียงต่อกัน	38
3.2.7 อัลกอริทึมการเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้ หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง	40
บทที่ 4 การทดลองและการวิเคราะห์ผล.....	46
4.1 ชุดข้อมูลที่ใช้ในการทดลอง	46
4.2 การวัดและประเมินผล	49

4.3 การทดลองเพื่อหาพารามิเตอร์ <i>nLevel</i> ที่ดีที่สุด.....	50
4.4 ผลการทดลอง.....	56
บทที่ 5 สรุปผลงานวิจัยและข้อเสนอแนะ.....	65
5.1 สรุปผลงานวิจัย.....	65
5.2 ข้อเสนอแนะ.....	66
บรรณานุกรม.....	67
ภาคผนวก.....	69
ภาคผนวก ก.....	70
ประวัติผู้เขียน.....	122



425352952

สารบัญรูป

หน้า

รูปที่ 1.1	แสดงการปรับแนวของ 2 อนุกรมเวลา (ก) การปรับแนวของไดนามิกไทม์วอร์ปปีง (ข) การปรับแนวของไดนามิกไทม์วอร์ปปีงแบบรูปร่าง	2
รูปที่ 1.2	อธิบายการทำงานของตัวอธิบายรูปร่าง HOG-1D ของลำดับย่อยอนุกรมเวลา	2
รูปที่ 2.1	กราฟแสดงตัวอย่างอนุกรมเวลาของชุดข้อมูล coffee จาก UCR Time Series Classification Archive [6].....	5
รูปที่ 2.2	กราฟแสดงตัวอย่างลำดับย่อยของอนุกรมเวลาของชุดข้อมูล coffee (เส้นประ) จาก UCR Time Series Classification Archive [6]	6
รูปที่ 2.3	แนวคิดการหาลำดับย่อยของอนุกรมเวลาด้วยเทคนิค Sliding Window	6
รูปที่ 2.4	การจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว.....	7
รูปที่ 2.5	การวัดระยะทางยุคลิดของ 2 อนุกรมเวลา.....	8
รูปที่ 2.6	การปรับแนวของ 2 อนุกรมเวลาด้วยไดนามิกไทม์วอร์ปปีง	10
รูปที่ 2.7	ขั้นตอนการทำงานของฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ในการตรวจจับภาพ บุคคล.....	10
รูปที่ 2.8	ขั้นตอนทำงานวิธีฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์.....	11
รูปที่ 2.9	ถังที่เก็บทิศทางของฮิสโตแกรมขนาด 0 ถึง 180 องศา.....	12
รูปที่ 2.10	ตัวอย่างการวิเคราะห์ความถดถอยเชิงเส้น	13
รูปที่ 2.11	ขั้นตอนการทำงานของไดนามิกไทม์วอร์ปปีงแบบรูปร่าง (ก) อนุกรมเวลาที่ป้อนเข้า (ข) แบ่งลำดับย่อยของอนุกรมเวลา (ค) คำนวณตัวอธิบายรูปร่าง (ง) แสดงลำดับ ย่อยของตัวอธิบายรูปร่าง (จ) ปรับแนวลำดับย่อยของตัวอธิบายรูปร่างด้วยไดนามิก ไทม์วอร์ปปีง (ฉ) ส่งผ่านการปรับแนวไปยังอนุกรมเวลาดั้งเดิม.....	15
รูปที่ 2.12	แผนภูมิแสดงขั้นตอนการหาลำดับย่อยของตัวอธิบายรูปร่างภายใต้กรอบการทำงาน ของไดนามิกไทม์วอร์ปปีงแบบรูปร่าง.....	22
รูปที่ 2.13	ตัวอย่างวิธีการที่นำมาใช้ในการอธิบายรูปร่างของอนุกรมเวลา	23
รูปที่ 2.14	ตัวอย่างการทำงานการแปลงเวฟเล็ตแบบไม่ต่อเนื่อง.....	24

รูปที่ 2.15 ตัวอย่างการลดมิติของข้อมูลด้วยวิธี PAA..... 25

รูปที่ 3.1 แผนภูมิแสดงขั้นตอนการทำงานของตัวอธิบายรูปร่าง HOG1D-L 30

รูปที่ 3.2 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรม
เวลาที่มีความยาว 30 31

รูปที่ 3.3 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรม
เวลาที่มีความยาว 30 เมื่อทำการแบ่งจุดในช่วงต่าง ๆ ของเวลาจากส่วนที่ไม่ทับซ้อน
กันออกเป็น 2 ส่วน 32

รูปที่ 3.4 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรม
เวลาที่มีความยาว 30 ในการคำนวณหาแนวเส้นความถดถอยเชิงเส้น จากจุดต่าง ๆ
ของเวลาในแต่ละช่วง 33

รูปที่ 3.5 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรม
เวลาที่มีความยาว 30 ที่แสดงทิศทางการวางตัวของค่าเกรเดียนท์ในแต่ละจุดต่าง ๆ 34

รูปที่ 3.6 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรม
เวลาที่มีความยาว 30 ที่มีการเปรียบเทียบทิศทางของแนวเส้นความถดถอยเชิงเส้น
ในช่วงนั้น 34

รูปที่ 3.7 แสดงขนาดและทิศทางของเกรเดียนท์ และเส้นความถดถอยเชิงเส้นภายใต้ส่วนทั้ง 2
ส่วน..... 35

รูปที่ 3.8 ถังเก็บทิศทางของฮิสโตแกรมขนาด -90 ถึง 90 องศา..... 36

รูปที่ 3.9 ตัวอย่างฮิสโตแกรมของเกรเดียนท์ 37

รูปที่ 3.10 ตัวอย่างการคำนวณเกรเดียนท์โหวตของจุดที่ 5 และจุดที่ 7 38

รูปที่ 3.11 ตัวอย่างแท่งฮิสโตแกรมที่เก็บสะสมทิศทางตามค่าเกรเดียนท์ของจุดลำดับย่อย 39

รูปที่ 3.12 ลำดับย่อยของตัวอธิบายรูปร่าง..... 39

รูปที่ 4.1 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล DisphalanxOutlineCorrect 55

รูปที่ 4.2 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Cricket_Z..... 56

รูปที่ 4.3 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล ToeSegmentation1 56

รูปที่ 4.4 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล 50words 62

รูปที่ 4.5 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล FaceUCR 62

รูปที่ 4.6 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Adiac..... 63

รูปที่ 4.7 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Strawberry 63

รูปที่ 4.8 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Computers 64

รูปที่ ก.1 กราฟแสดงคลาสของชุดข้อมูล 50Words..... 73

รูปที่ ก.2 กราฟแสดงคลาสของชุดข้อมูล Adiac..... 76

รูปที่ ก.3 กราฟแสดงคลาสของชุดข้อมูล ArrowHead..... 76

รูปที่ ก.4 กราฟแสดงคลาสของชุดข้อมูล Beef..... 77

รูปที่ ก.5 กราฟแสดงคลาสของชุดข้อมูล BeetleFly..... 77

รูปที่ ก.6 กราฟแสดงคลาสของชุดข้อมูล BirdChicken..... 77

รูปที่ ก.7 กราฟแสดงในแต่ละคลาสของชุดข้อมูล Car..... 78

รูปที่ ก.8 กราฟแสดงคลาสของชุดข้อมูล CBF..... 78

รูปที่ ก.9 กราฟแสดงคลาสของชุดข้อมูล ChlorineConcentration..... 79

รูปที่ ก.10 กราฟแสดงคลาสของชุดข้อมูล CinC_ECG_torso..... 79

รูปที่ ก.11 กราฟแสดงคลาสของชุดข้อมูล Coffee..... 80

รูปที่ ก.12 กราฟแสดงคลาสของชุดข้อมูล Computers 80

รูปที่ ก.13 กราฟแสดงคลาสของชุดข้อมูล Cricket_X..... 81

รูปที่ ก.14 กราฟแสดงคลาสของชุดข้อมูล Cricket_Y..... 82

รูปที่ ก.15 กราฟแสดงคลาสของชุดข้อมูล Cricket_Z..... 83

รูปที่ ก.16 กราฟแสดงคลาสของชุดข้อมูล DiatomSizeReduction..... 84

รูปที่ ก.17 กราฟแสดงคลาสของชุดข้อมูล DistalPhalanxOutlineAgeGroup 84

รูปที่ ก.18 กราฟแสดงคลาสของชุดข้อมูล DistalPhalanxOutlineCorrect..... 84

รูปที่ ก.19 กราฟแสดงคลาสของชุดข้อมูล DistalPhalanxTW 85

รูปที่ ก.20 กราฟแสดงคลาสของชุดข้อมูล Earthquakes 85

รูปที่ ก.21 กราฟแสดงคลาสของชุดข้อมูล ECG200..... 86

รูปที่ ก.22 กราฟแสดงคลาสของชุดข้อมูล ECG5000..... 86

รูปที่ ก.23 กราฟแสดงคลาสของชุดข้อมูล ECGFiveDays..... 87

รูปที่ ก.24 กราฟแสดงคลาสของชุดข้อมูล Face (all)..... 88

รูปที่ ก.25 กราฟแสดงคลาสของชุดข้อมูล Face (four)..... 88

รูปที่ ก.26 กราฟแสดงคลาสของชุดข้อมูล FacesUCR..... 90

รูปที่ ก.27 กราฟแสดงคลาสของชุดข้อมูล Fish 90

รูปที่ ก.28 กราฟแสดงคลาสของชุดข้อมูล FordA..... 91

รูปที่ ก.29 กราฟแสดงคลาสของชุดข้อมูล FordB..... 91

รูปที่ ก.30 กราฟแสดงคลาสของชุดข้อมูล Gun-Point 91

รูปที่ ก.31 กราฟแสดงคลาสของชุดข้อมูล Ham 92

รูปที่ ก.32 กราฟแสดงคลาสของชุดข้อมูล Haptics 92

รูปที่ ก.33 กราฟแสดงคลาสของชุดข้อมูล Herring..... 93

รูปที่ ก.34 กราฟแสดงคลาสของชุดข้อมูล InlineSkate..... 93

รูปที่ ก.35 กราฟแสดงคลาสของชุดข้อมูล InsectWingbeatSound 94

รูปที่ ก.36 กราฟแสดงคลาสของชุดข้อมูล ItalyPowerDemand 95

รูปที่ ก.37 กราฟแสดงคลาสของชุดข้อมูล LargeKitchenAppliances 95

รูปที่ ก.38 กราฟแสดงคลาสของชุดข้อมูล Lighting2 95

รูปที่ ก.39 กราฟแสดงคลาสของชุดข้อมูล Lighting7 96

รูปที่ ก.40 กราฟแสดงคลาสของชุดข้อมูล MALLAT..... 97

รูปที่ ก.41 กราฟแสดงคลาสของชุดข้อมูล Meat 97

รูปที่ ก.42 กราฟแสดงคลาสของชุดข้อมูล MedicalImages..... 98

รูปที่ ก.43 กราฟแสดงคลาสของชุดข้อมูล MiddlePhalanxOutlineAgeGroup..... 98

รูปที่ ก.44 กราฟแสดงคลาสของชุดข้อมูล MiddlePhalanxOutlineCorrect 99

รูปที่ ก.45 กราฟแสดงคลาสของชุดข้อมูล MiddlePhalanxTW 99

รูปที่ ก.46 กราฟแสดงคลาสของชุดข้อมูล MoteStrain 100

รูปที่ ก.47 กราฟแสดงคลาสของชุดข้อมูล OliveOil 100

รูปที่ ก.48 กราฟแสดงคลาสของชุดข้อมูล OSU Leaf..... 101

รูปที่ ก.49 กราฟแสดงคลาสของชุดข้อมูล PhalangesOutlinesCorrect..... 101

รูปที่ ก.50 กราฟแสดงคลาสของชุดข้อมูล Plane 102

รูปที่ ก.51 กราฟแสดงคลาสของชุดข้อมูล ProximalPhalanxOutlineAgeGroup 102

รูปที่ ก.52 กราฟแสดงคลาสของชุดข้อมูล ProximalPhalanxOutlineCorrect..... 103

รูปที่ ก.53 กราฟแสดงคลาสของชุดข้อมูล ProximalPhalanxTW..... 103

รูปที่ ก.54 กราฟแสดงคลาสของชุดข้อมูล RefrigerationDevices 104

รูปที่ ก.55 กราฟแสดงคลาสของชุดข้อมูล ScreenType 104

รูปที่ ก.56 กราฟแสดงคลาสของชุดข้อมูล ShapeletSim..... 104

รูปที่ ก.57 กราฟแสดงตัวอย่างคลาสของชุดข้อมูล ShapesAll..... 109

รูปที่ ก.58 กราฟแสดงคลาสของชุดข้อมูล SmallKitchenAppliances..... 109

รูปที่ ก.59 กราฟแสดงคลาสของชุดข้อมูล SonyAIBORobot Surface 109

รูปที่ ก.60 กราฟแสดงคลาสของชุดข้อมูล SonyAIBORobot Surfacell 110

รูปที่ ก.61 กราฟแสดงคลาสของชุดข้อมูล Strawberry 110

รูปที่ ก.62 กราฟแสดงคลาสของชุดข้อมูล Swedish Leaf..... 111

รูปที่ ก.63 กราฟแสดงคลาสของชุดข้อมูล Symbols..... 112

รูปที่ ก.64 กราฟแสดงคลาสของชุดข้อมูล Synthetic Control..... 112

รูปที่ ก.65 กราฟแสดงคลาสของชุดข้อมูล ToeSegmentation1 113

รูปที่ ก.66 กราฟแสดงคลาสของชุดข้อมูล ToeSegmentation2 113

รูปที่ ก.67 กราฟแสดงคลาสของชุดข้อมูล Trace..... 114

รูปที่ ก.68 กราฟแสดงคลาสของชุดข้อมูล Two Patterns 114



425352952

รูปที่ ก.69	กราฟแสดงคลาสของชุดข้อมูล TwoLeadECG.....	115
รูปที่ ก.70	กราฟแสดงคลาสของชุดข้อมูล uWaveGestureLibrary_X	115
รูปที่ ก.71	กราฟแสดงคลาสของชุดข้อมูล uWaveGestureLibrary_Y.....	116
รูปที่ ก.72	กราฟแสดงคลาสของชุดข้อมูล uWaveGestureLibrary_Z	117
รูปที่ ก.73	กราฟแสดงคลาสของชุดข้อมูล Wafer.....	117
รูปที่ ก.74	กราฟแสดงคลาสของชุดข้อมูล Wine	117
รูปที่ ก.75	กราฟแสดงคลาสของชุดข้อมูล WordSynonyms.....	119
รูปที่ ก.76	กราฟแสดงคลาสของชุดข้อมูล Worms.....	120
รูปที่ ก.77	กราฟแสดงคลาสของชุดข้อมูล WormsTwoClass.....	120
รูปที่ ก.78	กราฟแสดงคลาสของชุดข้อมูล Yoga.....	121



425352952

สารบัญตาราง

	หน้า
ตารางที่ 2.1 อัลกอริทึมไดนามิกไทม์วอร์ปิงแบบรูปร่าง	17
ตารางที่ 2.2 อัลกอริทึมสร้างกลุ่มตัวอย่างของลำดับย่อย.....	18
ตารางที่ 2.3 อัลกอริทึมหาระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่างด้วยวิธีการหาระยะทาง แบบยุคลิด	20
ตารางที่ 2.4 อัลกอริทึมใช้ในการปรับแนวของ 2 อนุกรมเวลาด้วยไดนามิกไทม์วอร์ปิง.....	20
ตารางที่ 3.1 อัลกอริทึม HOG1D-L.....	40
ตารางที่ 3.2 อัลกอริทึม histogramVoting	43
ตารางที่ 4.1 ตารางแสดงรายละเอียดของชุดข้อมูลที่ใช้ในการทดลอง	44
ตารางที่ 4.2 แสดงการเปรียบเทียบความแม่นยำจากการจำแนกประเภทข้อมูลโดยวิธีการเพื่อน บ้านที่ใกล้ที่สุด 1 ตัว ด้วยวิธีการ Leave-One-Out Cross Validation เพื่อหาเพื่อหา พารามิเตอร์ <i>nLevel</i> ที่ดีที่สุด	51
ตารางที่ 4.3 แสดงการเปรียบเทียบอัตราความผิดพลาดในการจำแนกประเภทข้อมูลโดยวิธีการ เพื่อนบ้านที่ใกล้ที่สุด 1 ตัว ควบคู่กับการวัดระยะห่างแบบไดนามิกไทม์วอร์ปิง ไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D และไดนามิกไทม์ วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG1D-L	57
ตารางที่ 4.4 สรุปผลการทดลอง.....	61

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ข้อมูลอนุกรมเวลา (Time series data) คือ ข้อมูลที่มีการเปลี่ยนแปลงไปตามจุดต่าง ๆ ของเวลา มีลำดับก่อนหลัง สามารถพบได้ทั่วไปในชีวิตประจำวันและเป็นที่น่าสนใจในหลาย ๆ ด้าน เช่น ด้านการแพทย์ ด้านธุรกิจ และด้านอุตุนิยมวิทยา นอกจากนี้ยังมีการนำข้อมูลประเภท รูปภาพ เสียง และวิดีโอมาแปลงเป็นข้อมูลอนุกรมเวลา ทำให้สามารถนำข้อมูลอนุกรมเวลามาวิเคราะห์เพื่อหา รูปแบบความสัมพันธ์ของข้อมูลที่จะนำไปประยุกต์ใช้ด้านต่าง ๆ หนึ่งในวิธีการที่นิยมใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา คือ การทำเหมืองข้อมูล หลังจากการทำเหมืองข้อมูลจะทำให้ได้ ความสัมพันธ์ของข้อมูลอนุกรมเวลา เพื่อนำไปใช้ประโยชน์ในการจำแนกประเภทข้อมูล

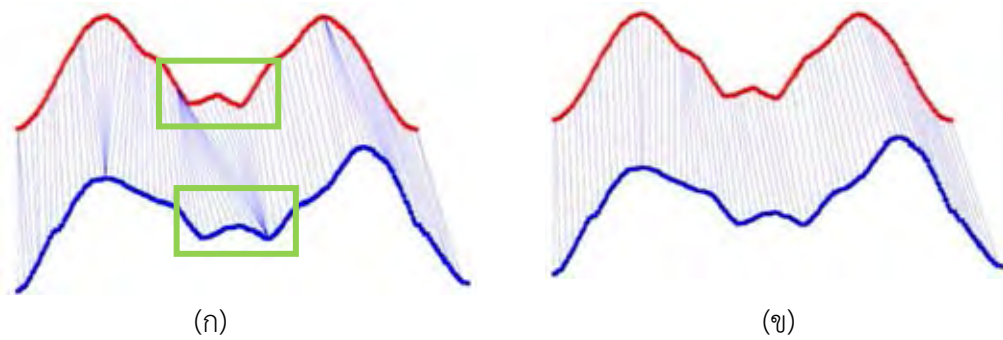
การจำแนกประเภทข้อมูลอนุกรมเวลา จะใช้ข้อมูลที่มีป้ายชื่อ (Labeled data) จำนวนหนึ่ง เพื่อทำการสอน (Training) ข้อมูลเหล่านั้นด้วยเกณฑ์การเรียนรู้บางอย่าง อัลกอริทึมในการจำแนกข้อมูลอนุกรมเวลามีอยู่หลายอัลกอริทึม เช่น ต้นไม้ตัดสินใจ (Decision tree) [1] ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine) [2] แบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว (1-Nearest Neighbor) ร่วมกับการวัดระยะทางแบบไดนามิกไทม์วอร์ปิง [3] เป็นต้น วิธีการที่ได้รับความนิยมและมีความแม่นยำสูงในการจำแนกประเภทข้อมูลคือ วิธีการจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว (1-Nearest Neighbor) ร่วมกับการวัดระยะทางแบบไดนามิกไทม์วอร์ปิง สำหรับข้อมูลอนุกรมเวลา ไดนามิกไทม์วอร์ปิง (Dynamic Time Warping) เป็นวิธีการวัดระยะทางที่มีความยืดหยุ่นในแกนเวลาและมีประสิทธิภาพที่ดีเมื่อทำงานร่วมกับการจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว ดังนั้นจึงมีการเสนอวิธีการจำนวนมากเพื่อเพิ่มความแม่นยำของการจำแนกข้อมูลให้มีประสิทธิภาพมากยิ่งขึ้น

จากงานวิจัยในปี 2016 [4] ได้มีการนำเสนอไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shape Dynamic Time Warping-shapeDTW) ซึ่งได้มีการนำมาประยุกต์ใช้ร่วมกับการจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว โดยใช้โครงสร้างพื้นที่ในแต่ละจุดในขั้นตอนการจับคู่ระหว่าง 2 อนุกรมเวลาโดยโครงสร้างรูปร่างที่เหมือนกันจะจับคู่ได้ดีกว่า จากรูปที่ 1.1 (ก) จะเห็นได้ว่าอนุกรมเวลา 2 อนุกรมที่มีโครงสร้างเฉพาะส่วนที่คล้ายคลึงกันหลายจุด จากการปรับแนวของไดนามิกไทม์วอร์ปิง นั้นไม่สามารถปรับแนวโครงสร้างเฉพาะส่วนที่มีความคล้ายคลึงกันได้ ในขณะที่ รูปที่ 1.1 (ข) ไดนามิกไทม์วอร์ปิงแบบรูปร่างได้ผลการปรับแนวที่มีความผิดพลาดน้อยกว่าการปรับแนวแบบ ไดนามิกไทม์วอร์ปิง ด้วยการพิจารณาโครงสร้างเฉพาะส่วนที่มีรูปร่างคล้ายคลึงกันระหว่าง 2 อนุกรมเวลา



425352952

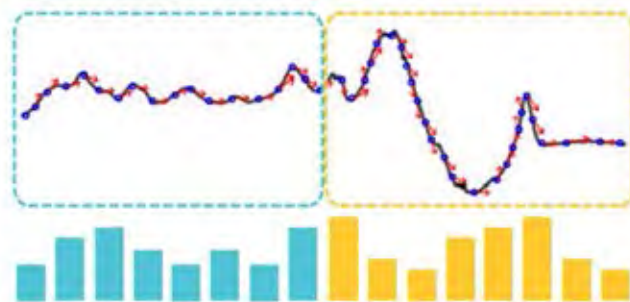
CU-Thesis 6070323921 thesis / rev: 06082562 14:42:31 / seq: 83



รูปที่ 1.1 แสดงการปรับแนวของ 2 อนุกรมเวลา (ก) การปรับแนวของไดนามิกไทม์วอร์ปิง
(ข) การปรับแนวของไดนามิกไทม์วอร์ปิงแบบรูปร่าง
(ที่มา : Jiaping Zhao and Laurent Ittiv [4])

ข้อเด่นอีกประการของไดนามิกไทม์วอร์ปิงแบบรูปร่างนั้นให้กรอบการทำงานทั่วไปที่ผู้ใช้สามารถออกแบบการอธิบายรูปร่างด้วยตัวเองซึ่งปรับเปลี่ยนได้ตามลักษณะข้อมูล ผลที่ได้คือไดนามิกไทม์วอร์ปิงแบบรูปร่างมีประสิทธิภาพที่ดีเยี่ยมเมื่อเทียบกับวิธีการไดนามิกไทม์วอร์ปิง [4] สำหรับแนวคิดสำคัญของการอธิบายรูปร่าง คือการแปลงลำดับของข้อมูลอนุกรมเวลาเป็นลำดับของตัวอธิบายรูปร่าง

จากงานวิจัยในปี 2016 [5] ได้นำเสนอตัวอธิบายรูปร่างที่พัฒนาขึ้นซึ่งมีชื่อว่า ฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์แบบ 1 มิติ (Histogram of Oriented Gradients 1D-HOG-1D) มาใช้เป็นตัวอธิบายรูปร่าง โดยพิจารณาอนุกรมเวลาของลำดับย่อยที่มีขนาดเท่ากัน ขั้นตอนวิธีของ HOG-1D จะอิงตามแนวคิดของฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์ (Histogram of Oriented Gradients-HOG) ซึ่งปรับเปลี่ยนมาเพื่อใช้ทำงานกับลำดับอนุกรมเวลาแบบ 1 มิติ โดยจะแบ่งลำดับอนุกรมเวลาออกเป็นลำดับย่อยที่มีขนาดเท่ากัน แล้วจึงใช้จุดของเวลาและจุดข้างเคียงในแต่ละลำดับย่อยเพื่อคำนวณหาขนาดของเกรเดียนต์และทิศทางของเกรเดียนต์เพื่อนำมาเก็บสะสมในฮิสโตแกรมของทิศทาง ดังรูปที่ 1.2 ซึ่งจุดเด่นของ HOG-1D นั้นจะทนต่อสิ่งรบกวนและไม่เปลี่ยนแปลงขนาดค่าคงที่ของแกน y



รูปที่ 1.2 อธิบายการทำงานของตัวอธิบายรูปร่าง HOG-1D ของลำดับย่อยอนุกรมเวลา

จากรูปที่ 1.2 แสดงการทำงานของตัวอธิบายรูปร่าง HOG-1D ของลำดับย่อยของอนุกรมเวลา โดยเริ่มจากรับลำดับย่อยของอนุกรมเวลาเข้ามา ลำดับย่อยดังกล่าวจะถูกแบ่งออกเป็นช่วงที่ทับซ้อนกันหรือไม่ทับซ้อนกันขึ้นอยู่กับข้อกำหนดค่าของผู้ใช้งาน จากงานวิจัย [4] ได้ทำการทดลองการแบ่งลำดับย่อยออกเป็นช่วงที่ไม่ทับซ้อนกัน โดยที่จุดของลำดับย่อยในแต่ละช่วงที่ถูกแบ่งจะถูกนำไปคำนวณด้วยฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ เพื่อหาขนาดของเกรเดียนท์ (Gradient Magnitude) และทิศทางตามค่าเกรเดียนท์ (Gradient Orientation) จากนั้นทิศทางและขนาดของเกรเดียนท์ที่คำนวณได้จะถูกนำมาเก็บสะสมในฮิสโตแกรมตามค่าทิศทางของเกรเดียนท์ โดยทิศทางตามค่าเกรเดียนท์ในแต่ละช่วงของฮิสโตแกรมจะวางตัวอยู่ในช่วง -90 ถึง 90 องศา ซึ่งในรูป 1.2 นี้จะทำการแบ่งทิศทางในช่วงนั้นออกเป็น 8 ส่วนที่มีขนาดเท่ากัน อย่างไรก็ตาม เนื่องจากการทำงานของตัวอธิบายรูปร่าง HOG-1D นั้นจะรวบรวมจำนวนการเกิดของทิศทางตามค่าเกรเดียนท์ของแต่ละจุดในลำดับย่อย โดยไม่ได้ทำการปรับให้เรียบ (Smoothing) เพื่อลดสัญญาณรบกวน (Noise) เช่นลำดับย่อยของอนุกรมเวลาที่ค่าภายในมีความผันผวนสูง ซึ่งหากนำขนาดของเกรเดียนท์ที่มีทิศทางที่แตกต่างจากทิศทางโดยรวมจากช่วงที่กำหนดไปเก็บรวบรวมไว้ในฮิสโตแกรมค่าเกรเดียนท์ อาจทำให้การจำแนกประเภทข้อมูลไม่แม่นยำ

เพื่อแก้ไขปัญหาดังกล่าว ในวิทยานิพนธ์นี้ ผู้วิจัยจึงมุ่งหวังที่จะพัฒนาตัวอธิบายรูปร่างใหม่ โดยการใช้ตัวอธิบายรูปร่างร่วมกับวิธีการวิเคราะห์ความถดถอยเชิงเส้น โดยการนำการวิเคราะห์ความถดถอยเชิงเส้นมาใช้เพื่อหาทิศทางเฉลี่ยในช่วงที่ถูกแบ่งและเลือกขนาดของเกรเดียนท์ที่มีทิศทางไปในแนวเดียวกันกับทิศทางเฉลี่ย เพื่อเป็นการเพิ่มประสิทธิภาพในการทำงานให้แก่ตัวอธิบายรูปร่าง HOG-1D และช่วยลดอัตราความผิดพลาดในการจำแนกประเภทข้อมูลของ HOG-1D ให้มีประสิทธิภาพเพิ่มมากยิ่งขึ้น

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อนำเสนอตัวอธิบายรูปร่าง ที่ใช้งานภายใต้โดนามิกไทม์วอร์ปิงแบบรูปร่าง
2. เพื่อลดอัตราความผิดพลาดในด้านการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้ตัวอธิบายรูปร่างให้มีประสิทธิภาพมากยิ่งขึ้น

1.3 ขอบเขตของงานวิจัย

1. พัฒนาตัวอธิบายรูปร่าง เพื่อใช้ในการจำแนกประเภทข้อมูลอนุกรมเวลาภายใต้การทำงาน ของโดนามิกไทม์วอร์ปิงแบบรูปร่าง
2. ทดสอบประสิทธิภาพของวิธีการที่นำเสนอ ด้วยการคำนวณอัตราความผิดพลาดในการ จำแนกประเภทข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว ภายใต้การทำงานของโดนามิกไทม์

วอร์ปิงแบบรูปร่าง โดยเปรียบเทียบกับวิธีไดนามิกใหม่วอร์ปิง และตัวอธิบายรูปร่าง HOG-1D ภายใต้การทำงานของไดนามิกใหม่วอร์ปิงแบบรูปร่าง เพื่อแสดงให้เห็นถึง ประสิทธิภาพในการทำงาน

3. ทำการทดลองด้วยชุดข้อมูลจาก UCR Time Series Classification Archive [6]

1.4 ประโยชน์ที่ได้รับจากงานวิจัย

1. ได้ตัวอธิบายรูปร่างภายใต้การทำงานของไดนามิกใหม่วอร์ปิงแบบรูปร่าง
2. ได้อัลกอริทึมในการปรับแนวของเวลาที่ทำให้ประสิทธิภาพแม่นยำกว่าไดนามิกใหม่วอร์ปิง ด้วยวิธีการจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว
3. ได้ตัวอธิบายรูปร่างที่ให้ประสิทธิภาพแม่นยำกว่าตัวอธิบายรูปร่าง HOG-1D ภายใต้การทำงานของไดนามิกใหม่วอร์ปิงแบบรูปร่าง ด้วยการทดลองจากชุดข้อมูล UCR Time Series Classification Archive [6]



425352952

บทที่ 2

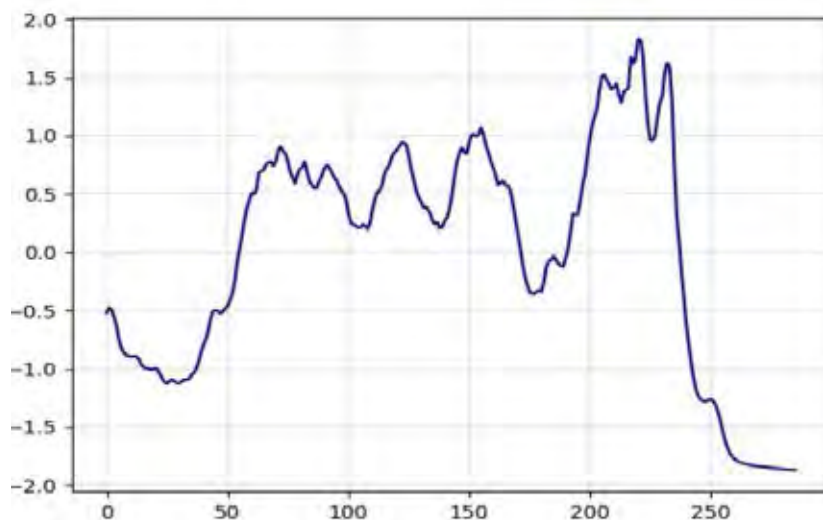
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในส่วนนี้กล่าวถึงหลักทฤษฎีที่เกี่ยวข้องในการพัฒนางานวิจัย และงานวิจัยที่เกี่ยวข้องตามลำดับ สำหรับทฤษฎีที่เกี่ยวข้อง ประกอบด้วย ข้อมูลอนุกรมเวลา (Time Series Data) ลำดับย่อยของอนุกรมเวลา (Time series subsequence) การจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด k ตัว (k -Nearest Neighbor Classification) ระยะทางยุคลิด (Euclidean Distance) ไดนามิกไทม์วอร์ปิง (Dynamic Time Warping) ฮิสโตแกรมของทิศทางการไล่ระดับ (Histogram of Oriented Gradients) และการวิเคราะห์ความถดถอยเชิงเส้น (Linear Regression Analysis) ในส่วนงานวิจัยที่เกี่ยวข้อง ประกอบด้วย ไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shape Dynamic Time Warping) และงานวิจัยที่เกี่ยวข้องกับตัวอธิบายรูปร่าง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 ข้อมูลอนุกรมเวลา (Time Series Data)

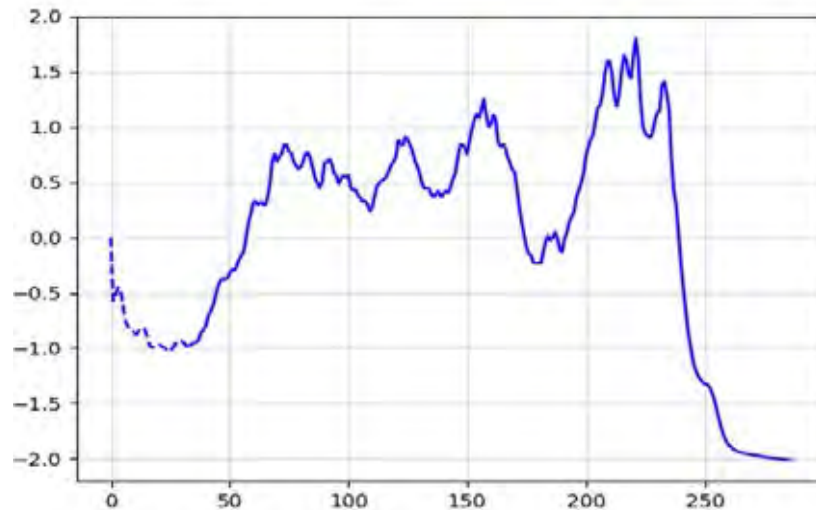
ข้อมูลอนุกรมเวลา คือ ข้อมูลใด ๆ ที่ได้รับการบันทึกค่าไว้ ณ จุดต่าง ๆ ของเวลา ข้อมูลอนุกรมเวลาสามารถพบเห็นได้ทั่วไปในชีวิตประจำวัน เช่น ข้อมูลราคาหุ้น หรือข้อมูลความรุนแรงของแผ่นดินไหวในปีต่าง ๆ หรือ อนุกรมเวลาของชุดข้อมูล coffee เป็นต้น รูปที่ 2.1 เป็นกราฟแสดงตัวอย่างอนุกรมเวลาของชุดข้อมูล coffee จาก UCR Time Series Classification Archive [6] เพื่อนำมาใช้ประโยชน์ในการคัดแยกสายพันธุ์ของเมล็ดกาแฟ



รูปที่ 2.1 กราฟแสดงตัวอย่างอนุกรมเวลาของชุดข้อมูล coffee จาก UCR Time Series Classification Archive [6]

2.1.2 ลำดับย่อยของอนุกรมเวลา (Time series subsequence)

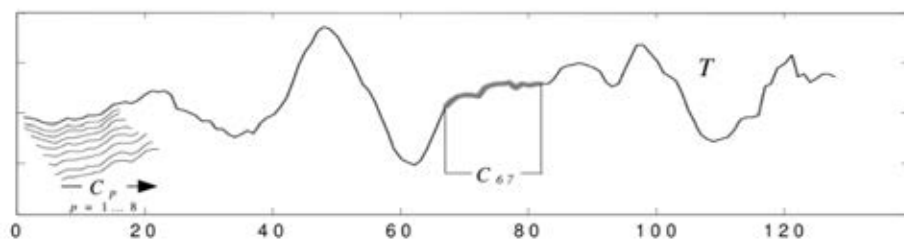
ลำดับย่อยของอนุกรมเวลา คือ อนุกรมเวลาที่พิจารณาเฉพาะบางส่วนของอนุกรมเวลาทั้งหมด ที่ได้รับการบันทึกค่าไว้ ณ จุดต่าง ๆ ดังแสดงเป็นเส้นประในรูปที่ 2.2 ซึ่งเป็นกราฟแสดงตัวอย่าง ลำดับย่อยของชุดข้อมูล coffee จาก UCR Time Series Classification Archive [6]



รูปที่ 2.2 กราฟแสดงตัวอย่างลำดับย่อยของอนุกรมเวลาของชุดข้อมูล coffee (เส้นประ) จาก UCR Time Series Classification Archive [6]

วิธีการในการหาลำดับอนุกรมย่อยของอนุกรมเวลาสามารถหาได้จากหลากหลายเทคนิค โดยงานวิจัยนี้ได้ใช้เทคนิค Sliding Window [7] ในการหาลำดับย่อยของอนุกรมเวลา โดยแนวคิดของเทคนิค Sliding Window สามารถคำนวณได้ดังนี้

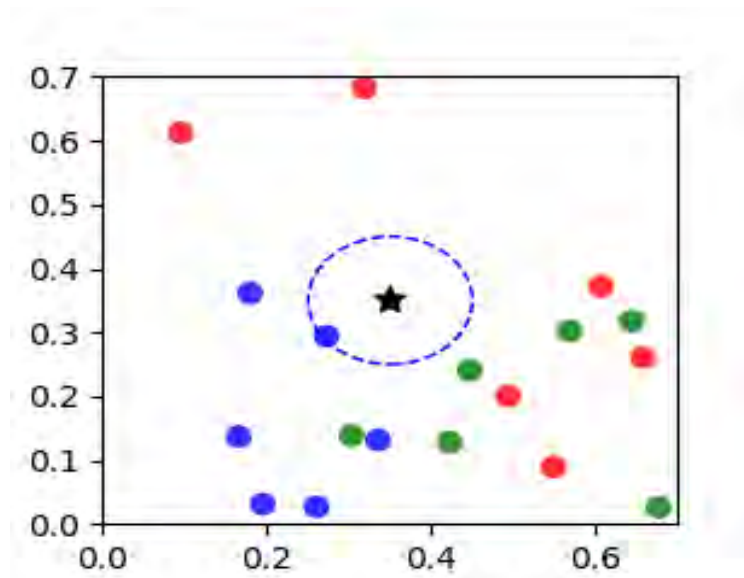
กำหนดให้อนุกรมเวลา $T = \{t_1, t_2, \dots, t_m\}$ มีความยาวเท่ากับ m และกำหนดให้ความยาวของลำดับอนุกรมย่อยเท่ากับ w จากนั้นกำหนดให้ S เป็นเมทริกซ์ที่ใช้ในการเก็บข้อมูลลำดับย่อยที่ได้จากเทคนิค Sliding Window มีจำนวนข้อมูลเท่ากับ $m - w + 1$ โดยให้ C_p คือลำดับย่อยของอนุกรมเวลา T จะได้ $C_p = t_p, \dots, t_{p+w-1}$ โดยที่ $1 \leq p \leq m - w + 1$ แสดงแนวคิดได้ดังรูปที่ 2.3



รูปที่ 2.3 แนวคิดการหาลำดับย่อยของอนุกรมเวลาด้วยเทคนิค Sliding Window (ที่มา : Eamonn Keogh และ Jessica Lin [7])

2.1.3 การจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด k ตัว (k -Nearest Neighbor Classification)

การจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด k ตัว เป็นวิธีการที่ใช้สำหรับการการจำแนกประเภทข้อมูล (Classification) โดยตัวแปรที่ใช้จำแนกข้อมูลจะประกอบไปด้วยตัวแปร k ตัว ซึ่งผู้ใช้สามารถกำหนดค่าตัวแปร k เพื่อระบุจำนวนข้อมูลที่ใกล้ข้อมูลที่ต้องการจำแนกประเภทข้อมูลมากที่สุดได้ เมื่อทราบข้อมูลที่ใกล้ที่สุดแล้ว จากนั้นจะสำรวจชื่อประเภทของข้อมูลที่ถูกระบุไว้ โดยข้อมูลที่นำเข้าไปทดสอบจะถูกกำหนดค่าตามประเภทของข้อมูลที่ใกล้ที่สุดและมีจำนวนมากที่สุดในจำนวน k ตัว ถ้าค่า k ถูกกำหนดให้มีค่าเท่ากับ 1 เช่นเดียวกับวิทยานิพนธ์นี้ การจำแนกประเภทของข้อมูลจะถูกกำหนดหลังจากได้เพื่อนบ้านที่ใกล้ที่สุดเพียง 1 ตัว ดังรูปที่ 2.4



รูปที่ 2.4 การจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว

2.1.4 ระยะทางยุคลิด (Euclidean Distance)

ระยะทางยุคลิด เป็นวิธีการวัดระยะทางระหว่างจุดของอนุกรมเวลา ณ จุดเวลาที่อยู่ตำแหน่งเดียวกัน แล้วคืนค่าเป็นระยะทางสำหรับการวัดความคล้ายคลึงกันระหว่างชุดข้อมูล ดังรูปที่ 2.5 การวัดระยะทางแบบยุคลิดสามารถคำนวณได้ดังนี้

กำหนดให้มีอนุกรมเวลา 2 อนุกรม ได้แก่ อนุกรมเวลา Q และ อนุกรมเวลา C ที่มีความยาว n เมื่อ

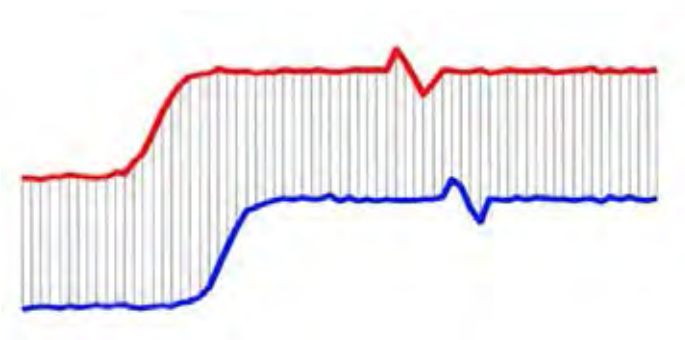
$$Q = q_1, q_2, \dots, q_i, \dots, q_n$$

$$C = c_1, c_2, \dots, c_i, \dots, c_n$$

ซึ่งสามารถคำนวณระยะทางยุคลิดได้จากสมการที่ 2.1

$$\text{Euclidean } (Q, C) = \sqrt{\sum_{i=1}^n (q_i - c_i)^2} \quad (2.1)$$

โดยที่ i คือ จุดเวลา ณ จุดใด ๆ ที่อยู่บนตำแหน่งเดียวกันระหว่าง 2 อนุกรมเวลา



รูปที่ 2.5 การวัดระยะทางยุคลิดของ 2 อนุกรมเวลา

(ที่มา : Chotirat Ann Ratanamahatana และ Eamonn Keogh [8])

2.1.5 ไดนามิกไทม์วอร์ปิง (Dynamic Time Warping-DTW)

ไดนามิกไทม์วอร์ปิง [9] เป็นวิธีการวัดระยะห่างระหว่างจุดของอนุกรมเวลา เพื่อค้นหาการปรับแนวที่เหมาะสมระหว่างอันดับของอนุกรมเวลา 2 ชุด แล้วคืนค่าเป็นระยะทางสำหรับการวัดความคล้ายคลึงกันระหว่างชุดข้อมูล โดยไดนามิกไทม์วอร์ปิงมีจุดเด่น คือ มีคุณสมบัติในการยืดหยุ่นของแนวแกนเวลาทำให้ทนทานต่อความคลาดเคลื่อนของข้อมูลเป็นอย่างดี จึงเป็นที่นิยมในการนำไปใช้งานและพัฒนาต่อยอดกันอย่างแพร่หลาย

ไดนามิกไทม์วอร์ปิงสามารถคำนวณได้ดังนี้ กำหนดให้มีอนุกรมเวลา 2 อนุกรม ได้แก่ อนุกรมเวลา Q ที่มีความยาว n และ อนุกรมเวลา C มีความยาว m เมื่อ

$$Q = q_1, q_2, \dots, q_i, \dots, q_n$$

$$C = c_1, c_2, \dots, c_j, \dots, c_m$$

จากนั้นคำนวณหาระยะทางระหว่าง 2 อนุกรม โดยกำหนดให้เมทริกซ์ d เป็นเมทริกซ์ที่เก็บระยะห่างของอนุกรมเวลาระหว่างจุด i และ j ของอนุกรมเวลา Q และ C ตามลำดับที่มีขนาดเท่ากับ $n \times m$ คำนวณได้จากสมการที่ 2.2

$$d_{i,j} = (q_i - c_j)^2 \quad (2.2)$$

เมื่อได้เมตริกซ์จากการคำนวณหาระยะห่างระหว่างจุดของอนุกรมเวลาแล้ว ในลำดับต่อไป จะทำการหาวิถีการวอร์ป (Warping Path) ที่เหมาะสมที่สุด โดยพิจารณาจากระยะทางสะสมของการวอร์ปที่มีค่าน้อยที่สุดระหว่าง 2 อนุกรม กำหนดให้ W แทนวิถีการวอร์ป โดยที่ตำแหน่งที่ k_{th} ของ w จะเขียนแทนด้วย $w_k = (i, j)_k$ แทนค่าคู่ลำดับระหว่าง 2 อนุกรมเวลา จะได้ $W = w_1, w_2, \dots, w_k, \dots, w_K$ โดย $\max(n, m) \leq K < n + m - 1$ ซึ่ง W_0 หรือ วิถีการวอร์ป ที่เหมาะสมที่สุดคือ วิถีการวอร์ปที่มีระยะทางสั้นที่สุด คำนวณได้ดังสมการที่ 2.3

$$DTW(Q, C) = \sqrt{\sum_{k=1}^K w_k} \quad (2.3)$$

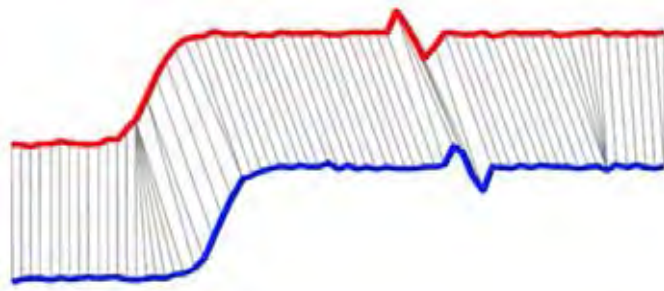
ซึ่งวิถีการวอร์ปที่เหมาะสมที่สุดของอนุกรมเวลา Q และ C สามารถคำนวณได้ด้วยวิธีการพลวัต (Dynamic Programming) เพื่อคำนวณหาระยะทางสะสมที่สั้นที่สุด (Cumulative Distance, $\gamma(i, j)$) คำนวณได้ดังสมการที่ 2.4

$$\gamma(i, j) = d(q_i, c_j) + \min \left\{ \begin{array}{l} \gamma(i-1, j-1), \\ \gamma(i-1, j), \\ \gamma(i, j-1) \end{array} \right\} \quad (2.4)$$

โดยมีเงื่อนไขดังต่อไปนี้

- เงื่อนไขขอบเขต (Boundary Condition) : ค่า i และ j ของอนุกรมเวลา Q และ C จะเริ่มต้นการปรับแนวที่ตำแหน่ง $(1, 1)$ และสิ้นสุดที่ตำแหน่ง (n, m)
- เงื่อนไขภาวะต่อเนื่อง (Continuity Condition) : ค่า i และ j ของอนุกรมเวลา Q และ C จะมีค่าเพิ่มทีละ 0 หรือ 1 ในแต่ละขั้นตอนของการปรับแนว
- เงื่อนไขทางเดียว (Monotonicity Condition) : ค่า i และ j ของอนุกรมเวลา Q และ C จะมีค่าเท่าเดิมหรือเพิ่มขึ้นได้เท่านั้น

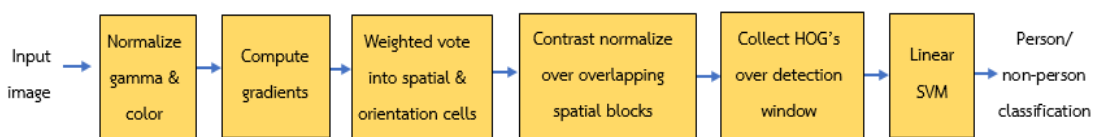
เมื่อคำนวณหาระยะทางสะสมที่น้อยที่สุดแล้ว จะทำการปรับแนวด้วยเส้นทางการวอร์ป โดยพิจารณาจุดข้อมูลที่มีระยะทางสะสมที่น้อยที่สุดและทำการจับคู่ระหว่างจุดข้อมูลนั้น จะได้ผลลัพธ์การปรับแนวที่เหมาะสมของ 2 อนุกรมเวลา ดังรูปที่ 2.6



รูปที่ 2.6 การปรับแนวของ 2 อนุกรมเวลาด้วยไดนามิกโทมวอร์ปิง
(ที่มา : Chotirat Ann Ratanamahatana และ Eamonn Keogh [8])

2.1.6 ฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ (Histogram of Oriented Gradients-HOG)

ฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ [10] เป็นวิธีการดึงลักษณะเด่นของภาพโดยใช้การกระจายตัวความเข้มของภาพเกรเดียนท์ และถูกนำมาใช้เพื่อตรวจจับภาพบุคคล (Human Detection) ซึ่งแสดงให้เห็นถึงลักษณะเด่นเฉพาะส่วนของบุคคล หรือวัตถุ โดยสามารถเก็บค่าได้โดยใช้การกระจายตัวของค่าความหนาแน่นเกรเดียนท์ แสดงขั้นตอนทำงาน ดังรูปที่ 2.7

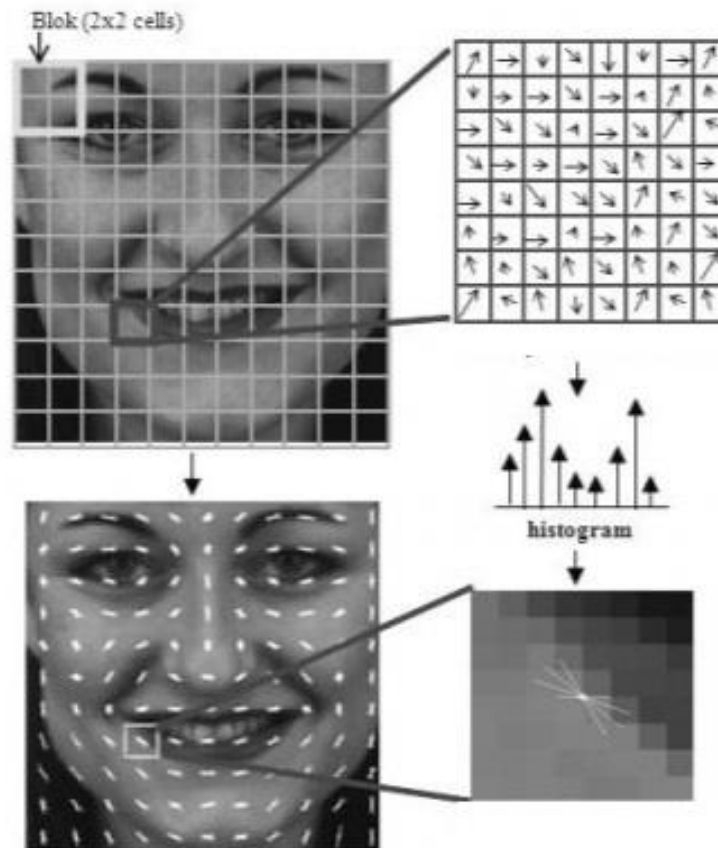


รูปที่ 2.7 ขั้นตอนการทำงานของฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ในการตรวจจับภาพบุคคล
(ที่มา : Navneet Dalal และ Bill Triggs [10])

จากรูปที่ 2.7 แสดงขั้นตอนการทำงานของฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์ในการตรวจจับภาพบุคคล ขั้นตอนแรก เริ่มที่การรับภาพเข้ามาและทำการนอร์มัลไลซ์ภาพจากความสว่าง (Gamma) และสี (Color) หลังจากนั้นคำนวณหาค่าเกรเดียนท์ (Gradient) แล้วทำการโหวต (Vote) หาค่าของทิศทาง (Orientation) ในแต่ละเซลล์ (Cell) และทำการนอร์มัลไลซ์ในแต่ละบล็อก (Block) ที่มีพื้นที่ทับซ้อนกัน จากนั้นทำการเก็บรวบรวมค่าที่คำนวณได้ทั้งหมดและจำแนกลักษณะเด่นเฉพาะส่วนด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีนแบบเชิงเส้น

วิธีฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์นั้น จะเริ่มจากการเตรียมภาพโดยการนอร์มัลไลซ์ภาพเพื่อดึงคุณลักษณะแสดงคุณลักษณะของวัตถุในภาพชัดเจน จากนั้นจะทำการแบ่งรูปภาพออกเป็นเซลล์ ในแต่ละเซลล์จะมีขนาด $m \times m$ พิกเซล โดย m คือจำนวนพิกเซลในแนวนอนและแนวตั้งของเซลล์ ซึ่งการกำหนดขนาดของเซลล์ขึ้นอยู่กับขนาดของคุณลักษณะ (Feature) โดยจะต้องเพียงพอที่จะตรวจจับคุณลักษณะที่สนใจได้ เช่น ศีรษะ หรือใบหน้า จากนั้นทำการแบ่งรูปภาพเป็น

บล็อก ซึ่งในแต่ละบล็อกจะมีขนาด $n \times n$ เซลล์ โดย n คือจำนวนเซลล์ในแนวนอนและแนวตั้งของบล็อก ดังรูปที่ 2.8



รูปที่ 2.8 ขั้นตอนทำงานวิธีฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์
(ที่มา : Peng Wang และ คณะ [11])

เมื่อแบ่งรูปภาพออกเป็นบล็อกจากรูปที่ 2.8 แล้ว จากนั้นทำการคำนวณค่าเกรเดียนท์ในแนวนอนและแนวตั้งของรูปภาพเพื่อนำไปคำนวณหาขนาดและทิศทางของเกรเดียนท์ในแต่ละบล็อก ดังสมการที่ 2.5 และ 2.6

$$G_x = f(x + 1, y) - f(x - 1, y) \quad (2.5)$$

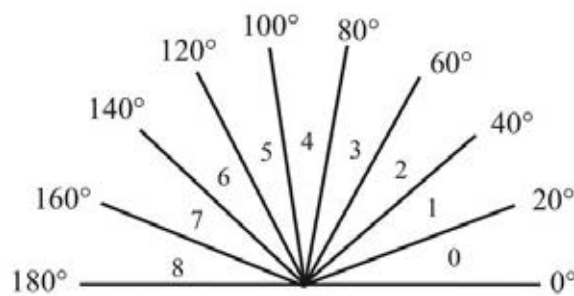
$$G_y = f(x, y + 1) - f(x, y - 1) \quad (2.6)$$

กำหนดให้ ค่า G_x และ G_y คือค่าเกรเดียนท์แนวนอนและแนวตั้งของรูปภาพ โดยที่ $f(x, y)$ คือค่าความสว่างที่ตำแหน่ง (x, y) จากนั้นทำการหาขนาดของเกรเดียนท์ (Gradient Magnitude) และทิศทางเกรเดียนท์ (Gradient Orientation) ในแต่ละเซลล์ คำนวณได้ดังสมการที่ 2.7 และ 2.8

$$M(x, y) = \sqrt{G_x^2 + G_y^2} \quad (2.7)$$

$$\theta(x, y) = \arctan\left(\frac{G_y}{G_x}\right) \quad (2.8)$$

โดยกำหนดให้ $M(x, y)$ คือขนาดของเกรเดียนท์ และ $\theta(x, y)$ คือ ทิศทางของเกรเดียนท์ จากนั้นสร้างฮิสโตแกรมจากการคำนวณความถี่ของค่าทิศทางเกรเดียนท์ที่ปรากฏในแต่ละบล็อก โดยค่าของทิศทางเกรเดียนท์จะถูกจัดเก็บลงไปในถัง (Orientation bin) ตามขนาดที่ได้กำหนดไว้ 0 ถึง 180 องศา ดังรูปที่ 2.9



รูปที่ 2.9 ถังที่เก็บทิศทางของฮิสโตแกรมขนาด 0 ถึง 180 องศา
(ที่มา : Sangeetha Damu และ P. Deepa [12])

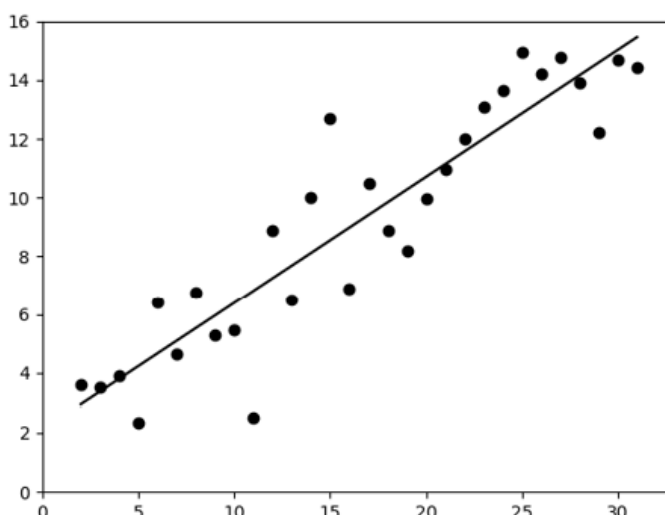
เนื่องจากค่าเกรเดียนท์นั้นจะเปลี่ยนแปลงเมื่อมีการปรับค่าแสงของภาพ ถ้าปรับแสงให้มืดขึ้นด้วยการหารค่าพิกเซลด้วย 2 จะทำให้ค่าเกรเดียนท์ลดลงครึ่งหนึ่งซึ่งส่งผลให้ค่าฮิสโตแกรมลดลงครึ่งหนึ่งด้วย ดังนั้นจึงต้องทำการนอร์มัลไลซ์ เพื่อไม่ให้ค่าเกรเดียนท์เปลี่ยนแปลงเมื่อมีการเปลี่ยนแปลงของค่าแสง ซึ่งการนอร์มัลไลซ์จะทำให้บล็อกโดยคำนวณได้จากสมการการนอร์มัลไลซ์แบบ L2-norm ดังสมการที่ 2.9

$$L2 - norm = \frac{v}{\sqrt{\|v\|_2^2 + e^2}} \quad (2.9)$$

โดยที่ v คือเวกเตอร์ของค่าฮิสโตแกรมในบล็อก $\|v\|$ คือวิธีการ k-norm แบบ L2-norm ของเวกเตอร์ v ของค่าฮิสโตแกรมในบล็อกและ e คือค่าคงที่มีค่าน้อยมากหรือสามารถละเว้นค่านี้ไว้ได้ หลังจากการนอร์มัลไลซ์จะนำค่าที่ได้เข้าสู่วิธีการจำแนกประเภทข้อมูล เช่น ซัพพอร์ตเวกเตอร์แมชชีน เพื่อใช้ในการทำ Human Detection ต่อไป

2.1.7 การวิเคราะห์ความถดถอยเชิงเส้น (Linear Regression Analysis)

การวิเคราะห์ความถดถอยเชิงเส้น เป็นเทคนิคที่ถูกใช้อย่างแพร่หลายในการวิเคราะห์ทางสถิติ โดยศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent variable) แทนด้วยตัวแปร x และตัวแปรตาม (Dependent variable) แทนด้วยตัวแปร y ซึ่งผลลัพธ์ที่ได้จากการวิเคราะห์ความถดถอยเชิงเส้นจะทำให้สามารถหาความสัมพันธ์หรือสมการที่ใช้ในการทำนายตัวแปรตาม y จากกลุ่มตัวแปรอิสระ x ดังรูปที่ 2.10



รูปที่ 2.10 ตัวอย่างการวิเคราะห์ความถดถอยเชิงเส้น

ในวิทยานิพนธ์นี้จะกล่าวถึงการวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression) เพื่อใช้ในการคำนวณทิศทางของแนวเส้นความถดถอยเชิงเส้น จากจุดต่าง ๆ ของเวลาซึ่งเป็นการหาความสัมพันธ์ระหว่างตัวแปร x และตัวแปร y ในลักษณะเส้นตรงโดยแสดงความสัมพันธ์ของทั้งสองตัวแปรในรูปของสมการเชิงเส้นตรง ดังสมการที่ 2.10

$$y = a + bx \quad (2.10)$$

จากสมการ 2.10 กำหนดให้ y คือค่าตัวแปรตาม a คือค่าสัมประสิทธิ์ความถดถอยซึ่งเป็นตัวกำหนดจุดตัดแกน y b คือค่าสัมประสิทธิ์ความถดถอยซึ่งเป็นตัวกำหนดความชันของเส้นสมการความถดถอยเชิงเส้น และ x คือค่าตัวแปรตาม สามารถคำนวณได้ดังสมการที่ 2.11 และ 2.12

$$a = \bar{y} - b\bar{x} \quad (2.11)$$

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.12)$$

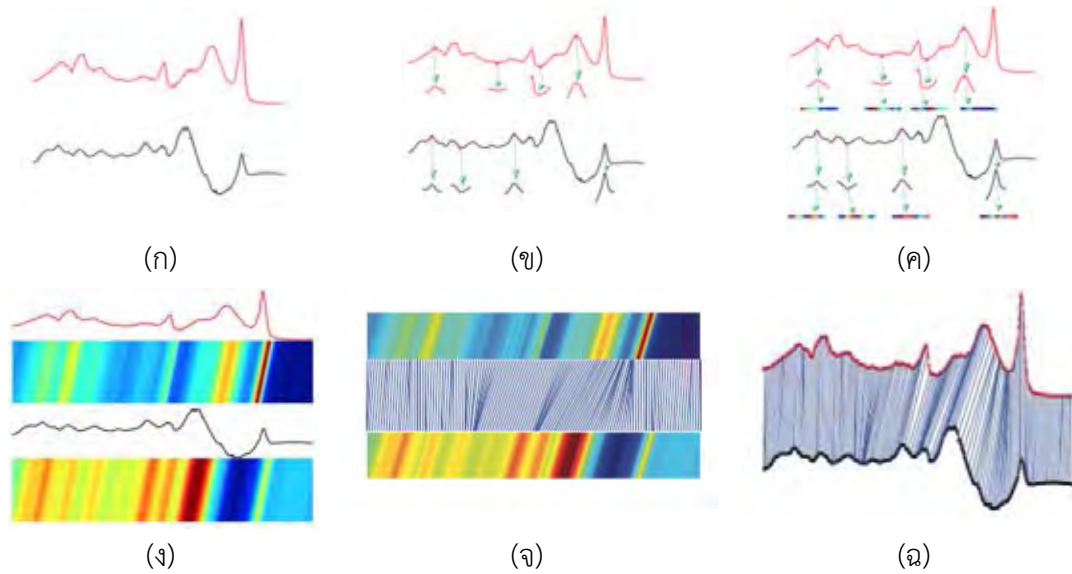
จากสมการ 2.11 และ 2.12 กำหนดให้ x_i คือค่าตัวแปรต้นที่ตำแหน่งจุดข้อมูล i y_i คือค่าตัวแปรตามที่ตำแหน่งจุดข้อมูล i \bar{x} คือค่าเฉลี่ยของตัวแปรต้น \bar{y} คือ ค่าเฉลี่ยของตัวแปรตาม และ n คือจำนวนความยาวของอนุกรมเวลา

2.2 งานวิจัยที่เกี่ยวข้อง

วิทยานิพนธ์นี้มุ่งหวังในการพัฒนาตัวอธิบายรูปร่างเพื่อเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาให้เพิ่มขึ้น เนื่องจากความประสบความสำเร็จในการพัฒนาไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shape Dynamic Time Warping-shapeDTW) [4] ที่สามารถเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลให้มีความแม่นยำมากขึ้นร่วมกับวิธีการจำแนกข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว โดยมีกรอบการทำงานการแปลงข้อมูลอนุกรมเวลาเป็นลำดับของตัวอธิบายรูปร่าง ในส่วนแรกผู้วิจัยจะกล่าวถึงขั้นตอนการทำงานของไดนามิกไทม์วอร์ปิงแบบรูปร่าง โดยคร่าว และในส่วนถัดไปจะกล่าวถึงงานวิจัยที่เกี่ยวข้องกับตัวอธิบายรูปร่างทั่วไปในรูปแบบต่าง ๆ

2.2.1 ไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shape Dynamic Time Warping-shapeDTW) [4]

วิธีการวัดระยะทางแบบไดนามิกไทม์วอร์ปิง เป็นวิธีการปรับแนวที่เหมาะสมโดยรวมของอนุกรมเวลาแต่ยังไม่สามารถปรับแนวเพื่อทำการเปรียบเทียบเฉพาะส่วนได้อย่างสมเหตุสมผล ไดนามิกไทม์วอร์ปิงแบบรูปร่างจึงนำเสนอแนวคิดการรวบรวมข้อมูลรูปร่างเฉพาะส่วนของบริเวณโดยรอบของแต่ละจุดของอนุกรมเวลาเข้าสู่การปรับแนวด้วยไดนามิกไทม์วอร์ปิง เพื่อผลการปรับแนวที่ดียิ่งขึ้น โดยไดนามิกไทม์วอร์ปิงแบบรูปร่างจะประกอบด้วย 2 ขั้นตอนหลัก ๆ ดังนี้ 1) แสดงแต่ละจุดของเวลาด้วยการใช้ตัวอธิบายรูปร่าง และ 2) ปรับแนวลำดับอนุกรมเวลาของตัวอธิบายรูปร่างด้วยไดนามิกไทม์วอร์ปิง โดยขั้นตอนการทำงานของไดนามิกไทม์วอร์ปิงแบบรูปร่าง แสดงได้ดังรูปที่ 2.11



รูปที่ 2.11 ขั้นตอนการทำงานของไดนามิกไทม์วอร์ปิงแบบรูปร่าง (ก) อนุกรมเวลาที่ป้อนเข้า (ข) แบ่งลำดับย่อยของอนุกรมเวลา (ค) คำนวณตัวอธิบายรูปร่าง (ง) แสดงลำดับย่อยของตัวอธิบายรูปร่าง (จ) ปรับแนวลำดับย่อยของตัวอธิบายรูปร่างด้วยไดนามิกไทม์วอร์ปิง (ฉ) ส่งผ่านการปรับแนวไปยังอนุกรมเวลาต้นฉบับ

(ที่มา : Jiaping Zhao and Laurent Ittiv [4])

ไดนามิกไทม์วอร์ปิงแบบรูปร่างสามารถคำนวณได้ดังนี้ กำหนดให้มีอนุกรมเวลา 2 อนุกรม ได้แก่ อนุกรมเวลา P ที่มีความยาว L_P และอนุกรมเวลา Q มีความยาว L_Q เมื่อ

$$P = p_1, p_2, \dots, p_i, \dots, p_{L_P}$$

$$Q = q_1, q_2, \dots, q_i, \dots, q_{L_Q}$$

การคำนวณไดนามิกไทม์วอร์ปิงแบบรูปร่าง จะมีการคำนวณลำดับย่อยของอนุกรมเวลา โดยกำหนดให้ S^P คืออนุกรมเวลาที่ใช้เก็บค่าลำดับย่อยของอนุกรมเวลา P มีความยาว L_P S^Q คืออนุกรมเวลาที่ใช้เก็บค่าลำดับย่อยของอนุกรมเวลา Q มีความยาว L_Q และ s_i คือลำดับย่อยของอนุกรมเวลา เมื่อ

$$S^P = s_{p_1}, s_{p_2}, \dots, s_{L_P}$$

$$S^Q = s_{q_1}, s_{q_2}, \dots, s_{L_Q}$$

เมื่อได้ลำดับย่อยของอนุกรมเวลาแล้วจะทำการแปลงลำดับย่อยด้วยตัวอธิบายรูปร่าง กำหนดให้ $F(.)$ คือ ฟังก์ชันคำนวณตัวอธิบายรูปร่าง (Shape descriptor function) ซึ่ง $F(.)$ จะทำการแปลงลำดับย่อยของอนุกรมเวลา S^P และ S^Q ไปเป็นอนุกรมเวลาของตัวอธิบายรูปร่าง

$$d^P = F(s_{p_1}), F(s_{p_2}), \dots, F(s_{L_P})$$

$$d^Q = F(s_{q_1}), F(s_{q_2}), \dots, F(s_{L_Q})$$

โดยกำหนดให้ d เป็นอนุกรมเวลาที่เก็บลำดับย่อยของอนุกรมเวลาที่ผ่านฟังก์ชันคำนวณตัวอธิบายรูปร่าง (Shape descriptor sequence) โดยมีความยาว L_P และ L_Q หลังจากนั้นจะใช้ไดนามิกไทม์วอร์ปิงเพื่อปรับแนวอนุกรมเวลาของตัวอธิบายรูปร่างทั้งสองอนุกรมเวลาและส่งผ่านการปรับแนวไปยังอนุกรมเวลาดั้งเดิม ซึ่งในตาราง 2.1 แสดงอัลกอริทึมการทำงานของไดนามิกไทม์วอร์ปิงแบบรูปร่าง

คำนิยามต่าง ๆ ที่ใช้ในอัลกอริทึมไดนามิกไทม์วอร์ปิงแบบรูปร่างมีดังนี้ [4]

1. อนุกรมเวลาแบบตัวแปรเชิงเดี่ยว (Univariate time series)

ในการประยุกต์ใช้ไดนามิกไทม์วอร์ปิง สามารถใช้ได้ทั้งค่าที่เป็นแบบตัวแปรแบบเชิงเดี่ยว (Univariate) และตัวแปรแบบพหุ (Multivariate) ในอัลกอริทึมไดนามิกไทม์วอร์ปิงแบบรูปร่าง กำหนดให้อนุกรมเวลาของ P และ Q เป็นแบบตัวแปรเชิงเดี่ยว มีความยาว L_P และ L_Q ตามลำดับ

2. กลุ่มตัวอย่างของลำดับย่อย (Sample subsequence)

กลุ่มตัวอย่างของลำดับย่อย คืออนุกรมเวลาที่เก็บลำดับย่อยของอนุกรมเวลา โดยใช้เทคนิค Sliding window มาใช้ในการหาลำดับย่อยของอนุกรมเวลา

3. การแปลงลำดับย่อย (Encode subsequences)

การแปลงลำดับย่อย คือการแปลงลำดับย่อยของอนุกรมเวลาให้เป็นลำดับของตัวอธิบายรูปร่าง ตามที่ผู้ใช้เลือกใช้งานจากฟังก์ชันคำนวณตัวอธิบายรูปร่าง หรือ $F(.)$

4. เมทริกซ์ที่เก็บการวอร์ปของอนุกรมเวลา (Warping matrices)

เมทริกซ์ที่เก็บการวอร์ปของอนุกรมเวลาเป็นเมทริกซ์ที่เก็บการวอร์ปของอนุกรมเวลา P และ Q กำหนดให้ \tilde{W}^P เป็นเมทริกซ์ที่แทนวิธีการวอร์ปของ P และ \tilde{W}^Q เมทริกซ์ที่แทนวิธีการวอร์ป Q

$$\tilde{W}^P = w_1^P, w_2^P, \dots, w_k^P, \dots, w_K^P$$

$$\tilde{W}^Q = w_1^Q, w_2^Q, \dots, w_k^Q, \dots, w_K^Q$$

5. การวัดระยะทางแบบไดนามิกไทม์วอร์ปิงแบบรูปร่าง (shapeDTW distance)
การวัดระยะทางแบบไดนามิกไทม์วอร์ปิงแบบรูปร่าง จะวัดความคล้ายคลึงระหว่างตัวอธิบายรูปร่างต่างๆ ด้วย $\|d_i^P - d_j^Q\|_2$

ตารางที่ 2.1 อัลกอริทึมไดนามิกไทม์วอร์ปิงแบบรูปร่าง

Algorithm 1 shape Dynamic Time Warping	
Inputs : univariate time series $P \in R^{L_P}$ and $Q \in R^{L_Q}$; subsequence length l ; shape descriptor function F	
1	Sample subsequence: $S^P \leftarrow P, S^Q \leftarrow Q$
2	Encode subsequences by shape descriptors: $d^P \leftarrow F(S^P), d^Q \leftarrow F(S^Q)$
3	Find cost distance between subsequence of shape descriptor: $\ d_i^P - d_j^Q\ _2$
4	Align descriptor sequences d^P and d^Q by DTW
Output : warping matrices: \tilde{W}^P and \tilde{W}^Q , shapeDTW distance: $\ \tilde{W}^P \cdot d^P - \tilde{W}^Q \cdot d^Q\ _{1,2}$	

จากตาราง 2.1 อัลกอริทึมนี้เป็นอัลกอริทึมที่ใช้ในการปรับแนวของ 2 อนุกรมเวลา กำหนดให้ P และ Q เป็นอนุกรมเวลา ที่มีความยาว L_P และ L_Q มีลำดับย่อยของอนุกรมเวลาความยาว l และมีฟังก์ชันคำนวณตัวอธิบายรูปร่าง F

บรรทัดที่ 1 สร้างกลุ่มตัวอย่างของลำดับย่อยของอนุกรมเวลา โดยกำหนดให้ S^P และ S^Q คืออนุกรมเวลาที่เก็บลำดับย่อยของอนุกรมเวลา P และ Q ตามลำดับ โดยรหัสเทียมของการสร้างกลุ่มตัวอย่างของลำดับย่อย แสดงดังตารางที่ 2.2

บรรทัดที่ 2 แปลงลำดับย่อยของอนุกรม S^P และ S^Q ด้วยฟังก์ชันคำนวณตัวอธิบายรูปร่าง หรือฟังก์ชันเชื่อมโยง $F(\cdot)$ ซึ่งจะได้ลำดับของตัวอธิบายรูปร่าง d^P และ d^Q ตามลำดับ

บรรทัดที่ 3 หาระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่าง d_i^P และ d_j^O ด้วยวิธีการหา ระยะทางแบบยูคลิด โดยรหัสเทียมสำหรับการหาระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่าง แสดงดังตารางที่ 2.3

บรรทัดที่ 4 หลังจากได้ระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่างในบรรทัดที่ 3 แล้ว จะนำระยะห่างดังกล่าวมาปรับแนวของ 2 อนุกรมเวลาด้วยไดนามิกไทม์วอร์ปิง โดยความแตกต่างที่สำคัญระหว่างไดนามิกไทม์วอร์ปิง และไดนามิกไทม์วอร์ปิงแบบรูปร่าง คือไดนามิกไทม์วอร์ปิงจะ วัดความคล้ายคลึงระหว่าง p_i และ q_i ด้วย $|p_i - q_i|$ ขณะที่ไดนามิกไทม์วอร์ปิงแบบรูปร่าง จะวัดความคล้ายคลึงระหว่างตัวอธิบายรูปร่างต่าง ๆ ด้วย $\|d_i^P - d_j^O\|_2$ และปรับแนวของ 2 อนุกรมเวลาด้วย ซึ่งไดนามิกไทม์วอร์ปิงแบบรูปร่างจะจัดการกับการวอร์ปเฉพาะส่วนเป็นหลักและ ปรับแนวของ 2 อนุกรมเวลาด้วยไดนามิกไทม์วอร์ปิง แสดงดังตารางที่ 2.4

ตารางที่ 2.2 อัลกอริทึมสร้างกลุ่มตัวอย่างของลำดับย่อย

Algorithm 2 samplingSubsequence(<i>ts</i> , <i>seqLen</i>)	
Input: univariate time series <i>ts</i> and subsequence length <i>seqLen</i>	
Output: samplingSubsequence	
1	samplingSubsequence [<i>ts</i> length]
2	for $i = 1$ to <i>ts</i> length
3	if $i < (seqLen / 2)$ then
4	tempSequence [(<i>seqLen</i> /2) - i]
5	for $j = 1$ to (<i>seqLen</i> /2) - i
6	tempSequence[j] \leftarrow <i>ts</i> [1]
7	end for
8	samplingSequence[i] \leftarrow tempSequence append <i>ts</i> [1: (<i>seqLen</i> /2) + i]
9	else if $i > ts$ length - (<i>seqLen</i> / 2) then
10	Declare tempSequence [$i + (seqLen/2) - ts$ length]
11	for $j = 1$ to $i + (seqLen/2) - ts$ length
12	tempSequence [j] \leftarrow [<i>ts</i> length]
13	end for

14	$sameplingSequence[i] \leftarrow ts[i - (seqLen/2) + 1]$ append $tempSequence$
15	else
16	$samplingSubsequence [i] \leftarrow ts[i - (seqLen/2) + 1 : i + (seqLen)]$
17	end if
18	end for

จากตาราง 2.2 อัลกอริทึมนี้เป็นอัลกอริทึมที่ใช้ในการหากลุ่มตัวอย่างของลำดับย่อย โดยนำข้อมูลที่นำเข้า กำหนดให้ ts คืออนุกรมเวลาที่จะทำการสร้างกลุ่มตัวอย่างของลำดับย่อย และ $seqLen$ คือความยาวของลำดับย่อยที่ต้องการโดยจะต้องเป็นจำนวนคู่ โดยให้ผลลัพธ์คือกลุ่มตัวอย่างของลำดับย่อย

บรรทัดที่ 1 ประกาศตัวแปรอาร์เรย์ $samplingSequence$ ขนาดเท่ากับความยาวของอนุกรมเวลา ts เพื่อใช้สำหรับเก็บตัวอย่างของลำดับย่อยที่ถูกสร้างขึ้น

บรรทัดที่ 2 วนลูปเพื่อสร้างลำดับย่อยตามจำนวนของความยาวอนุกรมเวลา ts

บรรทัดที่ 3 - 8 ถ้าค่า i น้อยกว่าค่า $seqLen$ ทหารด้วย 2 จะประกาศตัวแปร $tempSequence$ ขนาดเท่ากับ $seqLen$ ทหารด้วย 2 แล้วลบ i เพื่อเก็บค่าลำดับแรกของอนุกรมเวลา ts จำนวนเท่ากับ $seqLen$ ทหารด้วย 2 แล้วลบ i แล้วนำมาเรียงต่อกับอนุกรมเวลา ts ที่ลำดับที่ 1 จนถึงลำดับที่ $seqLen$ ทหารด้วย 2 แล้วบวก i จากนั้นนำค่าที่เรียงต่อกันแล้วใส่ค่าเข้าไปใน $samplingSequence$ ลำดับที่ i

บรรทัดที่ 9 - 14 ถ้าค่า i มากกว่าค่าความยาวของอนุกรมเวลา ts ลบด้วย $seqLen$ ทหารด้วย 2 จะประกาศตัวแปร $tempSequence$ ขนาดเท่ากับ $i + seqLen$ ทหารด้วย 2 แล้วลบด้วยความยาวอนุกรมเวลา ts เพื่อเก็บค่าลำดับสุดท้ายของอนุกรมเวลา ts จำนวนเท่ากับ $i + seqLen$ ทหารด้วย 2 แล้วลบด้วยความยาวอนุกรมเวลา ts แล้วนำอนุกรมเวลา ts ลำดับที่ i ลบด้วย $seqLen$ ทหารด้วย 2 แล้วบวก 1 มาเรียงต่อกันกับ $tempSequence$ จากนั้นนำค่าที่เรียงต่อกันแล้วใส่ค่าเข้าไปใน $samplingSequence$ ลำดับที่ i

บรรทัดที่ 15 -16 ถ้าไม่ตรงเงื่อนไขทั้งสองเงื่อนไขข้างต้น จะกำหนดให้ค่า $samplingSequence$ ลำดับที่ i เท่ากับอนุกรมเวลาลำดับที่ i ลบด้วย $seqLen$ ทหารด้วย 2 แล้วบวกด้วย 1 จนถึง i บวก $seqLen$



425352952

ตารางที่ 2.3 อัลกอริทึมหาระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่างด้วยวิธีการหาระยะทางแบบยุคลิด

Algorithm 3 Euclidean(dp, dq)	
Input: shape descriptor subsequence dp , shape descriptor subsequence dq	
Output: A distance of euclidean distance between two shape descriptor sequence	
1	$dist \leftarrow 0$ # Cumulative euclidean distance between two shape descriptor subsequences
2	For $i = 1$ to dp length
3	$dist \leftarrow dist + \text{sqrt}((dp[i] - dq[i])^2)$
4	end for

จากตาราง 2.3 อัลกอริทึมนี้เป็นอัลกอริทึมที่ใช้ในการหาระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่าง โดยข้อมูลที่นำเข้า กำหนดให้ dp และ dq คือลำดับย่อยของตัวอธิบายรูปร่างที่มีขนาดความยาวเท่ากัน และให้ผลลัพธ์คือระยะห่างระหว่างลำดับย่อยของตัวอธิบายรูปร่าง

บรรทัดที่ 1 กำหนดให้ $dist$ คือตัวแปรสำหรับเก็บค่าผลรวมของระยะห่างระหว่างค่าที่อยู่ในลำดับย่อยทั้งสองลำดับ

บรรทัดที่ 2 - 3 คำนวณระยะห่างระหว่างค่าที่อยู่ในลำดับย่อยทั้งสองลำดับด้วยจำนวนบวกของรากที่สองของการยกกำลังสองของผลต่างระหว่างลำดับย่อยแรกหลักที่ i ลบกับลำดับย่อยที่สองหลักที่ i จากนั้นนำไปบวกกับตัวแปร $dist$

ตารางที่ 2.4 อัลกอริทึมใช้ในการปรับแนวของ 2 อนุกรมเวลาด้วยไดนามิกไทม์วอร์ปิง

Algorithm 4 DTW(p, q)	
Input: shape descriptor sequence p , shape descriptor sequence q	
Output: A dynamic time warping distance from array $D [p \text{ length} + 1, q \text{ length} + 1]$	
1	$D [p \text{ length} + 1][q \text{ length} + 1]$
2	for $i = 1$ to p length + 1
3	for $j = 1$ to q length + 1
4	$D [i, j] \leftarrow \text{Inf}$
5	end for

6	end for
7	$D [1, 1] \leftarrow 0$
8	for $i = 1$ to p length
9	for $j = 1$ to q length
10	$dist \leftarrow \text{Euclidean} (p [i], q [j])$
11	$D [i+1, j+1] \leftarrow dist + \min ([D (i, j+1), D (i+1, j), D (i, j)])$
12	end for
13	end for

จากตาราง 2.4 อัลกอริทึมนี้เป็นอัลกอริทึมที่ใช้ในการปรับแนวระหว่างลำดับของตัวอธิบายรูปร่าง โดยข้อมูลที่นำเข้ากำหนดให้ p และ q คือลำดับของตัวอธิบายรูปร่าง โดยให้ผลลัพธ์คือระยะห่างระหว่าง 2 อนุกรมเวลา

บรรทัดที่ 1 ประกาศอาร์เรย์ D ขนาดสองมิติโดยมีขนาดเท่ากับความยาว $p + 1$ กับขนาดเท่ากับความยาว $q + 1$ เพื่อใช้สำหรับเก็บค่าระยะห่างระหว่างสองลำดับย่อย

บรรทัดที่ 4 กำหนดค่าภายในอาร์เรย์ D เท่ากับค่าอนันต์

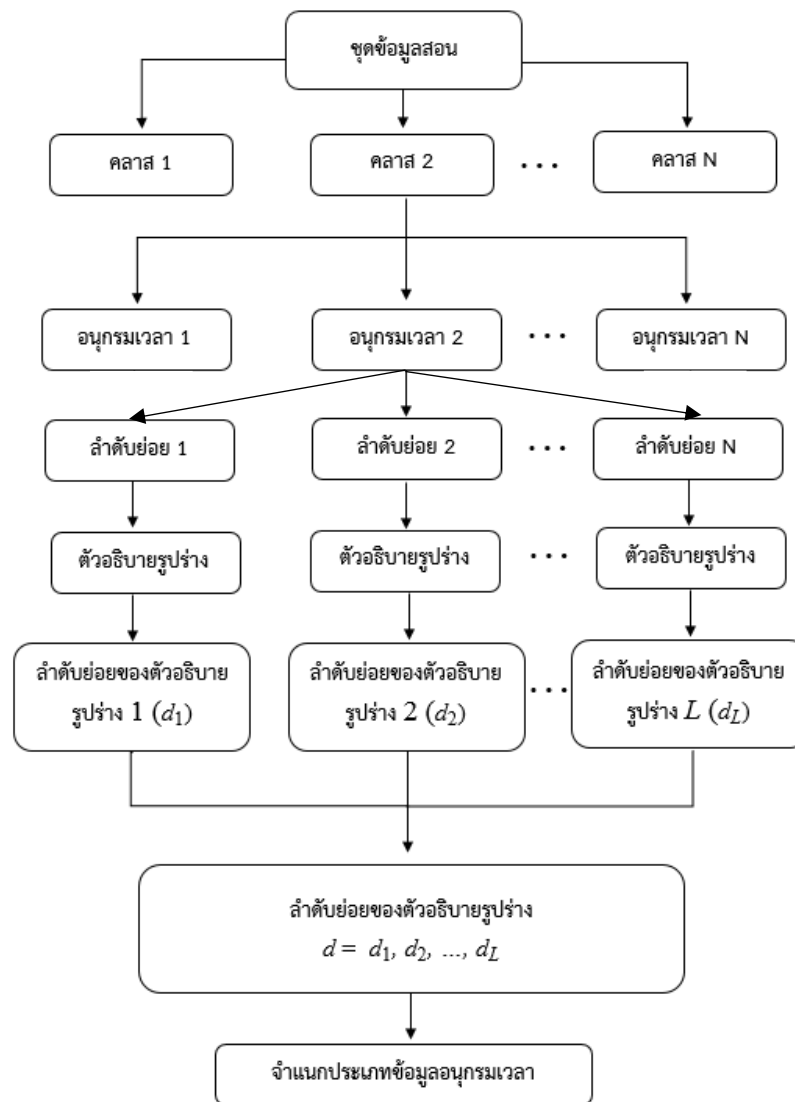
บรรทัดที่ 7 กำหนดค่าอาร์เรย์แถวที่ 1 หลักที่ 1 เท่ากับ 0

บรรทัดที่ 10 - 11 หาระยะห่างระหว่างสองลำดับย่อยที่อยู่ภายในตัวอธิบายรูปร่างด้วยฟังก์ชัน Euclidean จากนั้นนำค่าระยะห่างที่ได้บวกกับค่าระยะห่างที่น้อยที่สุดระหว่างอาร์เรย์ D แถวที่ i หลักที่ $j + 1$ อาร์เรย์ D แถวที่ $i + 1$ หลักที่ j และอาร์เรย์ D แถวที่ i หลักที่ j จากนั้นนำผลบวกไปใส่ในอาร์เรย์ D แถวที่ $i + 1$ หลักที่ $j + 1$

จากการทำงานของไดนามิกไทม์วอร์ปิงแบบรูปร่างที่ใช้ตัวอธิบายรูปร่างในการรวบรวมข้อมูลรูปร่างเฉพาะส่วนของบริเวณโดยรอบของแต่ละจุดของอนุกรมเวลาก่อนเข้าสู่การปรับแนวด้วยไดนามิกไทม์วอร์ปิง สามารถอธิบายเป็นแผนภูมิแสดงขั้นตอนการหาลำดับย่อยของตัวอธิบายรูปร่างได้ดังรูปที่ 2.12



425352952

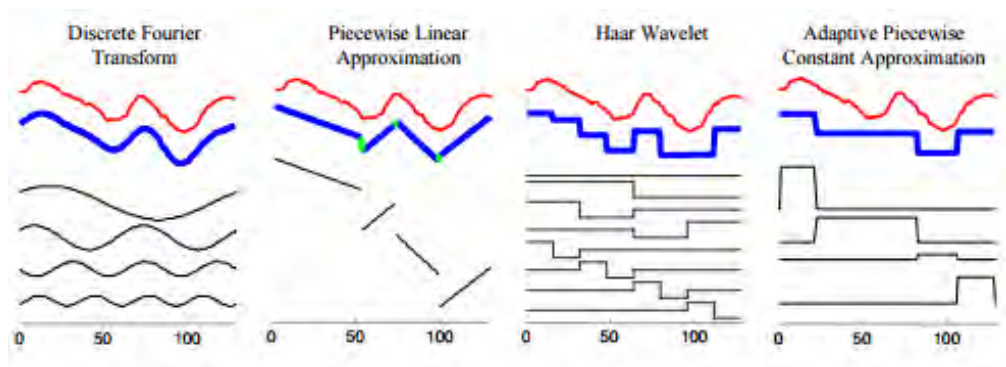


รูปที่ 2.12 แผนภูมิแสดงขั้นตอนการหาลำดับย่อยของตัวอธิบายรูปร่างภายใต้กรอบการทำงานของไดนามิกไทม์วอร์ปปีงแบบรูปร่าง

จากอัลกอริทึมไดนามิกไทม์วอร์ปปีงแบบรูปร่างจะให้ผลลัพธ์เป็นการปรับแนวระหว่าง 2 อนุกรมเวลาและระยะทางสำหรับการวัดความคล้ายคลึงกันของ 2 อนุกรมเวลาด้วยลำดับของตัวอธิบายรูปร่าง

2.2.2 งานวิจัยที่เกี่ยวข้องกับตัวอธิบายรูปร่าง

เนื่องจากไดนามิกไทม์วอร์ปิงแบบรูปร่างมีการปรับแนวในลักษณะทั่วไป โดยผู้ใช้สามารถออกแบบตัวอธิบายรูปร่างได้ และป้อนข้อมูลลงในไดนามิกไทม์วอร์ปิงแบบรูปร่าง เพื่อทำการปรับแนว ในส่วนนี้จะอธิบายถึงตัวอธิบายรูปร่างทั่วไปรูปแบบต่าง ๆ ซึ่งแต่ละตัวจะ เชื่อมโยงอนุกรมเวลาที่เก็บลำดับย่อยของอนุกรมเวลา S เข้ากับลำดับย่อยของตัวอธิบายรูปร่าง $d = F(S)$ โดยผู้ใช้สามารถเลือกตัวอธิบายรูปร่างและสามารถปรับการตั้งค่าต่าง ๆ ของตัวอธิบาย รูปร่างเพื่อให้ได้การแปลงลำดับย่อยของข้อมูลอนุกรมเวลาเป็นลำดับของตัวอธิบายรูปร่างได้ตามที่ ผู้ใช้ต้องการ โดยตัวอธิบายรูปร่างนั้นถูกนำเสนอด้วยกันหลายวิธีการ เช่น Discrete Fourier Transform, Piecewise Linear Approximation, Haar Wavelet และ Adaptive Piecewise Constant Approximation ดังรูปที่ 2.13



รูปที่ 2.13 ตัวอย่างวิธีการที่นำมาใช้ในการอธิบายรูปร่างของอนุกรมเวลา

(ที่มา : Eamonn Keogh และคณะ [13])

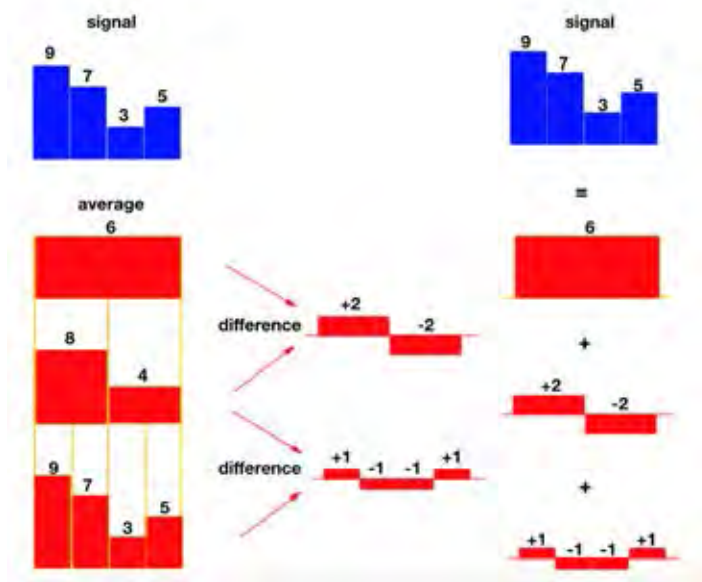
จากวิธีการที่สามารถนำมาใช้ในการอธิบายรูปร่างของลำดับย่อยของอนุกรมเวลานั้น มีหลากหลายวิธีด้วยกัน สำหรับในงานวิจัยนี้จะกล่าวถึงวิธีการเพียงบางวิธีการที่นำมาพัฒนาและใช้ เป็นแนวคิดในการพัฒนาตัวอธิบายรูปร่างในงานวิจัยนี้ ได้แก่

2.2.2.1 การแบ่งลำดับย่อยของอนุกรมเวลาแบบ Raw-subsequence

ลำดับย่อยของอนุกรมเวลาแบบ Raw-subsequence [4] จะนำค่าของลำดับย่อยของ อนุกรมเวลา s_i มาใช้เป็นลำดับย่อยของตัวอธิบายรูปร่าง โดยลำดับย่อยของอนุกรมเวลาแบบ Raw-subsequence ถูกสร้างจากจุดรอบ ๆ ของจุดเวลามบนอนุกรมเวลาที่มีความยาว l ที่กำหนด สามารถ นำลำดับย่อยมาใช้เป็นตัวอธิบายรูปร่างได้นั้นคือ $d_i = I(s_i) = s_i$ เมื่อ $I(\cdot)$ เป็นฟังก์ชัน เอกลักษณะ วิธีนี้สามารถหาลำดับย่อยรูปร่างเฉพาะส่วนได้ (local subsequence shape) ซึ่งช่วย หลีกเลี่ยงการจับคู่ของจุดที่มีค่าคล้ายกันแต่มีรูปร่างเฉพาะส่วนที่ต่างกัน

2.2.2.2 การแปลงเวฟเล็ทแบบไม่ต่อเนื่อง (Discrete Wavelet Transform-DWT)

การแปลงเวฟเล็ทแบบไม่ต่อเนื่อง [14, 15] เป็นอีกเทคนิคที่ใช้กันอย่างกว้างขวางในการประมาณค่าตัวอย่างอนุกรมเวลา เป็นวิธีการแปลงเวฟเล็ทแบบง่ายโดยจะใช้วิธีการแบบฮาร์ (Haar) โดยจะทำการแยกข้อมูลอนุกรมเวลาออกเป็น 3 ระดับ ซึ่งในระดับที่ 1 จะคำนวณค่าเฉลี่ยของทุกจุดในลำดับย่อย ในระดับที่ 2 จะแบ่งลำดับย่อยออกเป็นสองส่วนและคำนวณค่าเฉลี่ยในแต่ละส่วนสุดท้ายในระดับที่ 3 จะแบ่งลำดับย่อยออกเป็นสี่ส่วนและคำนวณค่าเฉลี่ยในแต่ละส่วนและนำค่าเฉลี่ยที่ได้จากการคำนวณตั้งแต่ระดับที่ 1 จนถึงระดับที่ 3 มาเรียงต่อกัน ซึ่งนำมาใช้เป็นลำดับย่อยของตัวอธิบายรูปร่าง d_i ของ s_i เช่น $F(.) = DWT$ โดยที่ $d_i = DWT(s_i)$ มีตัวอย่างการทำงานดังรูปที่ 2.14



รูปที่ 2.14 ตัวอย่างการทำงานการแปลงเวฟเล็ทแบบไม่ต่อเนื่อง

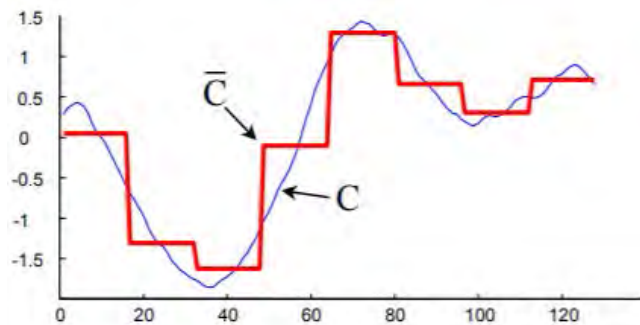
(ที่มา : Zbigniew R. Struzik และ Arno Siebes [16])

2.2.2.3 Piecewise Aggregate Approximation (PAA)

Piecewise Aggregate Approximation [17, 18] ถูกนำมาใช้เป็นตัวอย่างอธิบายรูปร่างที่ช่วยในการลดมิติของลำดับย่อยอนุกรมเวลา โดยเริ่มจากการแบ่งลำดับย่อยของอนุกรมเวลาเป็นช่วงที่เท่ากัน สำหรับในแต่ละช่วงจะถูกคำนวณค่าเฉลี่ยของทุกจุดในแต่ละส่วนนั้น ๆ เพื่อหาจุดตัวแทนของจุดทั้งหมด กำหนดให้ S คืออนุกรมเวลาที่เก็บลำดับย่อยของอนุกรมเวลา $S = s_1, s_2, \dots, s_l$ มีความยาว l โดย PAA จะทำการลดขนาดของลำดับย่อยของอนุกรมเวลาให้เหลือ w มิติ คำนวณได้จากสมการ

$$\bar{s}_i = \frac{w}{l} \sum_{j=\frac{l}{w}(i-1)+1}^{\frac{l}{w}i} s_j \quad (2.13)$$

จากสมการ 2.13 จะได้ค่า \bar{S} ที่ได้จากการลดมิติให้เหลือ w โดยที่ $\bar{S} = \bar{s}_1, \bar{s}_2, \dots, \bar{s}_w$ จากนั้นคำนวณหาเวกเตอร์ของค่าเฉลี่ยเหล่านี้จะให้ค่าประมาณของใช้ในตัวอธิบายรูปร่าง d_i ของ s_i เช่น $F(.) = PAA$ โดยที่ $d_i = PAA(s_i)$ แสดงตัวอย่างการลดมิติของข้อมูลด้วยวิธี PAA ดังรูปที่ 2.15



รูปที่ 2.15 ตัวอย่างการลดมิติของข้อมูลด้วยวิธี PAA

(ที่มา : Jessica Lin และ คณะ [19])

จากรูปที่ 2.15 แสดงตัวอย่างการลดขนาดข้อมูลของอนุกรมเวลา C ความยาว 128 มิติ ทำการลดขนาดของอนุกรมเวลา C ด้วยวิธี PAA ให้เหลือ 8 มิติ

จากตัวอธิบายรูปร่างทั้งสามแบบข้างต้นจะมีลักษณะเป็นการเข้ารหัสข้อมูลรูปร่างเฉพาะส่วน อย่างไรก็ตาม วิธีการดังกล่าวยังแปรเปลี่ยนไปตามการเปลี่ยนค่าของ y เพื่อให้เห็นภาพชัดขึ้น ยกตัวอย่างให้ลำดับย่อยของอนุกรมเวลาสองอนุกรมเวลา p และ q มีรูปร่างเหมือนกันทุกประการ แต่ p มีความสัมพันธ์ที่ค่า y เปลี่ยนไปเมื่อเทียบกับ q เช่น $p = q + \Delta \cdot l$ เมื่อ Δ คือ ขนาดความเปลี่ยนแปลงของการเปลี่ยนค่า y ดังนั้นตัวอธิบายรูปร่างภายใต้ลำดับย่อยของอนุกรมเวลา Raw-subsequence PAA และ DWT จะมีความแตกต่างกันไปตามค่า Δ โดยประมาณเช่นกัน ตัวอย่างเช่น $d(p) \approx d(q) + \Delta \cdot l$ แม้ว่าขนาดที่กล่าวถึงจะช่วยในการจำแนกประเภทอนุกรมเวลาได้ แต่ก็ยังไม่สามารถอธิบายได้ว่าลำดับย่อยของอนุกรมเวลาที่มีรูปร่างคล้ายกัน จะมีตัวอธิบายรูปร่างที่คล้ายกันด้วย ในงานวิจัย [4] จึงได้นำตัวอธิบายรูปร่าง HOG-1D ซึ่งการทำงานของ HOG-1D จะไม่แปรเปลี่ยนไปตามการเปลี่ยนค่าของ y

2.2.2.4 ฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์แบบ 1 มิติ (Histogram of Oriented Gradients 1D-HOG-1D)

ฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์แบบ 1 มิติ [5] ถูกออกแบบมาเพื่อใช้กับข้อมูลแบบ 1 มิติ ซึ่งเป็นตัวอธิบายรูปร่างที่เป็นพื้นฐานของวิทยานิพนธ์นี้ ในส่วนนี้จะอธิบายขั้นตอนการทำงานของ HOG-1D โดย HOG-1D ถูกพัฒนามาจากการนำแนวคิดหลักสำคัญจากฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์มาใช้ และปรับให้เข้ากับข้อมูลอนุกรมเวลาแบบ 1 มิติ ดังรูปที่ 1.1 ในบทที่ 1 โดยจะทำการแบ่งลำดับย่อยออกเป็นสองช่วงที่ไม่ทับซ้อนกัน และคำนวณหาขนาดและทิศทางของเกรเดียนต์เพื่อนำมาเก็บลงฮิสโตแกรม โดยการนำค่าฮิสโตแกรมของเกรเดียนต์จากทั้งสองช่วงของลำดับย่อยมาต่อกันจะได้ตัวอธิบายรูปร่างฮิสโตแกรมของทิศทางตามค่าเกรเดียนต์แบบ 1 มิติ โดยในแต่ละช่วงจะประกอบไปด้วยฮิสโตแกรมจำนวน 8 ถึง ที่มีขนาดของเกรเดียนต์เท่า ๆ กัน ตั้งแต่ -90 ถึง 90 องศา

โดยวิธีการ HOG-1D จะกำหนดให้ p คือจุดที่อยู่ภายในลำดับย่อยของอนุกรมเวลา จะได้ลำดับย่อยของอนุกรมเวลา s ที่มีความยาว l เมื่อ

$$s = p_1, p_2, \dots, p_l$$

หลังจากนั้นลำดับย่อยของอนุกรมเวลาจะถูกแบ่งออกเป็นช่วง I ที่มีความยาวคงที่ ที่ซ้อนทับหรือไม่ซ้อนทับกัน จำนวน n ช่วง เมื่อ

$$I = \{I_1, I_2, \dots, I_n\}$$

ภายในแต่ละช่วงของ I_i HOG-1D จะรวบรวมจุดของเวลาทั้งหมดใน I_i มาใส่ลงในฮิสโตแกรมและนำฮิสโตแกรมจำนวน n ช่วงมาเรียงต่อกัน เพื่อสร้างรูปแบบของตัวอธิบายรูปร่างของ s ขึ้นมาในขั้นตอนต่อไปจะกล่าวถึงรายละเอียดการนำ HOG-1D ไปใช้คำนวณในแต่ละช่วง

หลังจากแบ่งช่วงความยาวของอนุกรมตามช่วงที่กำหนดแล้ว จะคำนวณหาเกรเดียนต์ โดยกำหนดให้ช่วง $I = (p_{t_1}, p_{t_1+1}, \dots, p_{t_2})$ ด้วยช่วงเวลา (t_1, t_2) โดยคำนวณหาขนาดของเกรเดียนต์ที่จุดของเวลา t ($t_1 \leq t \leq t_2$) ได้จากสมการที่ 2.14

$$\begin{aligned} g_t &= \frac{\sigma \cdot (p_{t+1} - p_{t-1})}{(t+1) - (t-1)} \\ &= \sigma \cdot \frac{1}{2} (p_{t+1} - p_{t-1}) \end{aligned} \quad (2.14)$$



จากสมการที่ 2.14 ค่า $|g_t|$ คือขนาดของค่าเกรเดียนท์ p_t คือจุดของอนุกรมเวลาที่พิจารณา t คือจุดเวลาที่พิจารณา และ σ คือปัจจัยการปรับขนาดแบบโกลบอล (Global scaling factor) ซึ่งจะช่วยให้ปัญหาอนุกรมเวลาที่มีความถี่สูง หลังจากได้ขนาดของค่าเกรเดียนท์จากสมการที่ 2.14 โดยที่ $\arctan(g_t)$ คือทิศทางตามค่าเกรเดียนท์ ซึ่งวางตัวอยู่ในช่วง -90 ถึง 90 องศา หลังจากทราบขนาดและทิศทางตามค่าเกรเดียนท์ขั้นตอนต่อไปจะเป็นการรวบรวมค่าขนาดและทิศทางต่าง ๆ มาเก็บสะสม โดยค่าที่ได้จากการคำนวณขนาดเกรเดียนท์และทิศทางตามค่าเกรเดียนท์ของจุดอนุกรมเวลา จะถูกแทนด้วย b_i โดย i คือตำแหน่งของฮิสโตแกรมถึงที่ i^{th} ที่จะต้องนำค่าที่ได้ไปเก็บสะสม จากสมการที่ 2.15

$$b_i = |g_t| \cdot \exp \left\{ -\frac{1}{2} \frac{(\arctan(g_t) - \angle(b_i))^2}{\hat{\sigma}^2} \right\} \quad (2.15)$$

จากสมการที่ 2.15 $\angle(b_i)$ คือทิศทางของ b_i และ $\hat{\sigma}$ คือค่าปัจจัยปรับขนาดที่ใช้ระบุอัตราการลดระดับของคอร์เนล ซึ่งข้อดีของ HOG-1D จะทนทานต่อการรบกวนและไม่เปลี่ยนแปลงไปตามการเปลี่ยนขนาดของแกน y

บทที่ 3

แนวคิดและวิธีการดำเนินงาน

วิทยานิพนธ์นี้ได้นำเสนอตัวอธิบายรูปร่างภายใต้การทำงานของไดนามิกโหนดวอร์ปิงแบบรูปร่าง ซึ่งเป็นการนำแนวคิดสำคัญจากฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์แบบ 1 มิติ ซึ่งได้นำจุดเด่นของฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์มาประยุกต์ใช้สำหรับข้อมูลอนุกรมเวลาแบบ 1 มิติ เพื่อพัฒนาฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์แบบ 1 มิติให้มีประสิทธิภาพมากยิ่งขึ้น ในวิทยานิพนธ์นี้จึงได้นำแนวคิดหลักสำคัญของฮิสโตแกรมของทิศทางตามค่าเกรเดียนท์แบบ 1 มิติในการหาทิศทางและขนาดของเกรเดียนท์ และหลักการวิเคราะห์ความถดถอยมาปรับให้เข้ากับข้อมูลอนุกรมเวลา โดยแนวคิดในวิทยานิพนธ์นี้จะใช้การหาแนวเส้นความถดถอยเชิงเส้นจากการวิเคราะห์ความถดถอยของลำดับย่อยภายในพื้นที่ที่ถูกแบ่ง มาเปรียบเทียบกับทิศทางที่ได้จากจุดที่อยู่ภายใต้พื้นที่นั้น เพื่อให้มีความแม่นยำมากขึ้นเมื่อนำไปใช้เป็นตัวอธิบายรูปร่างภายใต้การทำงานของไดนามิกโหนดวอร์ปิงแบบรูปร่าง โดยแนวคิดในวิทยานิพนธ์นี้จะใช้ชื่อว่า HOG1D-L

ในบทที่ 3 จะกล่าวถึง การเตรียมข้อมูลก่อนนำข้อมูลไปใช้ในการทดลอง และการเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง

3.1 การเตรียมข้อมูล

สำหรับขั้นตอนการเตรียมข้อมูลก่อนการนำไปใช้ในการจำแนกประเภทข้อมูล จะทำการปรับข้อมูลอนุกรมให้เป็นบรรทัดฐานเดียวกันก่อน เนื่องจากข้อมูลอนุกรมเวลาอาจมีมาตราส่วนที่แตกต่างกันซึ่งอาจจะส่งผลทำให้ผลลัพธ์ในการคำนวณผิดพลาด โดยก่อนนำข้อมูลไปใช้ในการจำแนกประเภทข้อมูลหรือการวัดระยะห่างของข้อมูล จะทำการแปลงข้อมูลให้เป็นบรรทัดฐานเดียวกันก่อนด้วยการปรับมาตราส่วน (Scale) และแอมพลิจูด (Amplitude) ที่แตกต่างกัน ทำให้อยู่บนบรรทัดฐานเดียวกัน โดยในงานวิทยานิพนธ์นี้ใช้การทำให้เป็นบรรทัดฐานด้วยวิธีการ Z-Normalization ในการแปลงก่อนนำข้อมูลไปใช้ในการทดลอง

กำหนดให้อนุกรมเวลาที่ต้องการทำให้เป็นบรรทัดฐานเดียวกันคือ อนุกรมเวลา P มีความยาว n เมื่อ

$$P = p_1, p_2, \dots, p_i, \dots, p_n$$

จะทำให้เป็นบรรทัดฐานเดียวกัน โดยคำนวณจากสมการ 3.1



425352952

CD :Thesis 6070323921 thesis / rev: 06082562 14:42:31 / seq: 83

$$p'_i = \frac{p_i - \bar{p}}{SD} \quad (3.1)$$

จากสมการที่ 3.1 กำหนดให้ค่า p'_i คืออนุกรมเวลาที่ถูกรับให้เป็นบรรทัดฐานเดียวกัน p_i คืออนุกรมเวลาที่ยังไม่ถูกรับให้เป็นบรรทัดฐานกัน และ \bar{p} คือค่าเฉลี่ยของอนุกรมเวลา P ซึ่งสามารถคำนวณได้จากสมการที่ 3.2

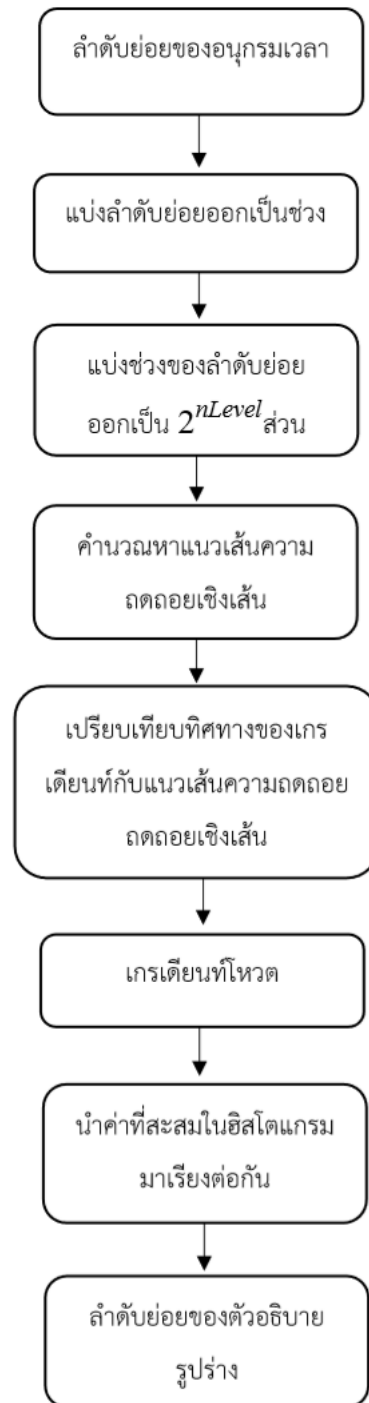
$$\bar{p} = \frac{1}{n} \sum_{i=1}^n p_i \quad (3.2)$$

จากนั้นเมื่อได้ค่าเฉลี่ยของอนุกรมเวลา P แล้วจะสามารถคำนวณหาค่า SD หรือส่วนเบี่ยงเบนมาตรฐานของอนุกรมเวลา P ได้ โดยคำนวณได้จากสมการที่ 3.3

$$SD = \sum_{i=1}^n \sqrt{\frac{(p_i - \bar{p})^2}{n}} \quad (3.3)$$

3.2 การเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง

ในวิทยานิพนธ์นี้ได้เสนอวิธีการจำแนกประเภทข้อมูลอนุกรมเวลาตัวอธิบายรูปร่าง ที่ใช้ชื่อว่า HOG1D-L โดยได้นำแนวคิดการวิเคราะห์ความถดถอยมาใช้ร่วมกับตัวอธิบายรูปร่างภายใต้กรอบการทำงานของไดนามิกไทม์วอร์ปปีงแบบรูปร่าง โดยกรอบการทำงานของตัวอธิบายรูปร่าง HOG1D-L สามารถอธิบายได้ดังรูปที่ 3.1



รูปที่ 3.1 แผนภูมิแสดงขั้นตอนการทำงานของตัวอธิบายรูปร่าง HOG1D-L

โดยแนวคิดที่วิทยานิพนธ์นี้นำเสนอคือ

3.2.1 การแบ่งลำดับย่อยออกเป็นช่วง

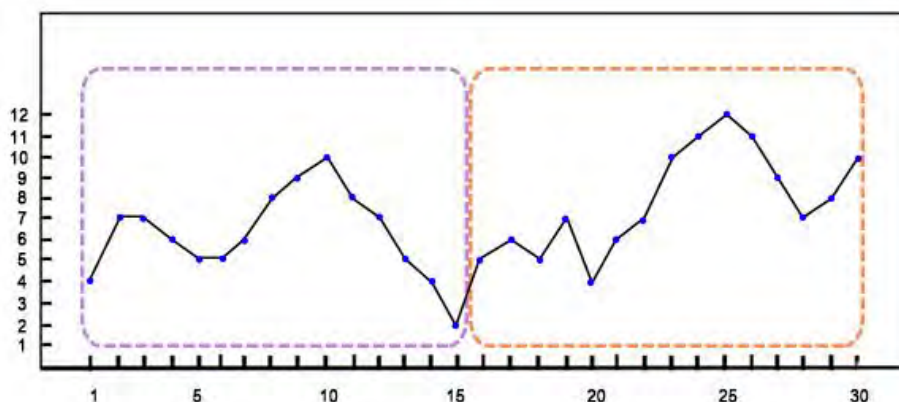
จากงานวิจัยที่เกี่ยวข้อง [4] กำหนดให้ S เป็นลำดับย่อยของอนุกรมเวลาที่มีความยาว l เมื่อ

$$S = p_1, p_2, \dots, p_l$$

โดยลำดับย่อยของอนุกรมเวลาจะถูกแบ่งความยาวออกเป็นช่วงที่ไม่ทับซ้อนกันและมีความยาวคงที่จำนวน $n = 2$ ช่วง จะได้ช่วงลำดับย่อยของอนุกรมเวลาเท่ากับ $I = \{I_1, I_2\}$ จากนั้นจะได้จุดของเวลาในแต่ละช่วง โดยกำหนดให้ช่วงของ I แสดงโดย

$$I = p_{t_1}, p_{t_1+1}, \dots, p_{t_2}$$

ซึ่งมีระยะเวลา t อยู่ระหว่าง t_1 ถึง t_2 ดังรูปที่ 3.2



รูปที่ 3.2 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรมเวลาที่มีความยาว 30

จากรูปที่ 3.2 จะทำการแบ่งลำดับย่อยของอนุกรมเวลาที่มีความยาวเท่ากับ 30 ออกเป็น 2 ช่วงที่ไม่ทับซ้อนกัน ซึ่งจะได้ความยาวในแต่ละช่วงมีความยาวเท่ากับ 15

3.2.2 การแบ่งช่วงของลำดับย่อยออกเป็น $nLevel$

ในวิทยานิพนธ์นี้ได้นำเสนอการแบ่งช่วงของ $I = \{I_1, I_2\}$ ออกเป็นจำนวนเท่ากับ 2^{nLevel} ส่วน เพื่อใช้ในการคำนวณความถดถอยเชิงเส้นสำหรับเปรียบเทียบทิศทางของเส้นความถดถอยเชิงเส้นกับทิศทางของเกรเดียนท์ภายใต้ ซึ่ง $nLevel$ คือค่าที่กำหนดขึ้นเพื่อใช้ในการแบ่งลำดับย่อยในแต่ละช่วง จากการทดลอง ได้ทำการแบ่งช่วงออกเป็น $nLevel = 1, 2, 3$ โดยผลลัพธ์ที่ดีที่สุดคือ

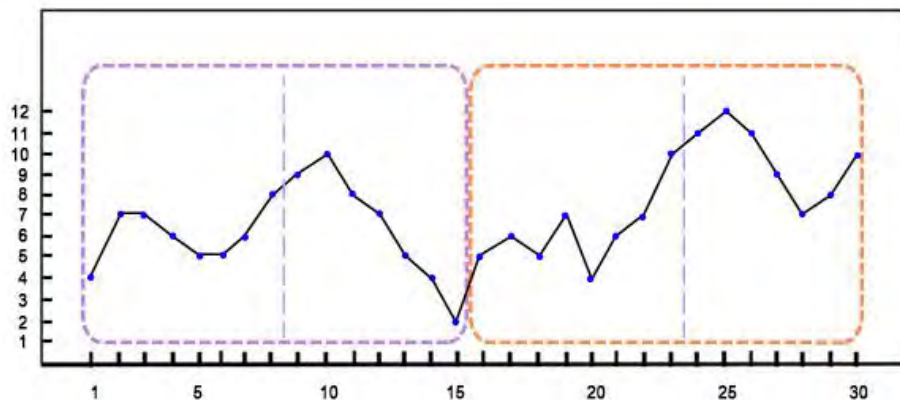
$nLevel = 1$ จะสามารถใช้งานได้ดีในการแบ่งช่วงต่าง ๆ ต่อประสิทธิภาพในการจำแนกประเภทข้อมูล ดังรูปที่ 3.3 โดยในแต่ละช่วง I จะได้

$$I_i = \{C_1, C_2, \dots, C_{2^{nLevel}}\}$$

โดย C คือส่วนที่ถูกแบ่งออกมาจากช่วง I_i มีจำนวน $L_c = \frac{l}{n \cdot 2^{nLevel}}$ ส่วน เมื่อ

$$C_i = c_1, c_2, \dots, c_{L_c}$$

กำหนดให้ c_i คือตำแหน่งจุดข้อมูลที่ i ภายใต้วงของ C



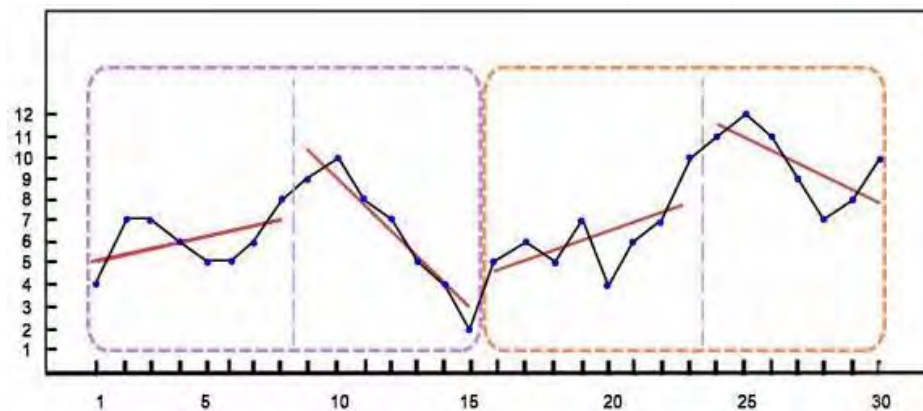
รูปที่ 3.3 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรมเวลาที่มีความยาว 30 เมื่อทำการแบ่งจุดในช่วงต่าง ๆ ของเวลาจากส่วนที่ไม่ทับซ้อนกันออกเป็น 2 ส่วน

3.2.3 การหาแนวเส้นความถดถอยเชิงเส้น

เพื่อเพิ่มประสิทธิภาพ HOG-1D ในการหาขนาดของเกรเดียนต์และทิศทางของข้อมูลที่อาจจะมีการคำนวณและผลลัพธ์ไม่สอดคล้องกับข้อมูล วิทยานิพนธ์นี้จึงได้นำเสนอการวิเคราะห์ความถดถอยมาปรับใช้กับข้อมูลอนุกรมเวลาเพื่อให้ข้อมูลที่นำไปเก็บในฮิสโตแกรมมีความแม่นยำมากขึ้น โดยใช้การหาแนวเส้นความถดถอยเชิงเส้น (Linear regression line) มาใช้เพื่อหาแนวโน้มทิศทางของอนุกรมเวลา โดยแบ่งข้อมูลอนุกรมเวลาออกเป็นช่วงย่อย ๆ ออกเป็น 2^{nLevel} ส่วน แล้วจะทำการลากแนวเส้นความถดถอยเชิงเส้นผ่านระหว่างจุดของลำดับย่อยของอนุกรมเวลาของส่วน C_i ในแต่ละช่วงคำนวณได้จากสมการที่ 3.4

$$b = \frac{\sum_{i=1}^{L_c} (c_i - \bar{c})(d_{t_i} - \bar{d}_t)}{\sum_{i=1}^{L_c} (c_i - \bar{c})^2} \quad (3.4)$$

จากสมการที่ 3.4 กำหนดให้ b คือค่าสัมประสิทธิ์ความถดถอยซึ่งเป็นตัวกำหนดความชันของเส้นสมการความถดถอยเชิงเส้น L_c คือจำนวนส่วนที่ถูกแบ่งออกมาจากช่วงของลำดับย่อย c_i คือตำแหน่งจุดข้อมูลที่ i ภายใต้อส่วนของ C ที่ถูกแบ่งออกมา \bar{c} คือค่าเฉลี่ยของตำแหน่งจุดข้อมูลทุกจุดที่อยู่ภายใต้อส่วนของ C d_{t_i} คือจุดที่อยู่ภายใต้อส่วนของ C และ \bar{d}_t คือค่าเฉลี่ยของข้อมูลอนุกรมเวลาทุกจุดข้อมูลที่อยู่ภายใต้อส่วนของ C ดังรูปที่ 3.4



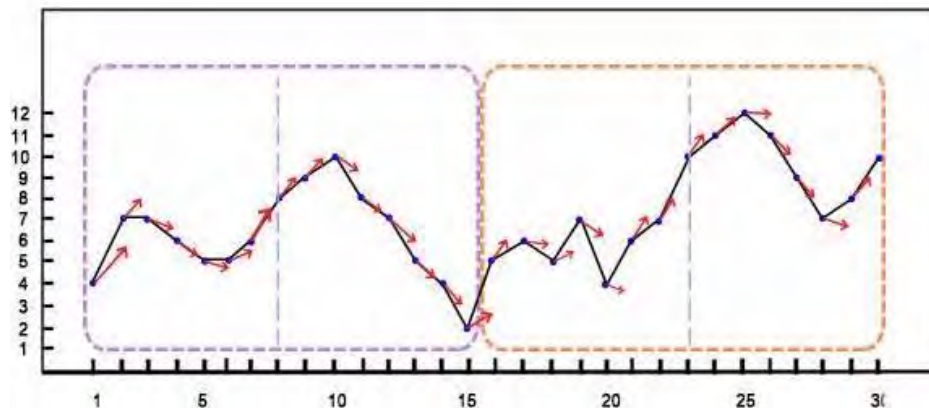
รูปที่ 3.4 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรมเวลาที่มีความยาว 30 ในการคำนวณหาแนวเส้นความถดถอยเชิงเส้น จากจุดต่าง ๆ ของเวลาในแต่ละช่วง

3.2.4 การหาขนาดและทิศทางของเกรเดียนท์

หลังจากที่ได้เส้นทิศทางของแนวเส้นความถดถอยเชิงเส้นจากสมการที่ 3.4 แล้ว ขั้นตอนต่อไปเป็นการคำนวณหาขนาดของเกรเดียนท์ (Gradient Magnitude) และการหาทิศทางตามค่าเกรเดียนท์ (Gradient Orientation) ของจุดในแต่ละส่วนจากการคำนวณ ดังสมการที่ 3.5

$$g_t = \frac{(p_{t+1} - p_{t-1})}{\sigma \cdot (t+1) - (t-1)} \quad (3.5)$$

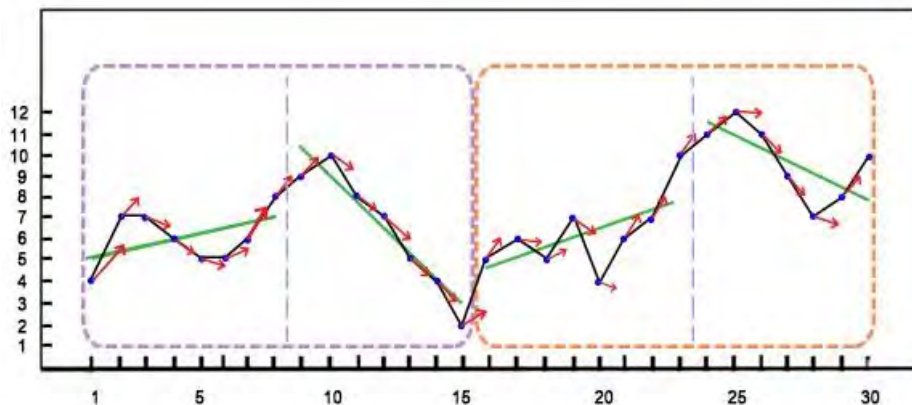
กำหนดให้ $\arctan(g_t)$ คือทิศทางตามค่าเกรเดียนท์ p คือจุดที่อยู่ภายในลำดับย่อยของอนุกรมเวลาช่วง $I = (p_{t_1}, p_{t_1+1}, \dots, p_{t_2})$ ด้วยช่วงเวลา (t_1, t_2) ดังรูปที่ 3.5 โดยที่ σ คือปัจจัยการปรับขนาดแบบโกลบอล (Global scaling factor) ซึ่งจะช่วยแก้ปัญหาอนุกรมเวลาที่มีความถี่สูง เมื่อค่าของจุดข้อมูลที่ p_t และ p_{t+1} มีค่าใกล้เคียงกันมากทำให้ขนาดของเกรเดียนท์จะมีค่าใกล้เคียง 0 เมื่อ σ ถูกกำหนดค่าให้เท่ากับ 1 ดังนั้นในงานวิจัย [4] จึงได้กำหนดให้ σ มีค่าเท่ากับ 0.1 เพื่อแก้ปัญหาค่าของจุดข้อมูลที่ส่งผลให้ขนาดของเกรเดียนท์ที่มีค่าใกล้เคียง 0



รูปที่ 3.5 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรมเวลาที่มีความยาว 30 ที่แสดงทิศทางการวางตัวของค่าเกรเดียนท์ในแต่ละจุดต่าง ๆ

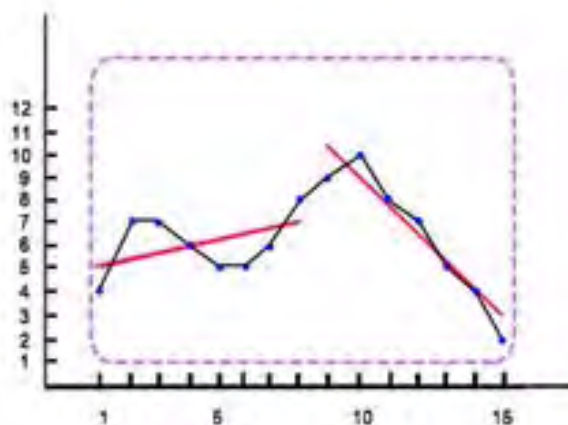
3.2.5 เปรียบเทียบแนวเส้นความถดถอยของการวิเคราะห์ความถดถอย

เมื่อได้ขนาดของเกรเดียนท์และทิศทางตามค่าเกรเดียนท์แล้ว จะนำทิศทางตามค่าเกรเดียนท์ ณ จุดนั้น ๆ ที่คำนวณได้เทียบกับเส้นทิศทางของเส้นความถดถอยเชิงเส้นของชุดข้อมูลช่วงนั้น หากทิศทางของเส้นความถดถอยเชิงเส้นมีค่ามากกว่าทิศทางเกรเดียนท์ของจุดข้อมูลเกิน 45 องศา จะไม่ถูกนำไปคิดรวมในฮิสโตแกรมของช่วงนั้น ดังรูปที่ 3.6



รูปที่ 3.6 ภาพตัวอย่างการทำงานของตัวอธิบายรูปร่าง HOG1D-L ของลำดับย่อยของอนุกรมเวลาที่มีความยาว 30 ที่มีการเปรียบเทียบทิศทางของแนวเส้นความถดถอยเชิงเส้นในช่วงนั้น

จากรูปที่ 3.6 ขนาดของเกรเดียนท์และทิศทางตามค่าเกรเดียนท์ ณ จุดของเวลาที่ 2 จะถูกเก็บรวบรวมในฮิสโตแกรมของเกรเดียนท์ เนื่องจากมีทิศทางสอดคล้องกันกับแนวเส้นความถดถอยเชิงเส้นในช่วง I_i



1.500	1.500	-0.500	-1.000
-0.250	0.500	1.250	1.500

ขนาดของเกรเดียนต์ภายใต้ส่วนที่หนึ่ง

56.31	56.31	-26.57	-45.00
-14.04	26.57	51.34	56.31

ทิศทางของเกรเดียนต์ภายใต้ส่วนที่หนึ่ง

1.000	-0.500	-1.500	-1.500
-1.500	-1.500	0.500	

ขนาดของเกรเดียนต์ภายใต้ส่วนที่สอง

45.00	-26.57	-56.31	-56.31
-56.31	-56.31	26.57	

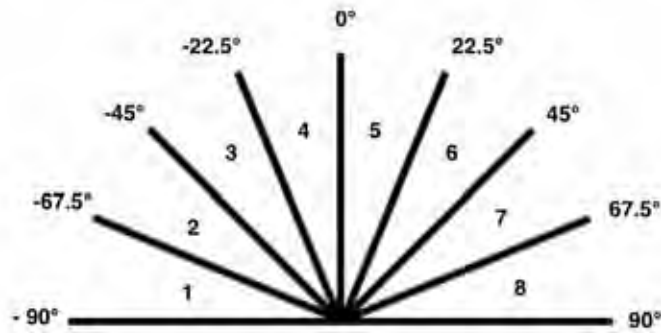
ทิศทางของเกรเดียนต์ภายใต้ส่วนที่สอง

รูปที่ 3.7 แสดงขนาดและทิศทางของเกรเดียนต์ และเส้นความถดถอยเชิงเส้นภายใต้ส่วนทั้ง 2 ส่วน

จากรูปที่ 3.7 แสดงขนาดและทิศทางของเกรเดียนต์ และเส้นความถดถอยเชิงเส้นซึ่งเป็นแนวโน้มของทิศทางของจุดภายใต้ส่วนทั้ง 2 ส่วน เพื่อใช้สำหรับการเปรียบเทียบกับทิศทางภายใต้ส่วนนั้น ๆ โดยเส้นความถดถอยเชิงเส้นภายใต้ส่วนที่หนึ่งมีทิศทางประมาณ 11.94 องศา ทำให้จุดที่ 4 ของส่วนที่หนึ่ง ซึ่งมีทิศทางของเกรเดียนต์เท่ากับ -45 องศา แตกต่างจากทิศทางของเส้นความถดถอยเชิงเส้นมากกว่า 45 องศา ทำให้จุดที่ 4 จะไม่ถูกนำไปคำนวณในเกรเดียนต์โหวต

3.2.6 เกรเดียนต์โหวต

การรวบรวมทิศทางต่าง ๆ และเก็บค่าเกรเดียนต์ของจุดในแต่ละช่วงลงในฮิสโตแกรมนั้น จะทำการแบ่งฮิสโตแกรมออกเป็นระยะเท่า ๆ กัน ตั้งแต่ -90 ถึง 90 องศา โดยจะมีจำนวนถัง (Bin) ของฮิสโตแกรมที่เก็บทิศทางในแต่ละช่วงจำนวน 8 ถัง ดังรูปที่ 3.8



รูปที่ 3.8 ถังเก็บทิศทางของฮิสโตแกรมขนาด -90 ถึง 90 องศา

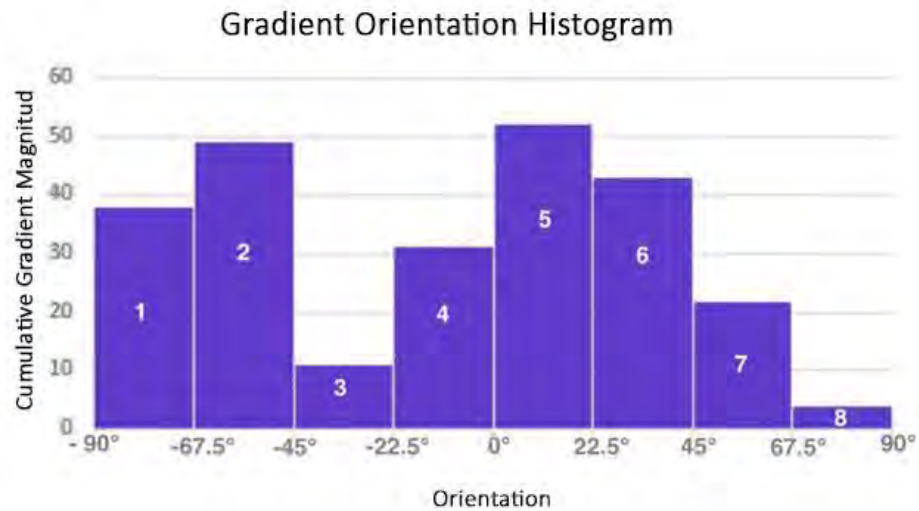
โดยขั้นตอนการนำขนาดของเกรเดียนท์ ($|g_t|$) และทิศทางของเกรเดียนท์ $\arctan(g_t)$ ของแต่ละจุดอนุกรมเวลา เข้ามาเก็บสะสมไว้ในถังของฮิสโตแกรม ขั้นตอนนี้เรียกว่า การโหวต (Voting) ซึ่งค่าที่ได้จากการคำนวณขนาดเกรเดียนท์และทิศทางตามค่าเกรเดียนท์ของจุดอนุกรมเวลา จะถูกแทนด้วย v_j โดย j คือตำแหน่งของฮิสโตแกรมถึงที่ j^{th} ที่จะต้องนำค่าที่ได้ไปเก็บสะสม ในกรณีที่ทิศทางของจุดนั้นอยู่ตั้งแต่ -90 องศา ถึงมุมศูนย์กลาง (Center angle) ของฮิสโตแกรมแรกและในกรณีที่ทิศทางของจุดมากกว่ามุมศูนย์กลางของฮิสโตแกรมสุดท้ายถึง 90 องศา จะใช้สมการที่ 3.6

$$v_j = |g_t| \quad (3.6)$$

แต่ในกรณีที่ทิศทางของจุดนั้นมีมุมมากกว่ามุมศูนย์กลางของฮิสโตแกรมแรกและมีมุมน้อยกว่าหรือเท่ากับมุมศูนย์กลางของฮิสโตแกรมสุดท้ายจะต้องคำนวณทั้งสองสมการ ดังสมการที่ 3.7

$$\begin{aligned} v_j &= |g_t| \cdot \cos(\alpha_t - \theta_j) \\ v_{j+1} &= |g_t| \cdot \cos(\theta_{j+1} - \alpha_t) \end{aligned} \quad (3.7)$$

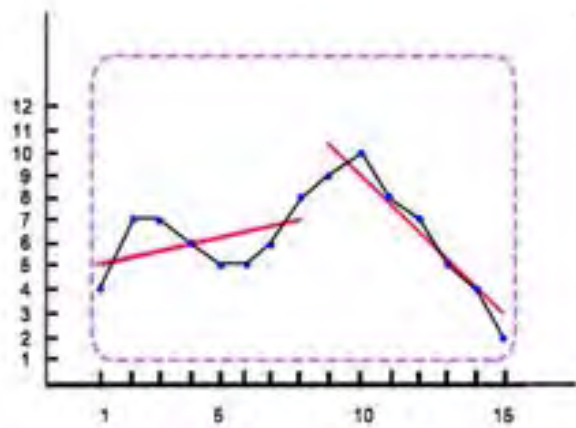
จากสมการ 3.8 กำหนดให้ θ_j คือมุมศูนย์กลางของฮิสโตแกรมที่ j^{th} ที่อยู่ติด α_t และมีค่าน้อยกว่า α_t โดย α_t คือทิศทางตามค่าเกรเดียนท์ของจุดอนุกรมเวลา แสดงการคำนวณดังรูปที่ 3.9



รูปที่ 3.9 ตัวอย่างฮิสโตแกรมของเกรเดียนท์

จากรูปที่ 3.10 แสดงวิธีการคำนวณเกรเดียนท์โหวตของจุดภายในช่วง โดยตารางขนาดของเกรเดียนท์และตารางทิศทางของเกรเดียนท์จะประกอบไปด้วยขนาดและทิศทางของเกรเดียนท์ของแต่ละจุดภายในช่วง ซึ่งขนาดและทิศทางของเกรเดียนท์สีแดงคือขนาดและทิศทางที่ถูกตัดออกหลังจากขั้นตอนการเปรียบเทียบแนวเส้นความถดถอยเชิงเส้น ทำให้เหลือขนาดและทิศทางของเกรเดียนท์ที่จะถูกนำไปคำนวณเกรเดียนท์โหวต 12 จุด จากรูป 3.10 ได้แสดงวิธีคำนวณเกรเดียนท์โหวตของจุดที่ 5 และ 7 โดยในจุดที่ 5 มีขนาดของเกรเดียนท์เท่ากับ -0.250 และมีทิศทางของเกรเดียนท์ที่ -14.04 องศา ซึ่งทิศทางที่อยู่มากกว่ามุมศูนย์กลางของฮิสโตแกรมแรก และมีค่าน้อยกว่าหรือเท่ากับมุมศูนย์กลางของฮิสโตแกรมสุดท้าย จะใช้สมการที่ 3.7 จากเงื่อนไขมุมศูนย์กลางของฮิสโตแกรมที่อยู่ติดและมีค่าน้อยกว่ามุม -14.04 คือมุม -33.75 ซึ่งเป็นมุมศูนย์กลางของฮิสโตแกรมถึงที่ 3 ทำให้สามารถคำนวณเกรเดียนท์โหวตได้ดังนี้

$$v_3 = -0.250 \cdot \cos(-14.04 - \theta_3) \text{ และ } v_4 = -0.250 \cdot \cos(\theta_4 - (-14.04))$$



1.500	1.500	-0.500	-1.000	-0.250
0.500	1.250	1.500	1.000	-0.500
-1.500	-1.500	-1.500	-1.500	0.500

ขนาดของเกรเดียนท์

56.31	56.31	-26.57	-45.00	-14.04
26.57	51.34	56.31	45.00	-26.57
-56.31	-56.31	-56.31	-56.31	26.57

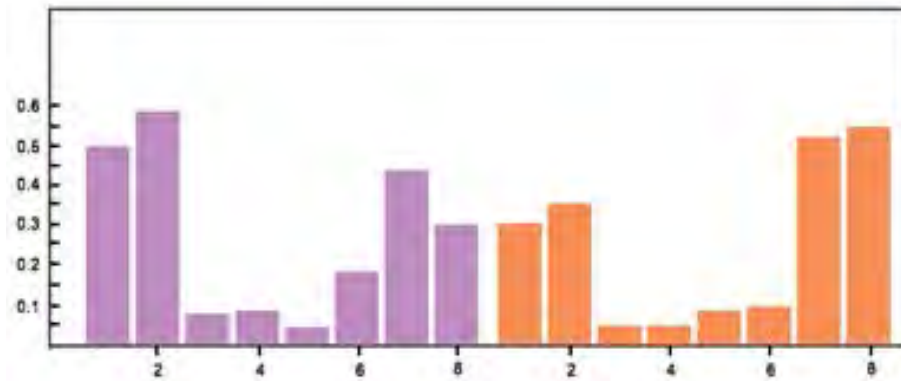
ทิศทางของเกรเดียนท์

$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
$\theta_1 = -78.75$	$\theta_2 = -56.25$	$\theta_3 = -33.75$	$\theta_4 = -11.25$	$\theta_5 = 11.25$	$\theta_6 = 33.75$	$\theta_7 = -56.25$	$\theta_8 = 78.75$
		$-0.25 \cdot \cos(-14.04 - \theta_3)$	$-0.25 \cdot \cos(-14.04 - \theta_4)$		$1.25 \cdot \cos(51.34 - \theta_6)$	$1.25 \cdot \cos(51.34 - \theta_7)$	

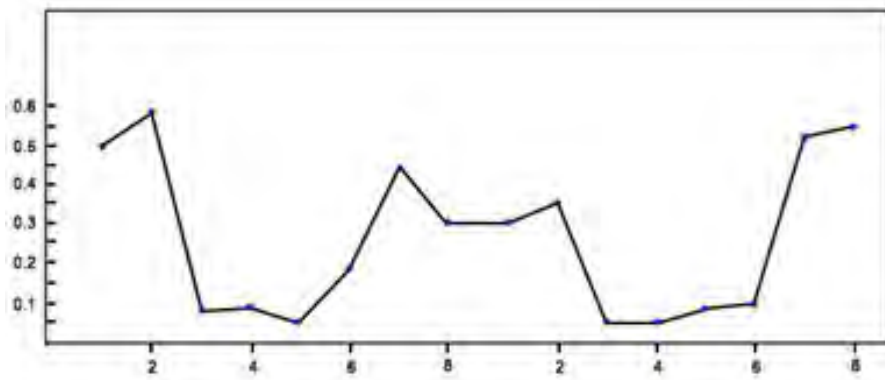
รูปที่ 3.10 ตัวอย่างการคำนวณเกรเดียนท์โหวตของจุดที่ 5 และจุดที่ 7

3.2.7 การนำค่าในฮิสโตแกรมของเกรเดียนท์มาเรียงต่อกัน

เมื่อรวบรวมค่าในฮิสโตแกรมของเกรเดียนท์แล้ว ดังรูปที่ 3.11 จากนั้นจะนำค่าฮิสโตแกรมของเกรเดียนท์จากทั้งสองลำดับย่อยมาต่อกัน ซึ่งค่าที่ได้นั้นจะกลายเป็นลำดับย่อยของตัวอธิบายรูปร่าง (Shape descriptor subsequence) แทนด้วย d_i ดังรูปที่ 3.12



รูปที่ 3.11 ตัวอย่างแห่งฮิสโตแกรมที่เก็บสะสมทิศทางตามค่าเกรเดียนท์ของจุดลำดับย่อย



รูปที่ 3.12 ลำดับย่อยของตัวอธิบายรูปร่าง

หลังจากที่ทำการแปลงอนุกรมเวลาที่เก็บลำดับย่อย S ให้กลายเป็นลำดับของตัวอธิบายรูปร่าง $d = (d_1, d_2, \dots, d_l)$ จากนั้นจะนำไดนามิกไทม์วอร์ปิงมาใช้เพื่อปรับแนวระหว่างลำดับของตัวอธิบายรูปร่างสองลำดับ โดยการหาระยะห่างระหว่างจุดในไดนามิกไทม์วอร์ปิงจะถูกแทนที่ด้วยการหาระยะห่างแบบยูคลิดระหว่างลำดับย่อยของตัวอธิบายรูปร่างในลำดับของตัวอธิบายรูปร่างและจะส่งผ่านวิธีการวอร์ปิงนี้กลับไปยังอนุกรมเวลาตั้งต้นของอนุกรมเวลา โดยวิธีการไดนามิกไทม์วอร์ปิง

3.2.7 อัลกอริทึมการเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้ หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง

ตารางที่ 3.1 อัลกอริทึม HOG1D-L

Algorithm 5 HOG1D-L	
Input: subsequence s	
Output: descriptorSequence	
1	$nInterval \leftarrow 2$ ($nInterval$ must be a number which is a subsequence length divisible)
2	$sCell \leftarrow s \text{ length} / nInterval$
3	$nLevel \leftarrow 1$
4	$dxScale \leftarrow 0.1$
5	$orientation \leftarrow 0.7854$ (comparative orientation in radians unit)
6	$descriptor [nInterval] [nbins]$
7	$descriptorSequence [nInterval \times nbins]$
8	for $i = 1$ to $nInterval$
9	for $j = 1$ to $sCell$
10	$idx \leftarrow (i - 1) * sCell + j$
11	$sLevel = \text{round}\left(\frac{sCell}{2^{nLevel}}\right)$
12	$edxRegression \leftarrow idx + sLevel - 1$
13	if $j - 1 \bmod sLevel = 0$ then
14	if $edxRegression > sCell$ then
15	$edxRegression \leftarrow sCell$
16	$sLevel \leftarrow edxRegression - idx + 1$
17	end if
18	for $k = 1$ to $sLevel$
19	$dxRegression[k] \leftarrow dxScale * k$
20	end for
21	$dyRegression \leftarrow [j : sLevel]$

22	$orientationRegression \leftarrow regression(dyRegression, dxRegression)$
23	end if
24	$sidx \leftarrow idx - 1$
25	$eidx \leftarrow idx + 1$
26	if $sidx < 1$ or $eidx > s$ length then
27	continue
28	end if
29	$dx \leftarrow 2 * dxScale$
30	$dy \leftarrow s[eidx] - s[sidx]$
31	$ang \leftarrow atan2(dy, dx)$
32	$mag \leftarrow \frac{dy}{dx}$
33	if $(orientationRegression - ang) > orientation$ then
34	continue
35	end if
36	$descriptor \leftarrow gradientVoting(descriptor, i, mag, ang)$ # Add gradient voting value to descriptor
37	end for
38	end for
39	for $j = 1$ to $nInterval$
40	for $k = 1$ to $nbins$
41	$descriptorSequence((j - 1) * nbins + k) \leftarrow descriptor(j, k)$
42	end for
43	end for

จากตาราง 3.1 อัลกอริทึมนี้เป็นอัลกอริทึมเพิ่มประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลาโดยใช้หลักการวิเคราะห์ความถดถอยร่วมกับแนวคิดของตัวอธิบายรูปร่าง HOG1D-L โดย

นำเข้าข้อมูล กำหนดให้ s คือตัวแปรที่เก็บค่าลำดับย่อยของอนุกรมเวลา และจะให้ผลลัพธ์เป็นลำดับย่อยของตัวอธิบายรูปร่าง

บรรทัดที่ 1 - 7 กำหนดตัวแปรต่าง ๆ ที่ใช้ดังนี้ กำหนดให้ตัวแปร s คือตัวแปรที่เก็บค่าลำดับย่อยของอนุกรมเวลา ตัวแปร $nInterval$ คือจำนวนช่วงของลำดับย่อย ตัวแปร $sCell$ คือจำนวนของข้อมูลในแต่ละช่วง ซึ่งมีค่าเท่ากับความยาวของลำดับย่อยหารด้วยจำนวนช่วง ตัวแปร $nLevel$ คือพารามิเตอร์ที่กำหนดเพื่อใช้ในการแบ่งช่วงออกเป็น 2^{nLevel} ส่วน ตัวแปร $dxScale$ คือปัจจัยการปรับขนาดแบบโกลบอล ตัวแปร $orientation$ คือค่าความต่างที่กำหนดสำหรับการเลือกทิศทางของเกรเดียนท์ของจุดสำหรับการนำเข้าฟังก์ชัน $gradientVoting$ ตัวแปร $descriptor$ คืออาร์เรย์ 2 มิติ ขนาด $nInterval$ และขนาด $nbins$ ใช้สำหรับเก็บสะสมผลรวมของขนาดเกรเดียนท์ในถังฮิสโตแกรมต่าง ๆ ในแต่ละช่วง และตัวแปรอาร์เรย์ $descriptorSequence$ ขนาด $nInterval$ และขนาด $nbins$ ใช้สำหรับเก็บค่าของฮิสโตแกรมทุกช่วงที่เรียงต่อกัน

บรรทัดที่ 10 - 12 ตัวแปร idx คือตัวแปรที่ใช้สำหรับการอ้างอิงจุดแรกของข้อมูลในลำดับย่อย ตัวแปร $sLevel$ คือจำนวนของข้อมูลในแต่ละส่วน $edxRegression$ คือ ตัวแปรที่ใช้สำหรับอ้างอิงจุดสุดท้ายของข้อมูลในลำดับย่อยที่จะใช้คำนวณความถดถอยเชิงเส้น

บรรทัดที่ 13 - 22 เมื่อค่า $j - 1$ สามารถหารจำนวน $sLevel$ แล้วเหลือเศษ 0 ได้ จะทำการคำนวณหาเส้นความถดถอยเชิงเส้นโดยเริ่มจากจุดข้อมูลที่ $j - 1$ จนถึงจุดที่ $j - 1$ บวกกับจำนวน $sLevel$ แต่ถ้าบวกกันแล้วมีค่ามากกว่า $sCell$ หรือมากกว่าขนาดของลำดับย่อย จะลดขนาดของ $sLevel$ ให้พอดีกับจุดที่เหลือที่เพื่อไม่ให้เกินจากส่วนที่ถูกแบ่ง โดย $orientationRegression$ คือตัวแปรที่เก็บทิศทางของเส้นความถดถอยเชิงเส้นไว้ใช้สำหรับเปรียบเทียบกับทิศทางของจุดต่าง ๆ ภายใต้อันส่วน

บรรทัดที่ 24 - 32 เป็นการคำนวณเพื่อหาขนาดและทิศทางเกรเดียนท์ของแต่ละจุด โดยจะใช้วิธีการนำจุดที่อยู่ถัดไปจากจุดนั้นลบด้วยจุดก่อนหน้าแล้วหารด้วยปัจจัยการปรับขนาดแบบโกลบอลคูณด้วย 2 จะทำให้ได้ขนาดของเกรเดียนท์ mag ของจุดนั้น และหาทิศทางของเกรเดียนท์ α จากฟังก์ชัน $atan2$ ซึ่งมีค่าอยู่ระหว่าง $-\pi$ และ π

บรรทัดที่ 33 - 34 เมื่อได้ขนาดและทิศทางของเกรเดียนท์แล้ว จะนำค่าทิศทางของ α มาเปรียบเทียบกับทิศทางเส้นความถดถอยเชิงเส้น $OrientationRegression$ ถ้าทิศทางของ $OrientationRegression$ มากกว่าทิศทางของเกรเดียนท์เกิน 45 องศา จะไม่นำขนาดของเกรเดียนท์จากจุดนั้นไปคำนวณในขั้นตอนถัดไป มิฉะนั้นจะนำขนาดและทิศทางของเกรเดียนท์นั้นเพื่อไปคำนวณเกรเดียนท์โหวต



4253522952

CD :Thesis 6070323221 thesis / revv: 06082562 14:42:31 / seq: 83

บรรทัดที่ 36 นำขนาดและทิศทางของเกรเดียนท์ไปคำนวณเกรเดียนท์โหวต เพื่อนำขนาดของเกรเดียนท์ไปเก็บสะสมไว้ในถังฮิสโตแกรมตามทิศทางของเกรเดียนท์นั้น โดยจะอยู่ในรูปของตัวแปร *descriptor*

บรรทัดที่ 39 - 43 นำค่าผลรวมขนาดของเกรเดียนท์ที่อยู่ในฮิสโตแกรมแต่ละถังของทั้ง 2 ช่วงของลำดับย่อยมาเรียงต่อกัน จะได้ลำดับของตัวอธิบายรูปร่าง (Shape descriptor sequence) ซึ่งในที่นี้จะมีขนาดความยาวเท่ากับ 16 ซึ่งเท่ากับจำนวนของถังฮิสโตแกรมของทั้งสองช่วงของลำดับย่อย

ตารางที่ 3.2 อัลกอริทึม histogramVoting

Algorithm 6 histogramVoting(descriptor, i, mag, ang)	
Inputs: An array descriptor , index of interval i , magnitude of gradient mag , orientation of gradient ang	
Output: An array descriptor which is cumulated magnitude of gradient	
1	$nbins \leftarrow 8$
2	$angles \leftarrow [-1.5708, -1.1781, -0.7854, -0.3927, 0, 0.3927, 0.7854, 1.1781, 1.57080]$
3	for $j=1$ to $angles$ length
4	$centerAngles[j] \leftarrow (angles[j] + angles[j+1])/2$
5	end for
6	for $j = 1$ to $angles$ length - 1
7	if $angles[j] \leq ang$ and $angles[j+1] > ang$ then
8	$n \leftarrow j$
9	else if $angles[j] > ang$ then
10	$n \leftarrow 1$
11	else if $angles[angles$ length] $< ang$ then
12	$n \leftarrow angles$ length - 1
13	end if
14	end for
15	if $n == 1$ and $ang \leq centerAngles[1]$ then
16	$descriptor[i, n] \leftarrow descriptor[i, n] + abs(mag)$



425352952

17	else if $n == nbins$ and $ang > centerAngles[nbins]$ then
18	$descriptor[i, n] \leftarrow descriptor[i, n] + abs(mag)$
19	else
20	if $abs(angles(n) - ang) > abs(angles(n+1) - ang)$ then
21	$ang1 \leftarrow centerAngles[n]$
22	$ang2 \leftarrow centerAngles[n+1]$
23	$descriptor[i, n] \leftarrow descriptor[i, n] + abs(mag) * \cos(ang1 - ang)$
24	$descriptor[i, n+1] \leftarrow descriptor[i, n+1] + abs(mag) * \cos(ang2 - ang)$
25	else
26	$ang1 \leftarrow centerAngles[n - 1]$
27	$ang2 \leftarrow centerAngles[n]$
28	$descriptor[i, n-1] \leftarrow descriptor[i, n-1] + abs(mag) * \cos(ang1 - ang)$
29	$descriptor[i, n] \leftarrow descriptor[i, n] + abs(mag) * \cos(ang2 - ang)$
30	end if
31	end if

จากตาราง 3.2 อัลกอริทึมนี้เป็นอัลกอริทึมที่ใช้ในการการรวบรวมทิศทางต่าง ๆ และเก็บค่าเกรเดียนท์ของจุดในแต่ละช่วงลงในฮิสโตแกรม โดยนำเข้าข้อมูล กำหนดให้ ตัวแปร *descriptor* คือ อาร์เรย์สองมิติที่เก็บผลรวมของขนาดเกรเดียนท์ ตัวแปร *i* คือช่วงของลำดับย่อย ตัวแปร *mag* คือขนาดของเกรเดียนท์และ ตัวแปร *ang* คือทิศทางของเกรเดียนท์

บรรทัดที่ 1 - 2 กำหนดตัวแปร *nbins* ซึ่งเป็นจำนวนถังของฮิสโตแกรม และ *angles* ที่เก็บค่าของมุมตั้งแต่ $-\pi$ ถึง π ที่ถูกแบ่งออกเป็น 9 ส่วนเท่า ๆ กันในหน่วยเรเดียนท์

บรรทัด 3 - 5 ให้ตัวแปร *centerAngles* เก็บค่ามุมศูนย์กลางของถังในฮิสโตแกรม จากการนำค่าเฉลี่ยของค่ามุมที่อยู่ติดกันสองค่าใน *angles*

บรรทัด 4 นำค่า *ang* มาตรวจสอบว่าค่าของทิศทางของเกรเดียนท์อยู่ในถังฮิสโตแกรมถึงใด แทนด้วยตัวแปร *n*

บรรทัด 15 - 16 เมื่อได้ตำแหน่งถังฮิสโตแกรมของเกรเดียนท์แล้ว จากนั้นจะทำการหาค่าเกรเดียนท์ลงถังฮิสโตแกรมตามเงื่อนไข เมื่อ *ang* อยู่ในถังฮิสโตแกรมที่ 1 และมีค่า *ang* น้อยกว่าหรือ

เท่ากับมุมศูนย์กลางของฮิสโตแกรมที่ 1 จะนำค่าสัมบูรณ์ของขนาดเกรเดียนท์เพิ่มเข้าไปในตัวแปร *descriptor* ที่แถว *i* หลักที่ 1

บรรทัดที่ 17 - 18 ถ้า *ang* อยู่ที่ถึงฮิสโตแกรมที่ถึงสุดท้าย หรือถึงที่ 8 และมีค่า *ang* มากกว่ามุมศูนย์กลางของฮิสโตแกรมสุดท้าย จะนำค่าสัมบูรณ์ของขนาดเกรเดียนท์เพิ่มเข้าไปในตัวแปร *descriptor* ที่แถว *i* หลักที่ 8

บรรทัดที่ 19 - 29 นอกจากสองเงื่อนไขที่กล่าวไป หรือค่า *ang* มีค่ามากกว่ามุมศูนย์กลางของฮิสโตแกรมถึงที่ 1 และน้อยกว่าหรือเท่ากับมุมศูนย์กลางของฮิสโตแกรมถึงสุดท้าย จากนั้นถ้าค่าสัมบูรณ์ของ *ang* ลบกับ *angles* หลักที่ *n* มากกว่าค่าสัมบูรณ์ของ *ang* ลบกับ *angles* หลักที่ *n* + 1 จากนั้น กำหนดให้ *ang1* คือมุมศูนย์กลางของฮิสโตแกรมที่ *n* และ *ang2* คือมุมศูนย์กลางของฮิสโตแกรมที่ *n*+1 จากนั้นนำค่าสัมบูรณ์ของขนาดเกรเดียนท์คูณกับ cosine ของ *ang1* ลบด้วย *ang* เพิ่มเข้าไปในตัวแปร *descriptor* ที่แถว *i* หลักที่ *n* และนำค่าสัมบูรณ์ของขนาดเกรเดียนท์คูณกับ cosine ของ *ang2* ลบด้วย *ang* เพิ่มเข้าไปในตัวแปร *descriptor* ที่แถว *i* หลักที่ *n* + 1 และถ้าค่าสัมบูรณ์ของ *ang* ลบกับ *angles* หลักที่ *n* น้อยกว่าหรือเท่ากับค่าสัมบูรณ์ของ *ang* ลบกับ *angles* หลักที่ *n*+1 จากนั้นกำหนดให้ *ang1* คือมุมศูนย์กลางของฮิสโตแกรมที่ *n* - 1 และ *ang2* คือมุมศูนย์กลางของฮิสโตแกรมที่ *n* จากนั้นนำค่าสัมบูรณ์ของขนาดเกรเดียนท์คูณกับ cosine ของ *ang1* ลบด้วย *ang* เพิ่มเข้าไปในตัวแปร *descriptor* ที่แถว *i* หลักที่ *n* - 1 และนำค่าสัมบูรณ์ของขนาดเกรเดียนท์คูณกับ cosine ของ *ang2* ลบด้วย *ang* เพิ่มเข้าไปในตัวแปร *descriptor* ที่แถว *i* หลักที่ *n*



4253522952

บทที่ 4

การทดลองและการวิเคราะห์ผล

ในบทนี้จะกล่าวถึงการประเมินประสิทธิผลจากการทดลองที่วิทยานิพนธ์นี้ได้นำเสนอ โดยจะเปรียบเทียบกับสิ่งที่นำเสนอกับมาตรฐานระยะทางไดนามิกไทม์วอร์ปิง และ ไดนามิกไทม์วอร์ปิง แบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D โดยใช้การจำแนกประเภทข้อมูลแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว

4.1 ชุดข้อมูลที่ใช้ในการทดลอง

ข้อมูลอนุกรมเวลาที่ใช้ในการทดลองในวิทยานิพนธ์นี้ ประกอบด้วยข้อมูลอนุกรมเวลาจากทั้งหมด 78 ชุดจาก UCR Time Series Classification Archive [6] ซึ่งเป็นข้อมูลอนุกรมเวลาที่มีการเปิดเผยเพื่อใช้สำหรับงานวิจัยที่เกี่ยวข้องกับอนุกรมเวลาและมีความหลากหลาย โดยข้อมูลดังกล่าวถูกแบ่งออกเป็นข้อมูลสอนและข้อมูลทดสอบ ซึ่งประกอบด้วย จำนวนคลาส จำนวนข้อมูลสอน ข้อมูลทดสอบ และความยาว ซึ่งแตกต่างกันออกไปตามชุดของข้อมูลอนุกรมเวลา ดังตารางที่ 4.1

ตารางที่ 4.1 ตารางแสดงรายละเอียดของชุดข้อมูลที่ใช้ในการทดลอง

ชื่อชุดข้อมูล	จำนวน คลาส	จำนวน ข้อมูลสอน	จำนวนข้อมูล ทดสอบ	ความยาวอนุกรม เวลา
50Words	50	450	455	270
Adiac	37	390	391	176
ArrowHead	3	36	175	251
Beef	5	30	30	470
BeetleFly	2	20	20	512
BirdChicken	2	20	20	512
Car	4	60	60	577
CBF	3	30	900	128
ChlorineConcentration	3	467	3840	166
CinC_ECG_torso	4	40	1380	1639

ชื่อชุดข้อมูล	จำนวน คลาส	จำนวน ข้อมูลสอน	จำนวนข้อมูล ทดสอบ	ความยาวอนุกรม เวลา
Coffee	2	28	28	286
Computers	2	250	250	720
Cricket_X	12	390	390	300
Cricket_Y	12	390	390	300
Cricket_Z	12	390	390	300
DiatomSizeReduction	4	16	306	345
DistalPhalanxOutlineAgeGroup	3	139	400	80
DistalPhalanxOutlineCorrect	2	276	600	80
DistalPhalanxTW	6	139	400	80
Earthquakes	2	139	322	512
ECG200	2	100	100	96
ECG5000	5	500	4500	140
ECGFiveDays	2	23	861	136
Face (all)	14	560	1690	131
Face (four)	4	24	88	350
FacesUCR	14	200	2050	131
Fish	7	175	175	463
FordA	2	1320	3601	500
FordB	2	810	3636	500
Gun-Point	2	50	150	150
Ham	2	109	105	431
Haptics	5	155	308	1092
Herring	2	64	64	512
InlineSkate	7	100	550	1882
InsectWingbeatSound	11	220	1980	256



425352952

CU Thesisis 6070323921 thesisis / revv: 06082562 14:42:31 / seq: 83

ชื่อชุดข้อมูล	จำนวน คลาส	จำนวน ข้อมูลสอน	จำนวนข้อมูล ทดสอบ	ความยาวอนุกรม เวลา
ItalyPowerDemand	2	67	1029	24
LargeKitchenAppliances	3	375	375	720
Lighting2	2	60	61	637
Lighting7	7	70	73	319
MALLAT	8	55	2345	1024
Meat	3	60	60	448
MedicalImages	10	381	760	99
MiddlePhalanxOutlineAgeGroup	3	154	400	80
MiddlePhalanxOutlineCorrect	2	291	600	80
MiddlePhalanxTW	6	154	399	80
MoteStrain	2	20	1252	84
OliveOil	4	30	30	570
OSU Leaf	6	200	242	427
PhalangesOutlinesCorrect	2	1800	858	80
Plane	7	105	105	144
ProximalPhalanxOutlineAgeGroup	3	400	205	80
ProximalPhalanxOutlineCorrect	2	600	291	80
ProximalPhalanxTW	6	205	400	80
RefrigerationDevices	3	375	375	720
ScreenType	3	375	375	720
ShapeletSim	2	20	180	500
ShapesAll	60	600	600	512
SmallKitchenAppliances	3	375	375	720
SonyAIBORobot Surface	2	20	601	70
SonyAIBORobot SurfaceII	2	27	953	65

ชื่อชุดข้อมูล	จำนวน คลาส	จำนวน ข้อมูลสอน	จำนวนข้อมูล ทดสอบ	ความยาวอนุกรม เวลา
Strawberry	2	370	613	235
Swedish Leaf	15	500	625	128
Symbols	6	25	995	398
Synthetic Control	6	300	300	60
ToeSegmentation1	2	40	228	277
ToeSegmentation2	2	36	130	343
Trace	4	100	100	275
Two Patterns	4	1000	4000	128
TwoLeadECG	2	23	1139	82
uWaveGestureLibrary_X	8	896	3582	315
uWaveGestureLibrary_Y	8	896	3582	315
uWaveGestureLibrary_Z	8	896	3582	315
Wafer	2	1000	6174	152
Wine	2	57	54	234
WordSynonyms	25	267	638	270
Worms	5	77	181	900
WormsTwoClass	2	77	181	900
Yoga	2	300	3000	426

จากตารางที่ 4.1 แสดงรายละเอียดของข้อมูลแต่ละชุดที่ใช้ในการทดลองเพื่อประเมินประสิทธิภาพในการจำแนกประเภทข้อมูลอนุกรมเวลา โดยชุดข้อมูลที่น่ามาใช้ในการทดลองจะมีความหลากหลายของชุดข้อมูล เหมาะที่จะนำมาใช้ในการทดลอง

4.2 การวัดและประเมินผล

สำหรับในวิทยานิพนธ์นี้ จะใช้วิธีการไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG1D-L ร่วมกับการจำแนกประเภทแบบเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว จากนั้นจะทำการวัดผล

โดยใช้อัตราความผิดพลาด (Error rate) ที่สามารถคำนวณได้จากสมการที่ 4.1 เปรียบเทียบกับการวัดระยะห่างไดนามิกโทมวอร์ปิงและไดนามิกโทมวอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D เพื่อใช้ในการวัดผลความน่าเชื่อถือในการจำแนกประเภทข้อมูล

$$\text{อัตราความผิดพลาด (Error rate)} = 1 - \frac{\text{จำนวนข้อมูลที่จำแนกได้ถูกต้องตามประเภทของข้อมูลเหล่านั้น}}{\text{จำนวนข้อมูลทั้งหมด}} \quad (4.1)$$

ผลลัพธ์ของอัตราความผิดพลาด ขึ้นอยู่กับอัลกอริทึมที่ใช้ทดสอบนั้นให้ผลลัพธ์ในการทำนายประเภทของข้อมูลได้ถูกต้องมากน้อยเพียงใด โดยสำหรับวิธีการที่จะนำมาเพื่อเปรียบเทียบนั้น จะใช้วิธีคำนวณหาระยะทางระหว่างอนุกรมเวลา 2 อนุกรมด้วยวิธีการไดนามิกโทมวอร์ปิง โดยการนำข้อมูลทดสอบไปหาระยะทางกับชุดข้อมูลสอน และนำคลาสของข้อมูลสอนที่มีระยะทางที่สั้นที่สุดมาเปรียบเทียบกับคลาสของข้อมูลทดสอบนั้น ถ้าคลาสดตรงกันคือจำแนกประเภทได้ถูกต้อง เมื่อนำข้อมูลทดสอบทั้งหมดไปหาระยะทางกับชุดข้อมูลสอน และจำแนกประเภทของข้อมูลเสร็จ จะสามารถคำนวณหาอัตราความผิดพลาดได้จากสมการที่ 4.1

สำหรับวิธีการไดนามิกโทมวอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D และไดนามิกโทมวอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG1D-L จะทำการสร้างกลุ่มตัวอย่างของลำดับย่อยจากข้อมูลอนุกรมเวลา จากนั้นแปลงลำดับย่อยของอนุกรมเวลาให้เป็นลำดับของตัวอธิบายรูปร่าง ก่อนที่จะนำไปหาระยะทางด้วยวิธีการไดนามิกโทมวอร์ปิง โดยทั้ง 3 วิธีการจะใช้ชุดข้อมูลอนุกรม เวลาจากทั้งหมด 78 ชุดจาก UCR Time Series Classification Archive [6]

4.3 การทดลองเพื่อหาพารามิเตอร์ $nLevel$ ที่ดีที่สุด

การทดลองเพื่อหาพารามิเตอร์ที่เหมาะสมที่จะนำมาใช้ในการทดลอง ซึ่งวิทยานิพนธ์นี้ได้ทำการทดลองเพื่อหาพารามิเตอร์ $nLevel$ ที่เหมาะสมกับการทดลองด้วยวิธีการ Leave-One-Out Cross Validation โดยการเลือกข้อมูล 1 ตัวจากชุดข้อมูลสอนออกมา แล้วนำข้อมูลที่เหลือในชุดข้อมูลไปทำการสอน (Training) กับวิธีการที่นำเสนอ จากนั้นนำมาทดสอบกับข้อมูล 1 ตัวที่ถูกเลือกออกมาในข้างต้น เมื่อทดสอบเสร็จจึงนำข้อมูลที่ใช้ทดสอบกลับเข้าไปยังชุดข้อมูลสอน แล้วเริ่มกระบวนการเลือกข้อมูลจากชุดข้อมูลใหม่ โดยการเลือกข้อมูลมาทดสอบต้องไม่ซ้ำกับข้อมูลที่เคยเลือกมา ความแม่นยำที่ได้จากการทดสอบพารามิเตอร์ที่กำหนดกับชุดข้อมูลต่าง ๆ แสดงในตาราง 4.2

ตารางที่ 4.2 แสดงการเปรียบเทียบความแม่นยำจากการจำแนกประเภทข้อมูลโดยวิธีการเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว ด้วยวิธีการ Leave-One-Out Cross Validation เพื่อหาเพื่อหาพารามิเตอร์ $nLevel$ ที่ดีที่สุด

ชื่อชุดข้อมูล	ความแม่นยำ (%)		
	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L ($nLevel = 1$)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L ($nLevel = 2$)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L ($nLevel = 3$)
50Words	0.767	0.767	0.767
Adiac	0.713	0.713	0.715
ArrowHead	0.889	0.889	0.889
Beef	0.433	0.400	0.433
BeetleFly	0.800	0.800	0.800
BirdChicken	0.900	0.900	0.900
Car	0.800	0.800	0.800
CBF	0.967	0.900	0.967
ChlorineConcentration	0.612	0.621	0.623
CinC_ECG_torso	0.825	0.825	0.825
Coffee	0.964	0.964	0.964
Computers	0.764	0.776	0.784
Cricket_X	0.820	0.828	0.828
Cricket_Y	0.800	0.805	0.792
Cricket_Z	0.780	0.792	0.780
DiatomSizeReduction	0.938	0.938	0.938
DistalPhalanxOutlineAgeGroup	0.662	0.698	0.691

ชื่อชุดข้อมูล	ความแม่นยำ (%)		
	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L (<i>nLevel</i> = 1)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L (<i>nLevel</i> = 2)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L (<i>nLevel</i> = 3)
DistalPhalanxOutlineCorrect	0.743	0.728	0.728
DistalPhalanxTW	0.655	0.648	0.655
Earthquakes	0.676	0.655	0.655
ECG200	0.820	0.810	0.820
ECG5000	0.926	0.934	0.926
ECGFiveDays	0.913	0.913	0.913
Face (all)	0.966	0.959	0.957
Face (four)	0.875	0.875	0.875
FacesUCR	0.930	0.920	0.925
Fish	0.903	0.903	0.903
FordB	0.648	0.658	0.658
Gun-Point	0.980	0.980	0.980
Ham	0.853	0.835	0.826
ItalyPowerDemand	0.940	0.925	0.925
LargeKitchenAppliances	0.891	0.877	0.872
Lighting2	0.883	0.867	0.850
Lighting7	0.771	0.743	0.771
MALLAT	0.927	0.927	0.927
Meat	0.850	0.850	0.867



425352952

CU Thesisis 6070323221 thesisis / revv: 06082562 14:42:31 / seq: 83

ชื่อชุดข้อมูล	ความแม่นยำ (%)		
	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L (<i>nLevel</i> = 1)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L (<i>nLevel</i> = 2)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L (<i>nLevel</i> = 3)
MedicalImages	0.735	0.735	0.753
MiddlePhalanxOutlineAgeGroup	0.539	0.552	0.565
MiddlePhalanxOutlineCorrect	0.746	0.729	0.742
MiddlePhalanxTW	0.578	0.546	0.565
MoteStrain	0.800	0.800	0.800
OliveOil	0.733	0.800	0.800
OSU Leaf	0.870	0.870	0.870
PhalangesOutlinesCorrect	0.788	0.788	0.793
Plane	1.000	1.000	1.000
ProximalPhalanxOutlineAgeGroup	0.763	0.763	0.763
ProximalPhalanxOutlineCorrect	0.833	0.842	0.840
RefrigerationDevices	0.845	0.848	0.821
ProximalPhalanxTW	0.688	0.722	0.673
ScreenType	0.608	0.603	0.595
ShapeletSim	0.900	0.850	0.900
SmallKitchenAppliances	0.696	0.733	0.725
SonyAIBORobot Surface	0.900	0.950	0.950
SonyAIBORobot Surfacell	0.852	0.889	0.852
Strawberry	0.930	0.930	0.932
Swedish Leaf	0.930	0.926	0.924
Symbols	1.000	1.000	1.000

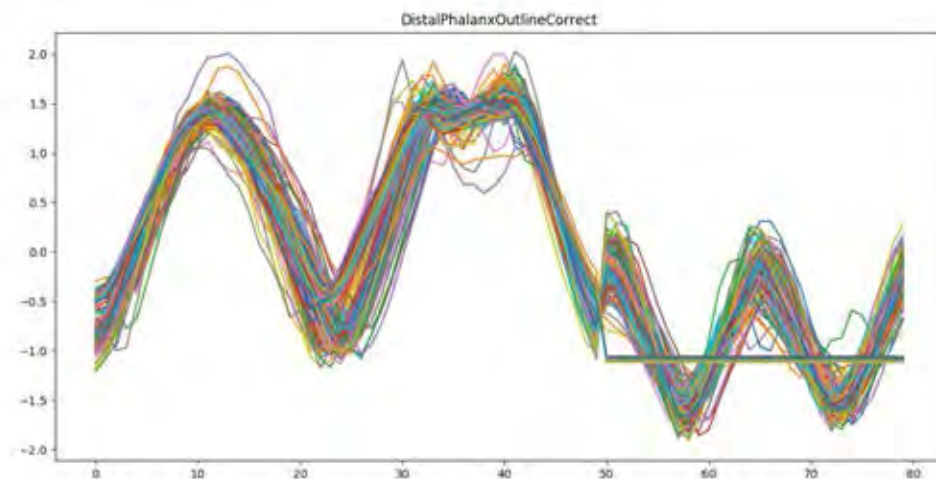
ชื่อชุดข้อมูล	ความแม่นยำ (%)		
	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L ($nLevel = 1$)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L ($nLevel = 2$)	ไดนามิกโทม วอร์ปิงแบบรูปร่าง ภายใต้ตัวอธิบาย รูปร่าง HOG1D-L ($nLevel = 3$)
Synthetic Control	0.887	0.873	0.853
ToeSegmentation1	0.700	0.700	0.725
ToeSegmentation2	0.917	0.917	0.917
Trace	1.000	1.000	1.000
Two Patterns	1.000	1.000	0.999
TwoLeadECG	0.957	0.957	0.957
uWaveGestureLibrary_X	0.732	0.733	0.733
uWaveGestureLibrary_Y	0.674	0.673	0.671
uWaveGestureLibrary_Z	0.722	0.721	0.721
Wafer	0.996	0.993	0.992
Wine	0.895	0.895	0.877
WordSynonyms	0.727	0.730	0.727
Worms	0.403	0.403	0.429
WormsTwoClass	0.623	0.597	0.597
Yoga	0.880	0.880	0.880

จากตาราง 4.2 ค่าพารามิเตอร์ $nLevel$ นั้น ใช้สำหรับการแบ่งจำนวนส่วนจากช่วงของลำดับย่อย เพื่อใช้ในการคำนวณความถดถอยเชิงเส้นสำหรับการหาแนวโน้มของทิศทางของจุดภายใต้ส่วนนั้น ๆ ในเปรียบเทียบความสอดคล้องของทิศทางของเกรเดียนท์ของจุดและทิศทางของเส้นความถดถอยเชิงเส้นภายใต้ส่วน

การวิเคราะห์เพื่อปรับพารามิเตอร์ $nLevel$ เพื่อหาความแม่นยำจากการจำแนกประเภทข้อมูลที่ได้จากทั้ง 3 ค่าของ $nLevel$ นั้น โดยจากค่าความแม่นยำที่ได้สามารถจำแนกเป็นชุดข้อมูลที่ได้ค่าความแม่นยำที่ดีในค่า $nLevel$ ต่าง ๆ ซึ่งแบ่งออกได้ 3 กรณีคือ กรณีที่ชุดข้อมูลได้ค่าความ

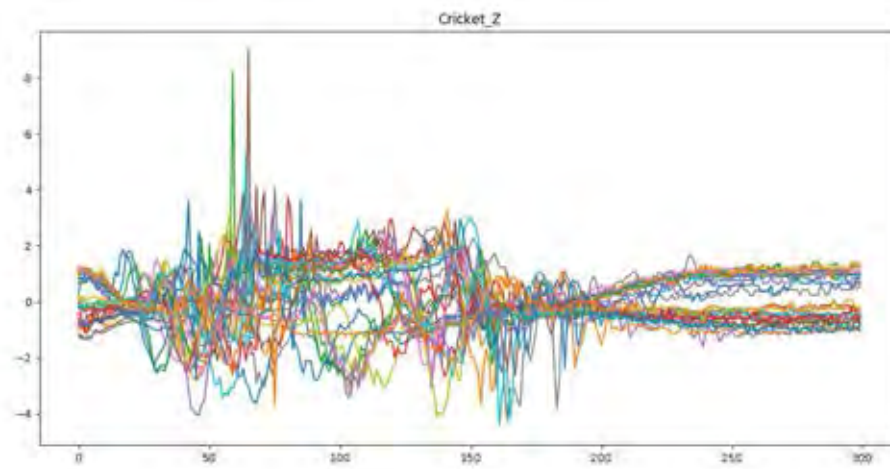
แม่นยำที่ดีเมื่อ $nLevel$ เท่ากับ 1 กรณีที่ชุดข้อมูลได้ค่าความแม่นยำที่ดีเมื่อ $nLevel$ เท่ากับ 2 และกรณีที่ชุดข้อมูลได้ค่าความแม่นยำที่ดีเมื่อ $nLevel$ เท่ากับ 3

กรณีที่ชุดข้อมูลได้ค่าความแม่นยำที่ดีเมื่อ $nLevel$ เท่ากับ 1 ในกรณีนั้น จะเป็นการแบ่งช่วงของลำดับย่อยออกเป็น 2^{nLevel} ส่วน หรือแบ่งออกเป็น 2 ส่วน ทำให้การคำนวณความถดถอยเชิงเส้นในกรณีนี้ จะมีจุดในการคำนวณมากกว่า $nLevel$ ค่าอื่น เช่นจากตัวอย่างการทำงานของตัวอธิบายรูปร่างที่ใช้ลำดับย่อยความยาวเท่ากับ 30 ทำให้ในแต่ละส่วนนั้นมีจุด 7 ถึง 8 จุดในการคำนวณความถดถอยเชิงเส้น ทำให้กรณีนี้เหมาะแก่การจำแนกประเภทข้อมูลอนุกรมเวลา ที่ข้อมูลมีความซับซ้อนน้อย เช่น ชุดข้อมูล DisphalanxOutlineCorrect ดังรูปที่ 4.1



รูปที่ 4.1 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล DisphalanxOutlineCorrect

กรณีที่ชุดข้อมูลได้ค่าความแม่นยำที่ดีเมื่อ $nLevel$ เท่ากับ 2 ในกรณีนี้จะคล้ายกับในกรณีแรก โดยส่วนที่นำมาคำนวณหาความถดถอยเชิงเส้นต่อช่วงจะมากขึ้น และจุดในส่วนที่ใช้คำนวณจะลดลง ทำให้กรณีนี้สามารถใช้งานได้ดีในกรณีที่ชุดข้อมูลที่มีความถี่ของความซับซ้อนมากขึ้น เช่นชุดข้อมูล Cricket_Z ดังรูปที่ 4.2



รูปที่ 4.2 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Cricket_Z

กรณีที่ชุดข้อมูลได้ค่าความแม่นยำดีเมื่อ $nLevel$ เท่ากับ 3 ในกรณีนี้ จุดที่ใช้ในการคำนวณความถดถอยเชิงเส้นของแต่ละส่วนจะน้อยที่สุด ทำให้ข้อมูลที่จะได้ผลลัพธ์ที่ดีจะต้องมีความถี่ของความขรุขระที่สูง เพื่อที่จะไม่สูญเสียโครงสร้างของรูปร่างของลำดับย่อยนั้น เช่น ชุดข้อมูล ToeSegmentation1 ดังรูปที่ 4.3



รูปที่ 4.3 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล ToeSegmentation1

4.4 ผลการทดลอง

จากวิทยานิพนธ์นี้ ผู้วิจัยได้ใช้หลักการในการพัฒนาตัวอธิบายรูปร่างแบบ HOG1D-L คือการนำข้อมูลอนุกรมเวลามาปรับให้เรียบ (Smoothing) โดยการกำจัดทิศทางของจุดที่ไม่สอดคล้องกับแนวเส้นความถดถอยเชิงเส้นภายใต้ส่วน ดังที่ได้อธิบายไปในบทที่ 1 และจากการทดลอง

เพื่อหาอัตราความผิดพลาดในการจำแนกประเภทข้อมูลในชุดข้อมูลอนุกรมเวลาของ UCR Time Series Classification Archive [6] จากการทดลองในวิทยานิพนธ์นี้จะเปรียบเทียบกับการวัดระยะทางแบบไดนามิกไทม์วอร์ปิง และเทียบกับไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D สำหรับผลการทดลองในชุดข้อมูลจำนวน 78 ชุด จาก UCR Time Series Classification Archive [6] แสดงในตาราง 4.3

ตารางที่ 4.3 แสดงการเปรียบเทียบอัตราความผิดพลาดในการจำแนกประเภทข้อมูลโดยวิธีการเพื่อนบ้านที่ใกล้ที่สุด 1 ตัว ควบคู่กับการวัดระยะห่างแบบไดนามิกไทม์วอร์ปิง ไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D และไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG1D-L

ชื่อชุดข้อมูล	อัตราความผิดพลาด (%)		
	ไดนามิกไทม์วอร์ปิง	ไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D	ไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG1D-L
50Words	0.310	0.242	0.237
Adiac	0.396	0.269	0.274
ArrowHead	0.297	0.177	0.177
Beef	0.367	0.267	0.267
BeetleFly	0.300	0.200	0.200
BirdChicken	0.250	0.050	0.050
Car	0.267	0.133	0.133
CBF	0.003	0.080	0.123
ChlorineConcentration	0.352	0.355	0.370
CinC_ECG_torso	0.349	0.209	0.207
Coffee	0.000	0.036	0.036
Computers	0.300	0.356	0.336
Cricket_X	0.246	0.208	0.197

ชื่อชุดข้อมูล	อัตราความผิดพลาด (%)		
	ไดนามิกใหม่ วอร์ปิง	ไดนามิกใหม่วอร์ปิง แบบรูปร่างภายใต้ตัว อธิบายรูปร่าง HOG-1D	ไดนามิกใหม่วอร์ปิง แบบรูปร่างภายใต้ตัว อธิบายรูปร่าง HOG1D-L
Cricket_Y	0.256	0.226	0.280
Cricket_Z	0.246	0.208	0.200
DiatomSizeReduction	0.033	0.069	0.069
DistalPhalanxOutlineAgeGroup	0.208	0.233	0.218
DistalPhalanxOutlineCorrect	0.232	0.228	0.228
DistalPhalanxTW	0.290	0.290	0.280
Earthquakes	0.258	0.258	0.257
ECG200	0.230	0.100	0.110
ECG5000	0.076	0.071	0.072
ECGFiveDays	0.232	0.057	0.057
Face (all)	0.192	0.238	0.229
Face (four)	0.170	0.091	0.091
FacesUCR	0.095	0.081	0.062
Fish	0.177	0.051	0.051
FordA	0.438	0.279	0.272
FordB	0.406	0.261	0.267
Gun-Point	0.093	0.007	0.007
Ham	0.533	0.457	0.467
Haptics	0.623	0.562	0.542
Herring	0.469	0.500	0.500
InlineSkate	0.616	0.629	0.626
InsectWingbeatSound	0.645	0.584	0.585
ItalyPowerDemand	0.050	0.103	0.069
LargeKitchenAppliances	0.205	0.160	0.152

ชื่อชุดข้อมูล	อัตราความผิดพลาด (%)		
	ไดนามิกใหม่ วอร์ปิง	ไดนามิกใหม่วอร์ปิง แบบรูปร่างภายใต้ตัว อธิบายรูปร่าง HOG-1D	ไดนามิกใหม่วอร์ปิง แบบรูปร่างภายใต้ตัว อธิบายรูปร่าง HOG1D-L
Lighting2	0.131	0.115	0.164
Lighting7	0.274	0.233	0.260
MALLAT	0.066	0.062	0.065
Meat	0.067	0.100	0.100
MedicalImages	0.263	0.264	0.259
MiddlePhalanxOutlineAgeGroup	0.250	0.260	0.258
MiddlePhalanxOutlineCorrect	0.352	0.250	0.242
MiddlePhalanxTW	0.416	0.429	0.421
MoteStrain	0.165	0.110	0.102
OliveOil	0.167	0.100	0.100
OSU Leaf	0.409	0.132	0.149
PhalangesOutlinesCorrect	0.272	0.261	0.256
Plane	0.000	0.000	0.000
ProximalPhalanxOutlineAgeGroup	0.195	0.210	0.234
ProximalPhalanxOutlineCorrect	0.216	0.206	0.210
ProximalPhalanxTW	0.263	0.275	0.250
RefrigerationDevices	0.536	0.507	0.520
ScreenType	0.603	0.525	0.533
ShapeletSim	0.350	0.028	0.033
ShapesAll	0.232	0.112	0.112
SmallKitchenAppliances	0.357	0.301	0.288
SonyAIBORobot Surface	0.275	0.193	0.206
SonyAIBORobot Surfacell	0.169	0.174	0.164
Strawberry	0.060	0.051	0.054

ชื่อชุดข้อมูล	อัตราความผิดพลาด (%)		
	ไดนามิกไทม์ วอร์ปิง	ไดนามิกไทม์วอร์ปิง แบบรูปร่างภายใต้ตัว อธิบายรูปร่าง HOG-1D	ไดนามิกไทม์วอร์ปิง แบบรูปร่างภายใต้ตัว อธิบายรูปร่าง HOG1D-L
Swedish Leaf	0.208	0.085	0.088
Symbols	0.050	0.039	0.038
Synthetic Control	0.007	0.153	0.107
ToeSegmentation1	0.228	0.101	0.097
ToeSegmentation2	0.162	0.138	0.138
Trace	0.000	0.000	0.000
Two Patterns	0.000	0.001	0.002
TwoLeadECG	0.096	0.006	0.006
uWaveGestureLibrary_X	0.273	0.263	0.263
uWaveGestureLibrary_Y	0.366	0.358	0.357
uWaveGestureLibrary_Z	0.342	0.338	0.340
Wafer	0.020	0.010	0.010
Wine	0.426	0.537	0.519
WordSynonyms	0.351	0.260	0.259
Worms	0.536	0.475	0.470
WormsTwoClass	0.337	0.287	0.287
Yoga	0.164	0.117	0.116
Mean	0.2546	0.2101	0.2099

จากตารางที่ 4.3 ตัวเลขที่เป็นตัวหนาจะแสดงอัตราความผิดพลาดที่น้อยที่สุดเมื่อเทียบกับไดนามิกไทม์วอร์ปิง ไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG-1D และไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้ตัวอธิบายรูปร่าง HOG1D-L จากตารางจะเห็นได้ว่า HOG1D-L ให้ค่าอัตราความผิดพลาดในการจำแนกประเภทข้อมูลอนุกรมเวลาในภาพรวมที่น้อยเมื่อเทียบกับไดนามิกไทม์วอร์ปิงสรุปได้ ดังตารางที่ 4.4

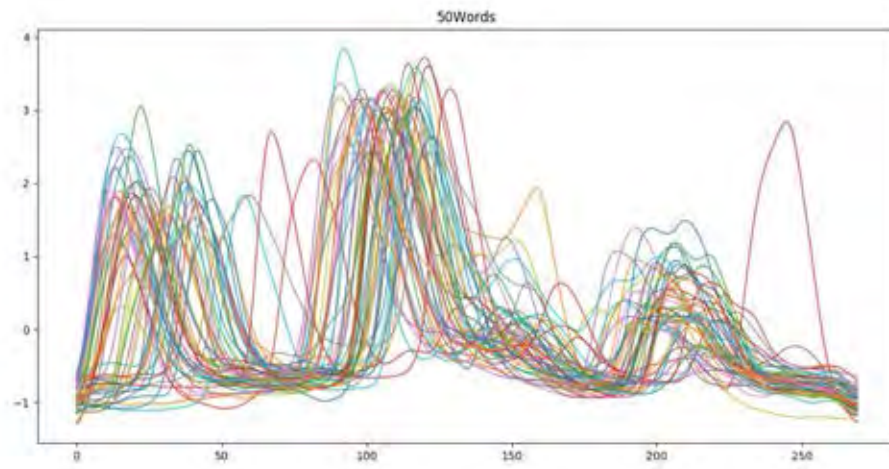
ตารางที่ 4.4 สรุปผลการทดลอง

ผลการทดลอง	จำนวนผลการทดลองของตัวอธิบายรูปร่าง HOG1D-L เมื่อเปรียบเทียบกับ	
	ไดนามิกไทม์วอร์ปิง	ไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้การทำงานของ HOG-1D
ชนะ	57	32
เสมอ	2	23
แพ้	19	23

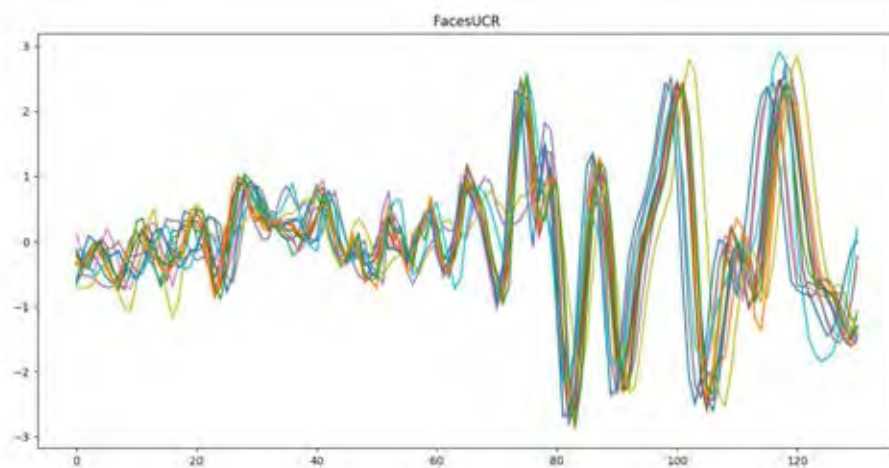
จากตารางที่ 4.4 จะพบว่าจากการทดลองกับข้อมูล UCR Time Series Classification Archive [6] จำนวน 78 ชุดข้อมูล ค่าอัตราความผิดพลาดของไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้การทำงานของตัวอธิบายรูปร่าง HOG1D-L มีค่าน้อยกว่าเมื่อเทียบกับอัตราความผิดพลาดของไดนามิกไทม์วอร์ปิง และไดนามิกไทม์วอร์ปิงแบบรูปร่างภายใต้การทำงานของ HOG-1D

จากอัตราความผิดพลาดที่กล่าวไป ทำให้สามารถแบ่งลักษณะของข้อมูลอนุกรมเวลาที่ใช้ในการทดลองได้ 2 กรณี คือกรณีลักษณะของข้อมูลอนุกรมเวลาที่แตกต่างกันในบางช่วง และลักษณะของข้อมูลอนุกรมเวลาที่มีความคล้ายกันในคลาส

กรณีที่ 1 ลักษณะของข้อมูลอนุกรมเวลาที่แตกต่างกันในบางช่วง ในกรณีนี้คือการที่ข้อมูลอนุกรมเวลาในคลาสเดียวกัน มีลักษณะของข้อมูลที่แตกต่างกันในบางช่วง ดังรูปที่ 4.4 ข้อมูลอนุกรมเวลาจากชุดข้อมูล 50words และรูปที่ 4.5 ข้อมูลอนุกรมเวลาจากชุดข้อมูล FaceUCR จะเห็นว่ากราฟของข้อมูลอนุกรมเวลาของทั้งสองรูปนั้นจะมีความแตกต่างกันในบางช่วง และจากผลการทดลองในตารางที่ 4.3 จะเห็นได้ว่าผลลัพธ์ของตัวอธิบายรูปร่างแบบ HOG1D-L จะได้ผลลัพธ์ที่ดีกว่าในกรณีที่ชุดข้อมูลมีลักษณะของข้อมูลแตกต่างกันในบางช่วง



รูปที่ 4.4 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล 50words

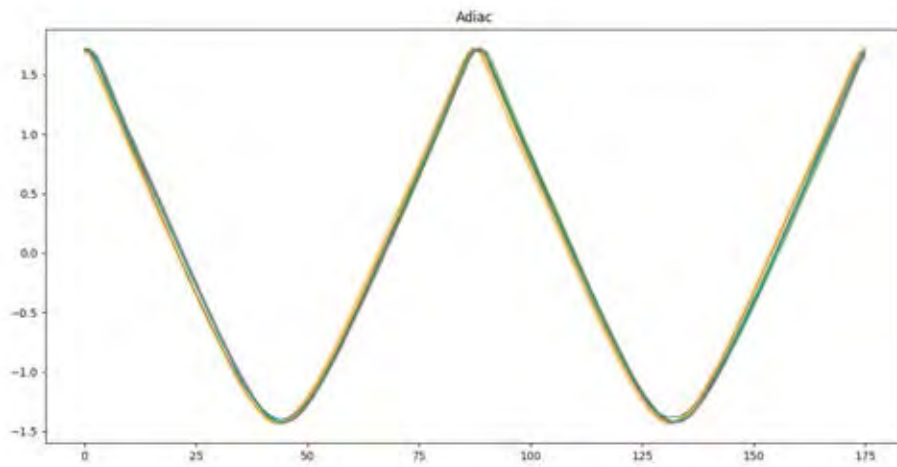


รูปที่ 4.5 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล FaceUCR

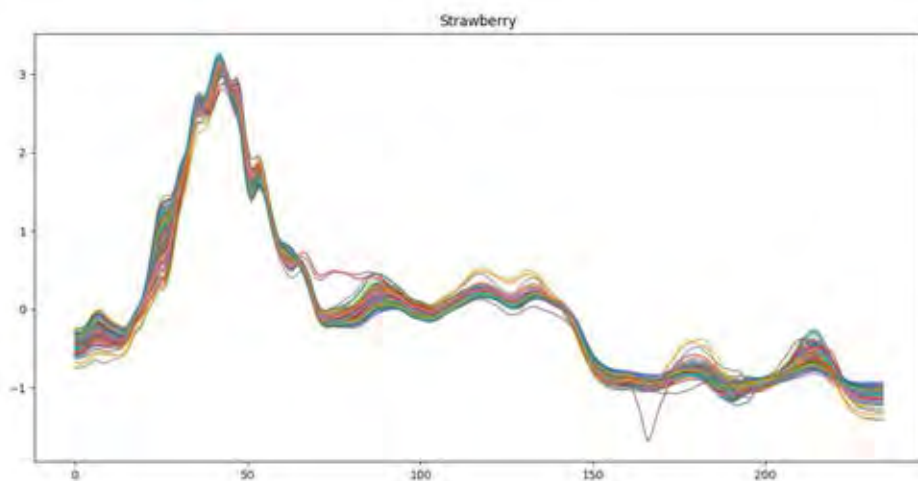
กรณีที่ 2 ลักษณะของข้อมูลอนุกรมเวลาที่มีความคล้ายกัน ในกรณีนี้คือการที่ข้อมูลอนุกรมเวลาในคลาสเดียวกัน มีลักษณะของข้อมูลที่คล้ายคลึงกัน ดังรูปที่ 4.6 ข้อมูลอนุกรมเวลาจากชุดข้อมูล Adiac และรูปที่ 4.7 ข้อมูลอนุกรมเวลาจากชุดข้อมูล Strawberry จะเห็นว่ากราฟของข้อมูลอนุกรมเวลาของทั้งสองรูปนั้นจะมีกราฟที่คล้ายกัน และจากผลการทดลองในตารางที่ 4.3 จะเห็นว่าผลลัพธ์ของตัวอธิบายรูปร่างแบบ HOG1D-L จะได้ผลลัพธ์ที่แย่กว่าตัวอธิบายรูปร่างแบบ HOG-1D ในกรณีที่ชุดข้อมูลมีลักษณะของข้อมูลที่คล้ายคลึงกัน เนื่องมาจากการตัดทิศทางที่ไม่สอดคล้องออกไป อาจทำให้เกิดการสูญเสียโครงสร้างของรูปร่างจากลำดับย่อย



425352952



รูปที่ 4.6 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Adiac



รูปที่ 4.7 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Strawberry

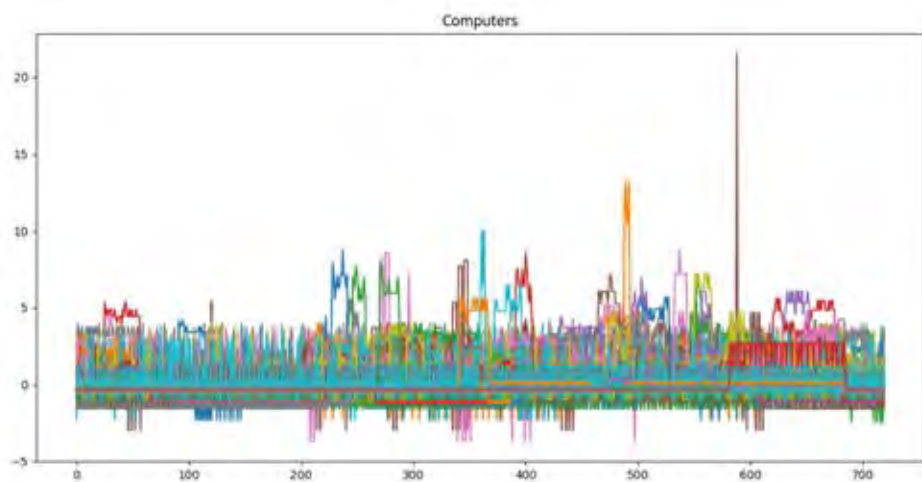
จากทั้ง 2 กรณีนั้น สรุปได้ว่า ลักษณะของข้อมูลที่มีความแตกต่างกันในบางช่วง จะได้ผลลัพธ์ความแม่นยำในการจำแนกประเภทข้อมูลที่ตีเมื่อใช้ตัวอธิบายรูปร่างแบบ HOG1D-L ในทางกลับกัน เมื่อลักษณะของข้อมูลมีความแตกต่างกันน้อย ผลลัพธ์ความแม่นยำในการจำแนกประเภทข้อมูลจากตัวอธิบายรูปร่างแบบ HOG1D-L ก็จะใกล้เคียงกับตัวอธิบายรูปร่างแบบ HOG-1D และเมื่อลักษณะของข้อมูลมีความคล้ายกัน จะได้ผลลัพธ์ที่ดีกว่าเมื่อใช้ตัวอธิบายรูปร่างแบบ HOG-1D

ในส่วนของชุดข้อมูลที่วิธีการไดนามิกไดนามิกไทม์วอร์ปิงภายใต้ตัวอธิบายรูปร่างแบบ HOG-1D ที่มีอัตราความผิดพลาดที่สูงกว่าวิธีการไดนามิกไทม์วอร์ปิง เนื่องมาจากชุดข้อมูลดังกล่าว



425352952

เช่น Computers มีความขรุขระและมีความถี่สูง ดังรูปที่ 4.8 จะทำให้ยากต่อการจำแนกโครงสร้างเฉพาะส่วนเมื่อแปลงข้อมูลอนุกรมเวลาเป็นลำดับของตัวอธิบายรูปร่างแบบ HOG-1D สำหรับกรณีชุดข้อมูล ItalyPowerDemand ความยาวของข้อมูลอนุกรมเวลามีความยาวอยู่ที่ 24 ในขณะที่การแบ่งลำดับย่อยของตัวอธิบายรูปร่างในการทดลองคือ 30 ทำให้ลำดับย่อยของแต่ละจุดในข้อมูลอนุกรมเวลาจะความคล้ายคลึงกัน ทำให้ไม่สามารถจำแนกโครงสร้างเฉพาะส่วนได้



รูปที่ 4.8 ข้อมูลอนุกรมเวลาคลาสที่ 1 จากชุดข้อมูล Computers

บทที่ 5

สรุปผลงานวิจัยและข้อเสนอแนะ

วิทยานิพนธ์นี้ นำเสนอตัวอธิบายรูปร่าง HOG1D-L ซึ่งพัฒนามาจากตัวอธิบายรูปร่าง HOG1-D โดยตัวอธิบายรูปร่างที่นำเสนอนี้ใช้วิธีการปรับให้เรียบ (Smoothing) ของลำดับย่อยในขั้นตอนของตัวอธิบายรูปร่าง HOG-1D โดยการแบ่งช่วงของลำดับย่อยเป็นจำนวนส่วนตามที่กำหนด แล้วใช้วิธีการวิเคราะห์ความถดถอยเชิงเส้นในการหาทิศทางของแนวเส้นความถดถอยเชิงเส้นในแต่ละส่วน ทิศทางของจุดภายใต้ส่วนที่ไม่สอดคล้องกับแนวเส้นความถดถอยเชิงเส้นในส่วนนั้น จะไม่ถูกนำไปโหวต (voting) ลงถึงฮิสโตแกรม ทำให้ลำดับย่อยของตัวอธิบายรูปร่างนั้นมีความเหมาะสมในการนำไปจำแนกประเภทของข้อมูลมากยิ่งขึ้น โดยวิธีการที่นำเสนอ นั้น สามารถให้ความแม่นยำในการจำแนกประเภทข้อมูลอนุกรมเวลาได้ดีกว่าวิธีการจำแนกประเภทข้อมูลแบบเพื่อนบ้านใกล้ที่สุด 1 ตัว ร่วมกับการวัดระยะทางโคนามิกโทมวอร์ปเป็นอย่างมากและให้ความแม่นยำที่ดีกว่าตัวอธิบายรูปร่างแบบ HOG-1D ภายใต้การทำงานของโคนามิกโทมวอร์ปแบบรูปร่าง ดังที่ได้นำเสนอผลการทดลองและการวิเคราะห์ผลการทดลองในบทที่ผ่านมา โดยสามารถสรุปผลการวิจัย อภิปรายผลและข้อเสนอแนะได้ดังนี้

5.1 สรุปผลงานวิจัย

โคนามิกโทมวอร์ปเป็นวิธีการวัดระยะทางที่นิยมใช้ในการจำแนกประเภทข้อมูลอนุกรมเวลา โดยขั้นตอนของโคนามิกโทมวอร์ปจะหาการปรับแนวโดยรวมที่เหมาะสม ซึ่งโคนามิกโทมวอร์ปจะใช้วิธีการจับคู่ระหว่างจุดข้อมูลที่มีระยะห่างสั้นที่สุด ภายใต้เงื่อนไขและขอบเขตบางอย่าง จากรูปที่ 1.1 (ก) ระยะห่างระหว่างจุดที่สั้นที่สุดที่ไม่ได้อยู่ในตำแหน่งที่เหมาะสม ทำให้การปรับแนวนั้นมีความไม่เหมาะสมเมื่อเทียบกับรูปที่ 1.1 (ข) จึงมีงานวิจัยที่นำเสนอ ชื่อว่าโคนามิกโทมวอร์ปแบบรูปร่างขึ้นมาเพื่อพิจารณาโครงสร้างของข้อมูลโดยรวมข้อมูลรูปร่างเฉพาะส่วนบริเวณรอบ ๆ แต่ละจุดเข้าสู่การปรับแนวด้วยโคนามิกโทมวอร์ป โดยส่วนที่มีโครงสร้างข้อมูลคล้ายกันจะทำการจับคู่กัน โดยขั้นตอนหลัก ๆ ของโคนามิกโทมวอร์ปแบบรูปร่างประกอบด้วย 3 ขั้นตอนคือ (1) การสร้างกลุ่มตัวอย่างของลำดับย่อยจากข้อมูลอนุกรมเวลา (2) แปลงลำดับย่อยด้วยตัวอธิบายรูปร่างจากกรอบการทำงานของโคนามิกโทมวอร์ปแบบรูปร่างที่ให้ผู้ใช้งานสามารถออกแบบหรือนำตัวอธิบายรูปร่างมาใช้ (3) เมื่อได้ลำดับของตัวอธิบายรูปร่างที่ผ่านการคำนวณด้วยตัวอธิบายรูปร่างแล้วจะทำการปรับแนวด้วยโคนามิกโทมวอร์ปแบบรูปร่าง

ตัวอธิบายรูปร่างมีหลายวิธีซึ่งวิธีที่น่าสนใจและให้ความแม่นยำ คือ ฮิสโตแกรมตามทิศทางของเกรเดียนท์แบบ 1 มิติ ที่มีการนำแนวคิดของฮิสโตแกรมตามค่าเกรเดียนท์ที่ใช้ในการตรวจจับผู้คน หรือตรวจจับวัตถุ มาใช้ในการตรวจจับโครงสร้างของอนุกรมเวลาโดยวิธีการฮิสโตแกรมตามค่าเกรเดียนท์แบบ 1 มิติ โดยจะคำนวณหาขนาดและทิศทางตามค่าเกรเดียนท์ของแต่ละจุดข้อมูล มาเก็บลงยังถึงฮิสโตแกรม ซึ่งค่าแต่ละจุดจะนำมาพิจารณาร่วมกับโครงสร้างของอนุกรมเวลาเพื่อใช้ในการเปรียบเทียบความเหมือนกันของโครงสร้างระหว่างอนุกรมเวลาและนำไปปรับแนวต่อไป แต่วิธีการนี้มีข้อเสียคือหากเก็บข้อมูลที่ไม่สอดคล้องลงในถึงฮิสโตแกรมจะส่งผลให้การจำแนกประเภทข้อมูลผิดพลาด

ดังนั้นในวิทยานิพนธ์นี้จึงได้นำวิธีการวิเคราะห์การถดถอยเข้ามาเพื่อหาแนวโน้มของทิศทางของอนุกรมเวลา โดยแบ่งลำดับย่อยของอนุกรมเวลาออกเป็นส่วนย่อย ๆ แล้วสร้างเส้นของการวิเคราะห์ความถดถอยเชิงเส้นให้เป็นตัวเปรียบเทียบทิศทางเฉลี่ยของชุดข้อมูลในช่วงส่วนนั้น ๆ โดยพิจารณาข้อมูลในชุดอนุกรมเวลาในส่วนนั้น ๆ เทียบกับเส้นการถดถอยของข้อมูล หากจุดของข้อมูลมีทิศทางแตกต่างจากเส้นทิศทางของเส้นความถดถอยเชิงเส้นมากกว่า 45 องศา ข้อมูลนั้นจะถูกตัดออกไม่นำมาเก็บในถึงฮิสโตแกรม ซึ่งจากบทที่ 4 ตารางที่ 4.3 แสดงให้เห็นอัตราความผิดพลาดจากการทดลอง จะเห็นได้ว่าตัวอธิบายรูปร่าง HOG1D-L สามารถลดอัตราความผิดพลาดข้อมูลลดลงกว่าวิธีการ HOG-1D และยังให้ผลที่ดีกว่าเมื่อเทียบกับวิธีไดนามิกโทมวอร์ปิงที่เป็นที่นิยมใช้ในการปรับแนว

5.2 ข้อเสนอแนะ

ตัวอธิบายรูปร่าง HOG1D-L นั้นให้ผลลัพธ์ที่ดีขึ้นเมื่อเทียบกับตัวอธิบายรูปร่าง HOG-1D โดยการกำจัดทิศทางของจุดที่ไม่สอดคล้องเมื่อเทียบกับทิศทางของแนวเส้นความถดถอยเชิงเส้นในส่วนนั้นซึ่งต้องแลกกับการสูญเสียข้อมูล วิธีการแก้ปัญหาที่ดีกว่าการตัดข้อมูลออกควรเป็นการหาค่าตัวแทนเพื่อแทนข้อมูลที่ถูกต้องออกไป เช่น ตัวอธิบายรูปร่าง HOG1D-L ที่กำจัดทิศทางที่ไม่สอดคล้องออกไป ควรเปลี่ยนจากการกำจัดทิศทางนั้น เป็นการแทนทิศทางที่ถูกกำจัดออกไปนั้นด้วยทิศทางเฉลี่ยหรือทิศทางของแนวเส้นความถดถอยเชิงเส้น ซึ่งจะทำให้ลำดับย่อยของตัวอธิบายรูปร่างนั้นมีประสิทธิภาพและคุณภาพที่ดียิ่งขึ้น

บรรณานุกรม

- Peng Wang, Fredrick Barrett, Elizabeth Martin, Marina Milanova, Raquel E. Gur, Ruben C. Gur, Christian Kohler and Ragini Verma (2009). Automated Video Based Facial Expression Analysis of Neuropsychiatric Disorders. *J Neurosci Methods*. 168(1): 224–238.
12. Sangeetha Damu and P. Deepa (2017). A low-cost and high-performance architecture for robust human detection using histogram of edge oriented gradients. *Microprocessors and Microsystem*. 53(C): 106–119.
 13. Jessica Lin, Eamonn Keogh, Li Wei and Stefano Lonardi (2007). Experiencing SAX: a novel symbolic representation of time series. *Data Mining and Knowledge Discovery*. 107–144.
 14. Jin Wang, ping Liu, Mary She and Saeid Nahavandi (2013). Bag-of-words representation for biomedical time series classification. *Biomedical Signal Processing and Control*. 8(6): 634–644.
 15. Xiaopeng Xi, Eamonn Keogh, Christian Shelton, Li Wei, and Chotirat Ann Ratanamahatana (2006). Fast time series classification using numerosity reduction. *Proceedings of the 23rd international conference on Machine learning*. 1033–1040.
 16. Zbigniew R. Struzik and Arno Siebes (1999). The Haar Wavelet Transform in the Time Series Similarity Paradigm. *Proceedings of the Third European Conference on Principles of Data Mining and Knowledge Discovery*. 12-22.
 17. Eamonn Keogh, Kaushik Chakrabarti, Michael Pazzani and Sharad Mehrotra (2001). Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases. *Knowledge and Information Systems*. 3(3): 263–286.
 18. Byoung-Kee Yi and Christos Faloutsos (2000). Fast Time Sequence Indexing for Arbitrary Lp Norms. *Proceedings of the 26th International Conference on Very Large Data Bases*. 385–394.

19. Jessica Lin, Eamonn Keogh, Stefano Lonardi, Stefano, Patel Pranav Patel (2002). Finding Motifs in Time Series. Proceedings of the Second Workshop on Temporal Data Mining.



425352952

CU iThesis 6070323921 thesis / recv: 06082562 14:42:31 / seq: 83

ภาคผนวก



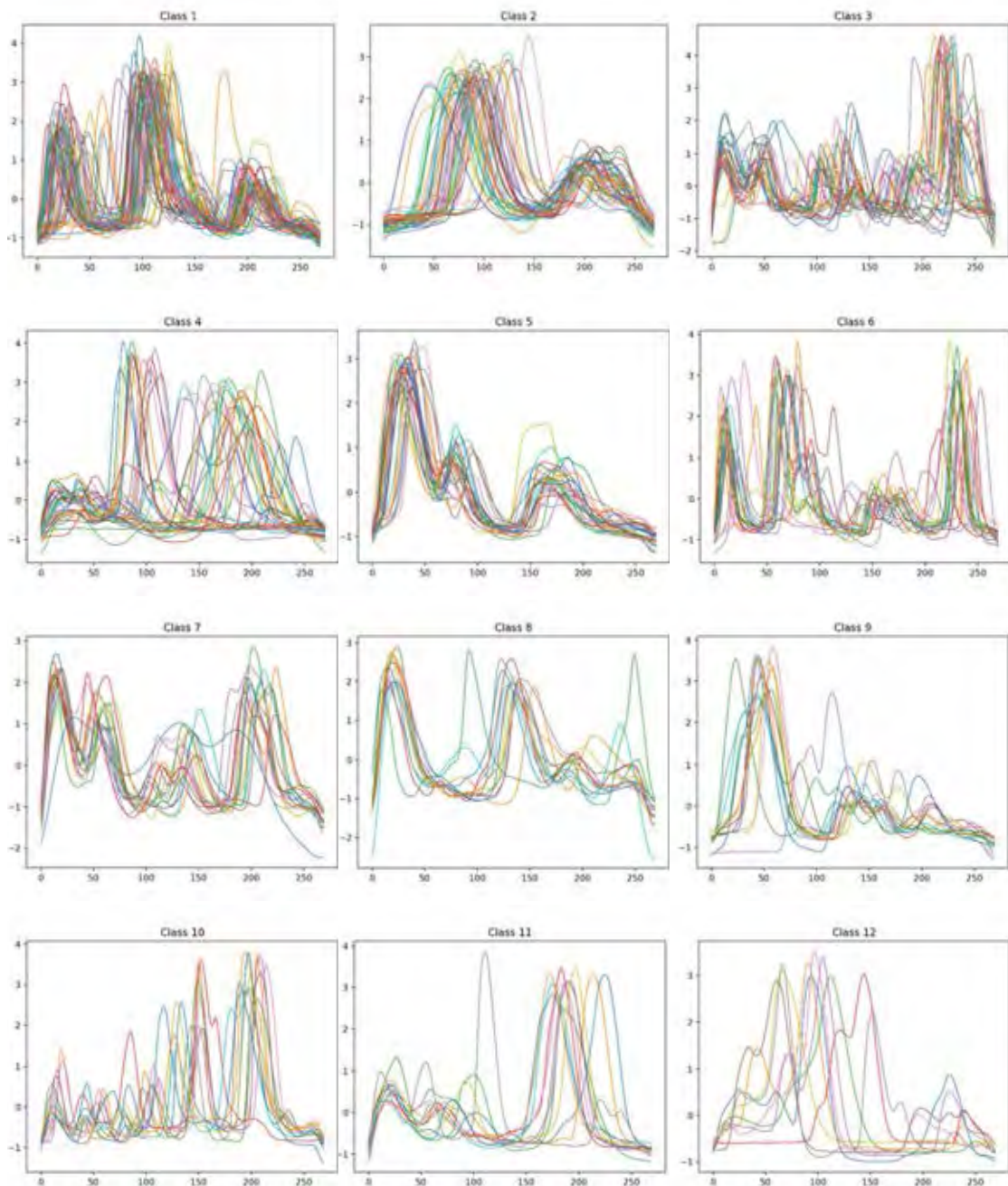
4253522952

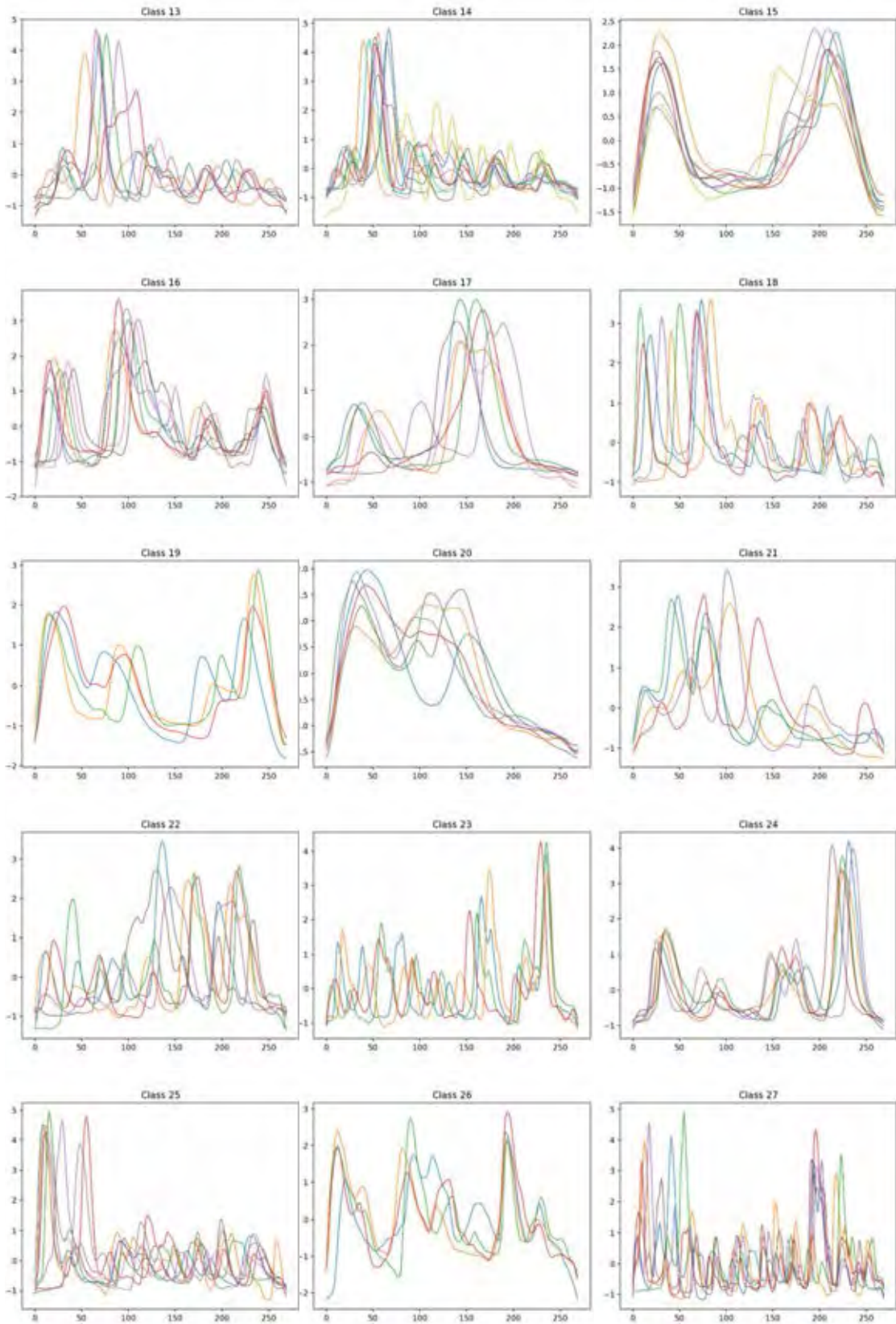
CU iThesis 607032321 thesis / recv: 06082562 14:42:31 / seq: 83

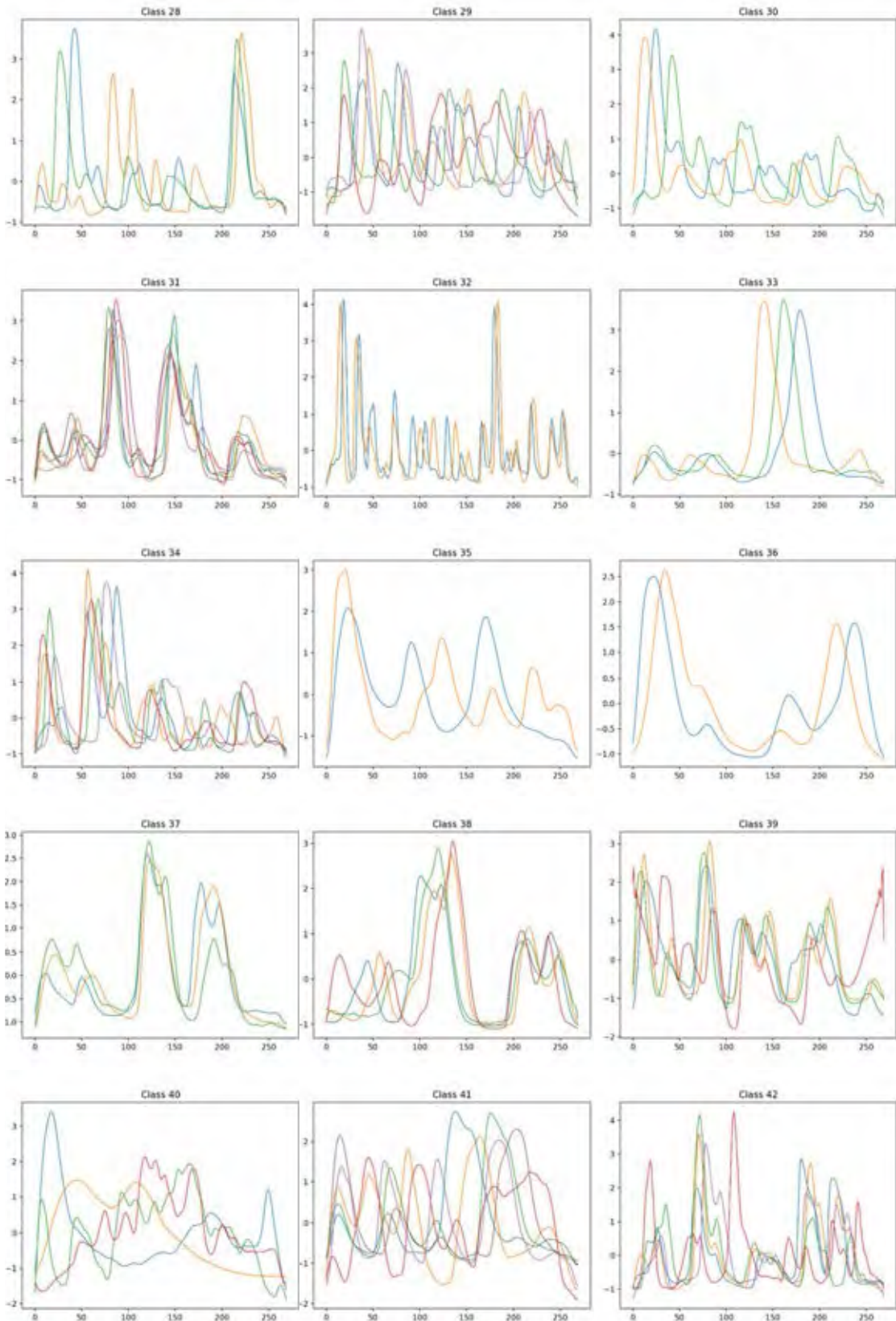
ภาคผนวก ก

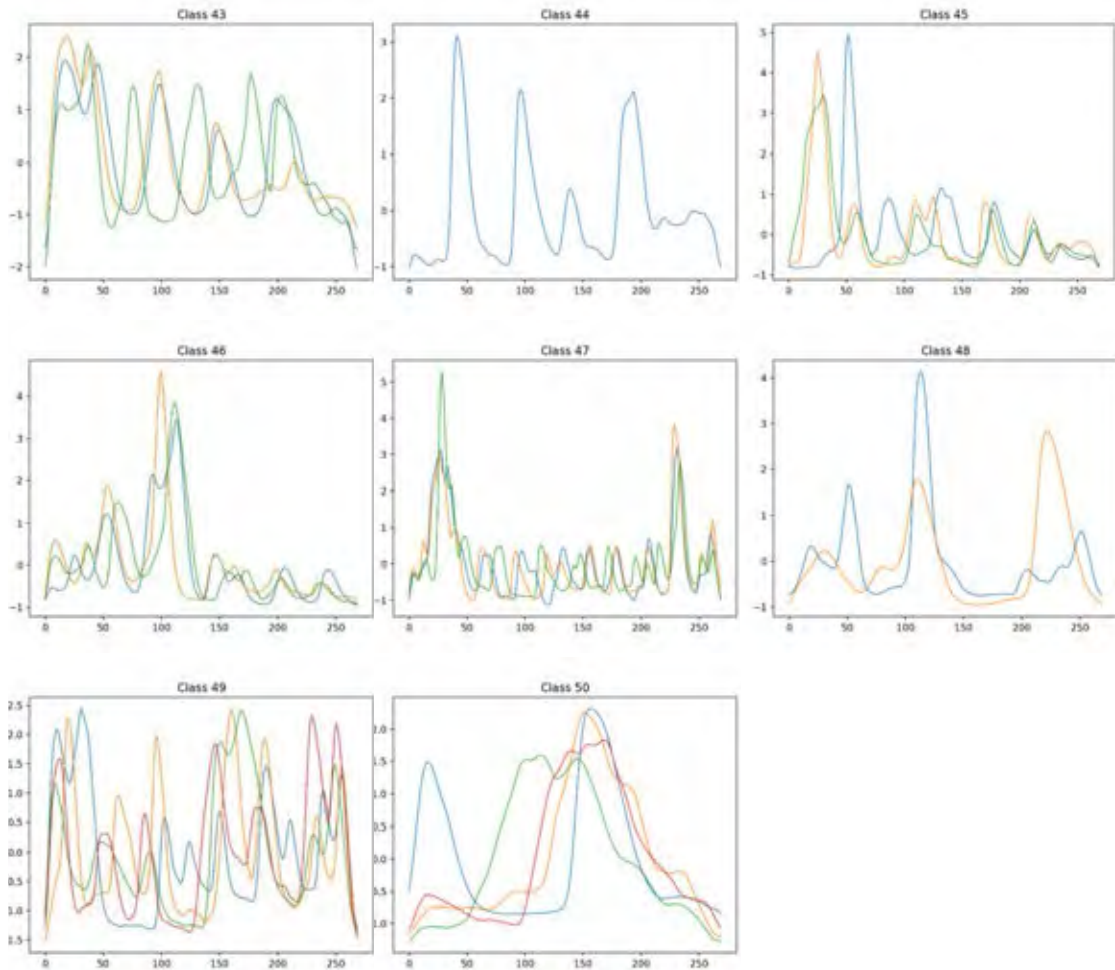
ข้อมูลที่ใช้ในการทดลองในงานวิทยานิพนธ์ มีทั้งหมด 78 ตัว จาก UCR Time Series Classification Archive [6] โดยแสดงข้อมูลและคลาสของข้อมูลทั้งหมดในชุดข้อมูลจากชุดข้อมูลการสอน

ตัวอย่างชุดข้อมูล 50Words มีข้อมูลทั้งหมด 50 คลาส



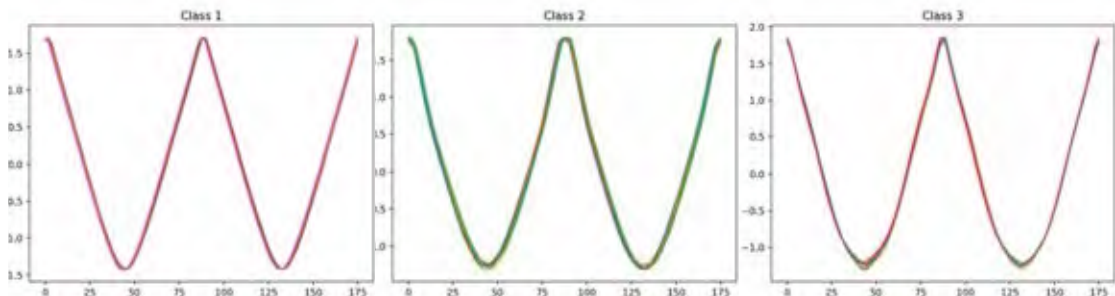


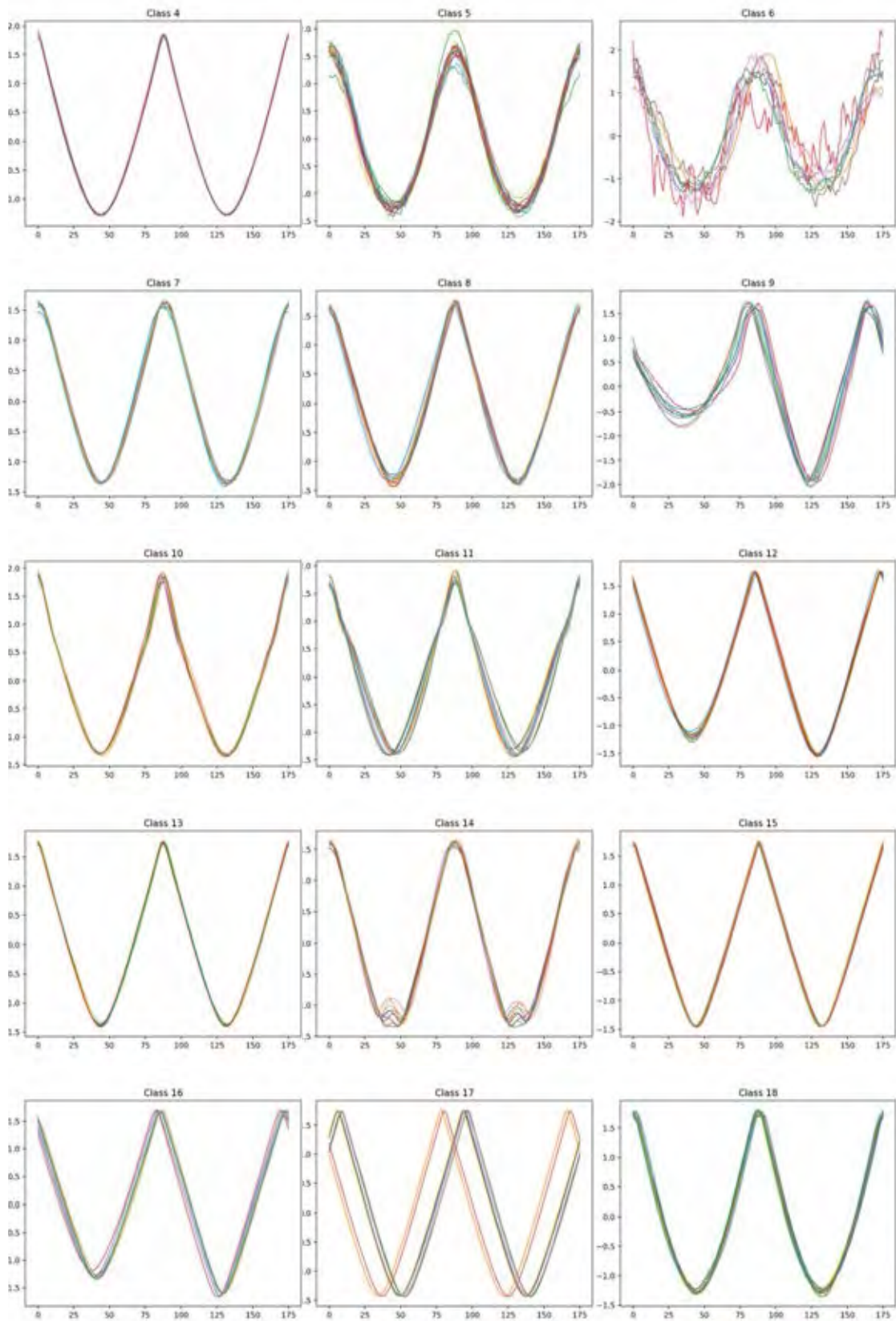


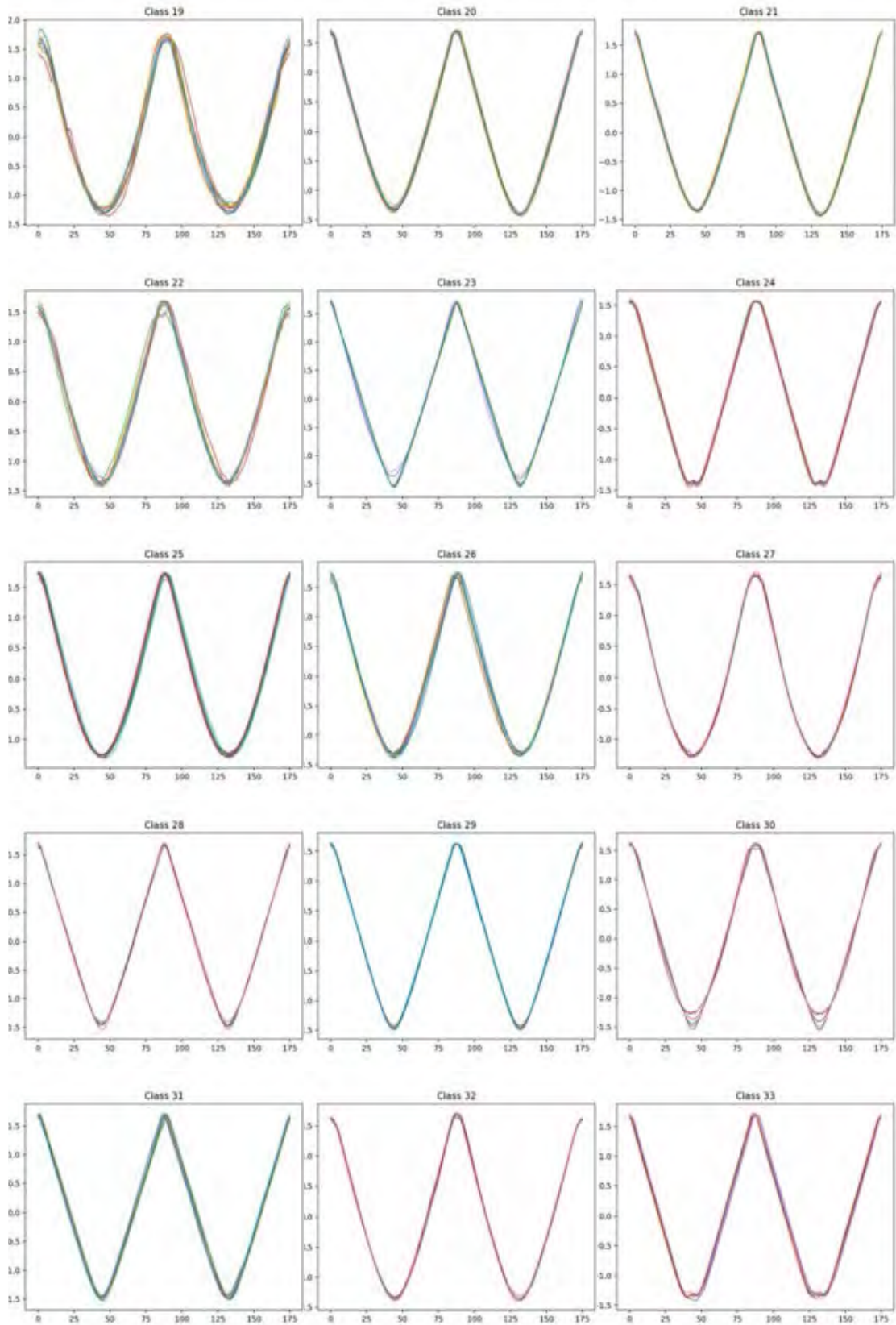


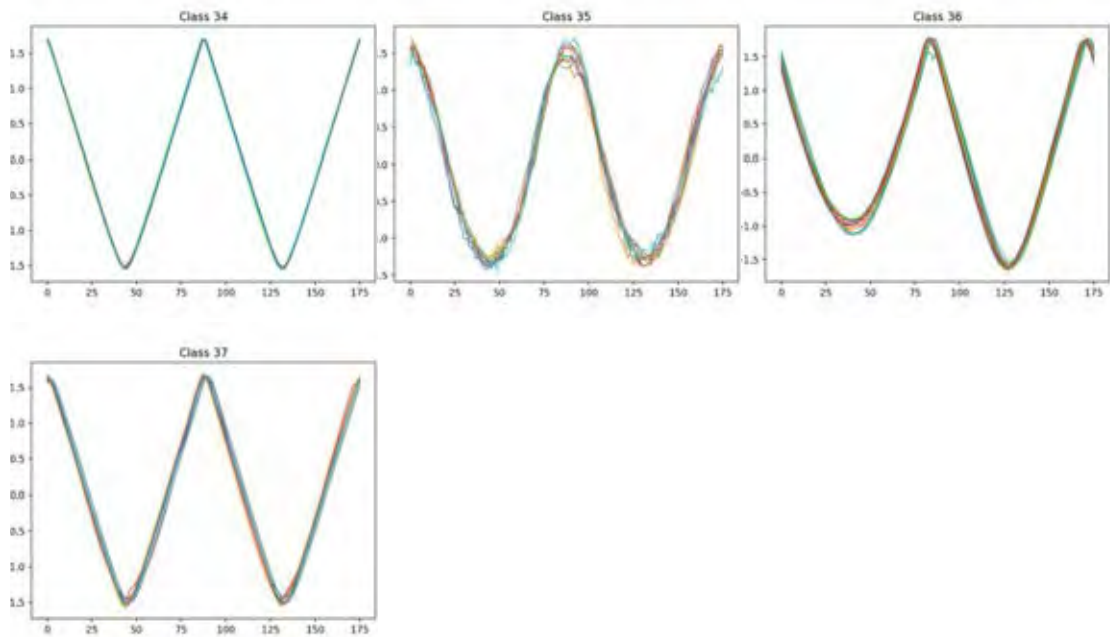
รูปที่ ก.1 กราฟแสดงคลาสของชุดข้อมูล 50Words

ตัวอย่างชุดข้อมูล Adiac มีข้อมูลทั้งหมด 37 คลาส



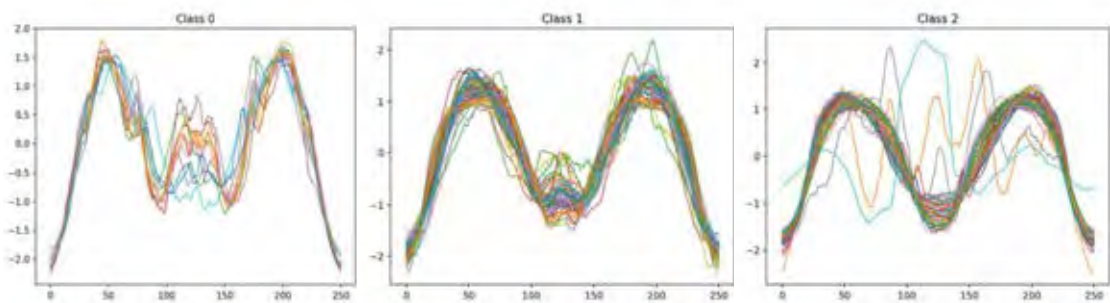






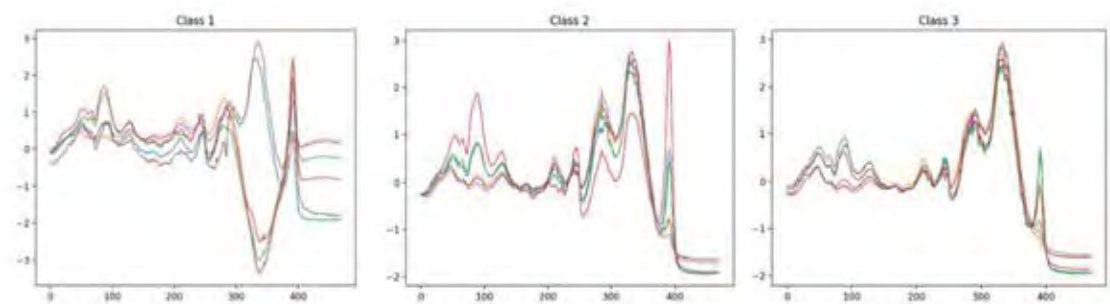
รูปที่ ก.2 กราฟแสดงคลาสของชุดข้อมูล Adiac

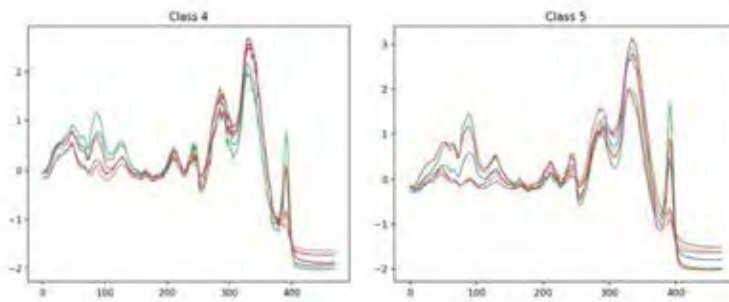
ตัวอย่างชุดข้อมูล ArrowHead มีข้อมูลทั้งหมด 3 คลาส



รูปที่ ก.3 กราฟแสดงคลาสของชุดข้อมูล ArrowHead

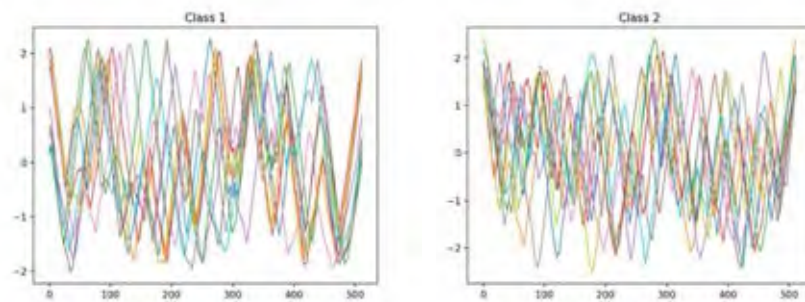
ตัวอย่างชุดข้อมูล Beef มีข้อมูลทั้งหมด 5 คลาส





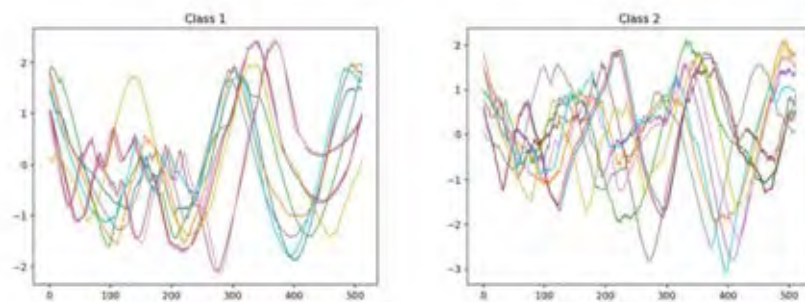
รูปที่ ก.4 กราฟแสดงคลาสของชุดข้อมูล Beef

ตัวอย่างชุดข้อมูล BeetleFly มีข้อมูลทั้งหมด 2 คลาส



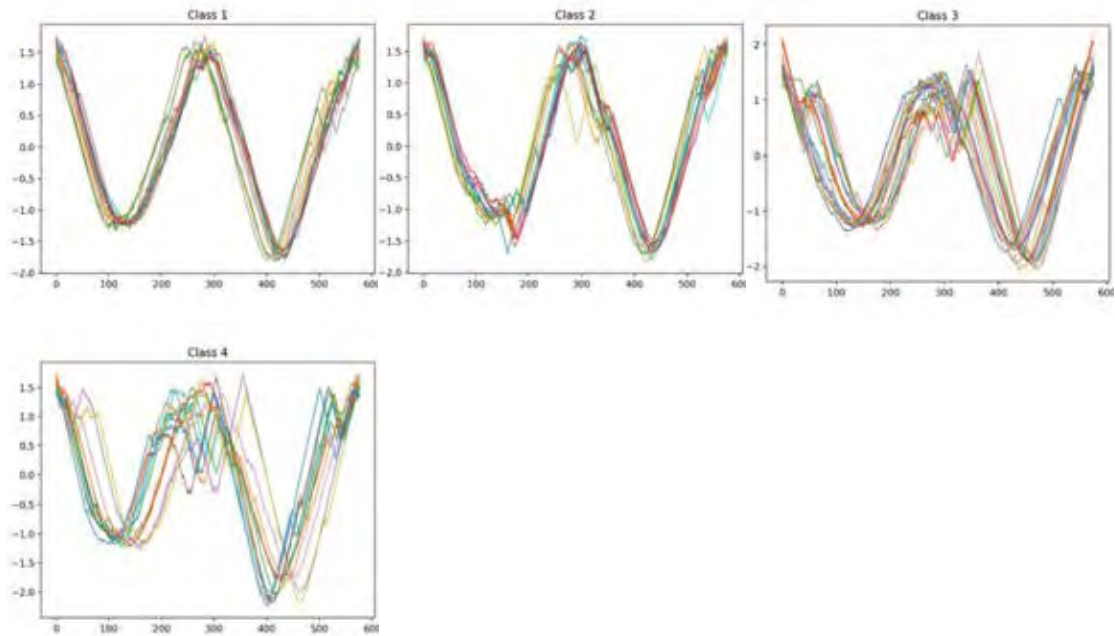
รูปที่ ก.5 กราฟแสดงคลาสของชุดข้อมูล BeetleFly

ตัวอย่างชุดข้อมูล BirdChicken มีข้อมูลทั้งหมด 2 คลาส



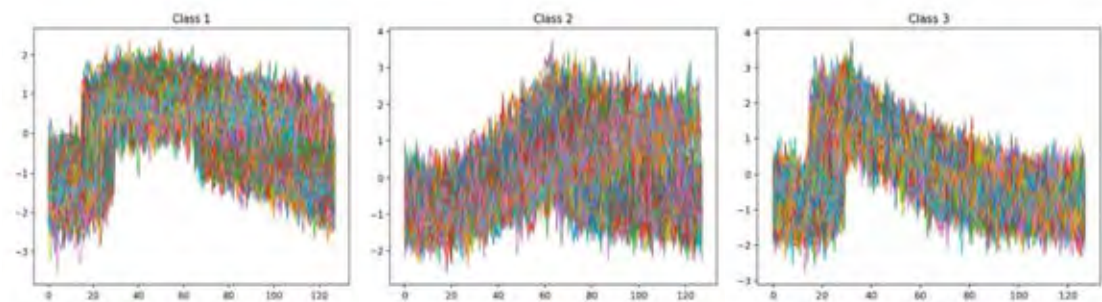
รูปที่ ก.6 กราฟแสดงคลาสของชุดข้อมูล BirdChicken

ตัวอย่างชุดข้อมูล Car มีข้อมูลทั้งหมด 4 คลาส



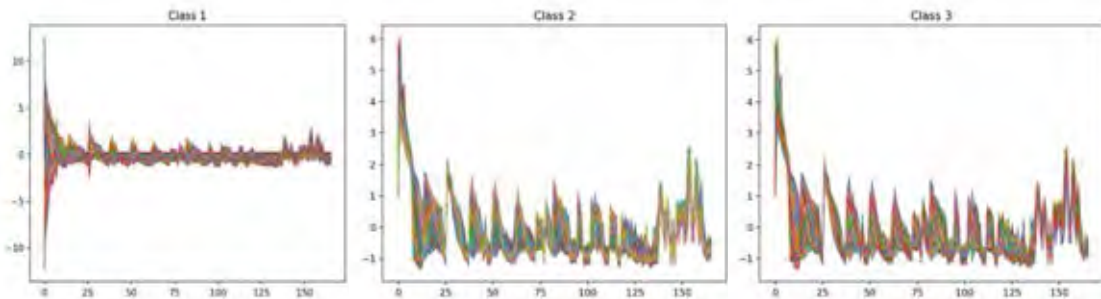
รูปที่ ก.7 กราฟแสดงในแต่ละคลาสของชุดข้อมูล Car

ตัวอย่างชุดข้อมูล CBF มีข้อมูลทั้งหมด 3 คลาส



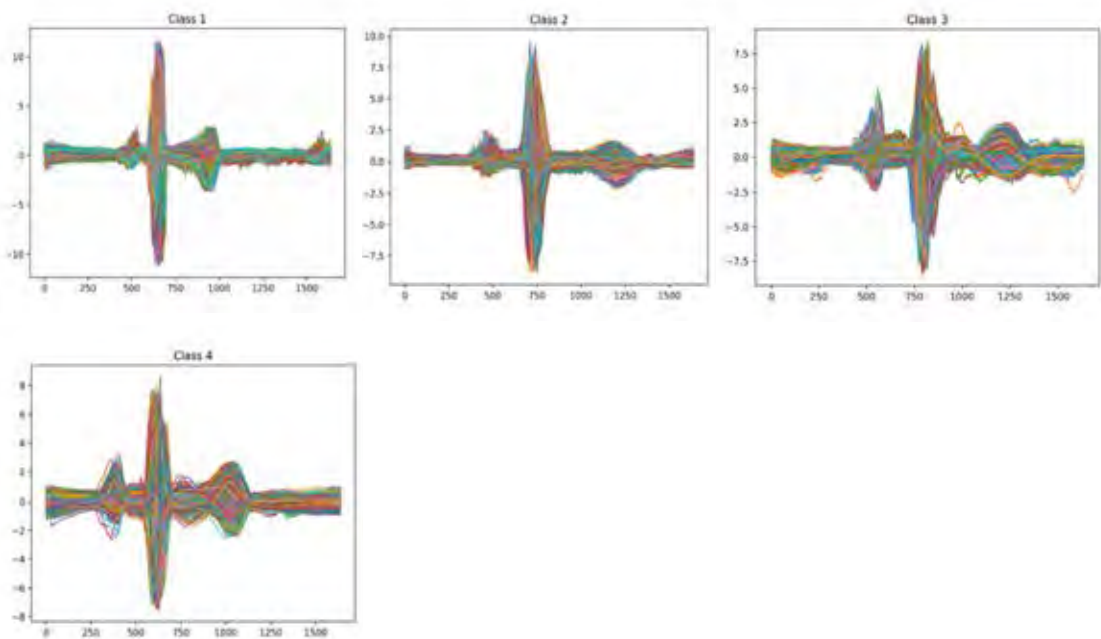
รูปที่ ก.8 กราฟแสดงคลาสของชุดข้อมูล CBF

ตัวอย่างชุดข้อมูล ChlorineConcentration มีข้อมูลทั้งหมด 3 คลาส



รูปที่ ก.9 กราฟแสดงคลาสของชุดข้อมูล ChlorineConcentration

ตัวอย่างชุดข้อมูล CinC_ECG_torso มีข้อมูลทั้งหมด 4 คลาส

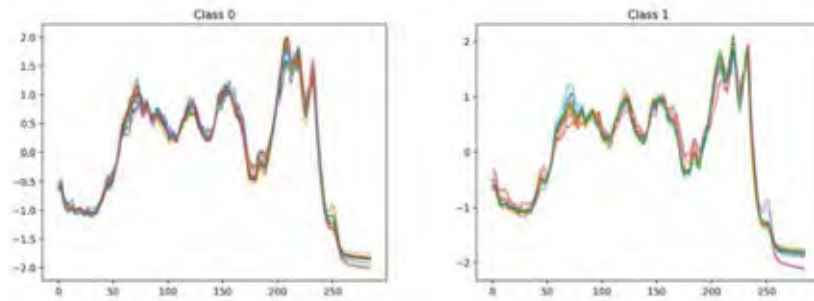


รูปที่ ก.10 กราฟแสดงคลาสของชุดข้อมูล CinC_ECG_torso



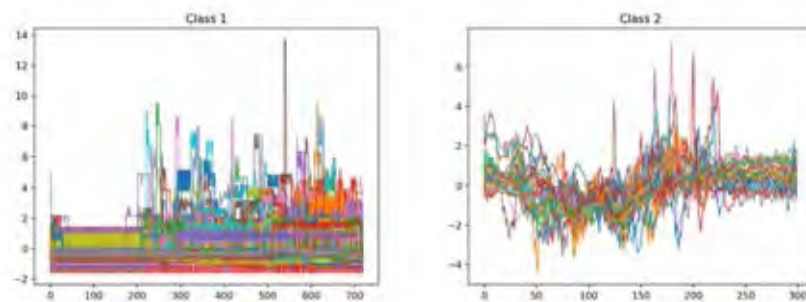
425352952

ตัวอย่างชุดข้อมูล Coffee มีข้อมูลทั้งหมด 2 คลาส



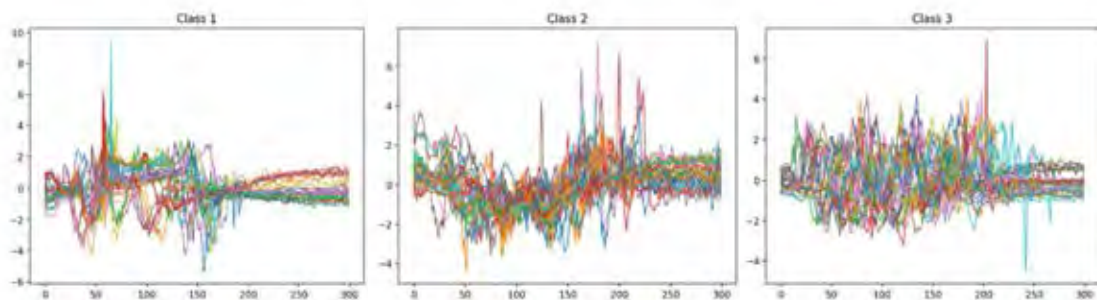
รูปที่ ก.11 กราฟแสดงคลาสของชุดข้อมูล Coffee

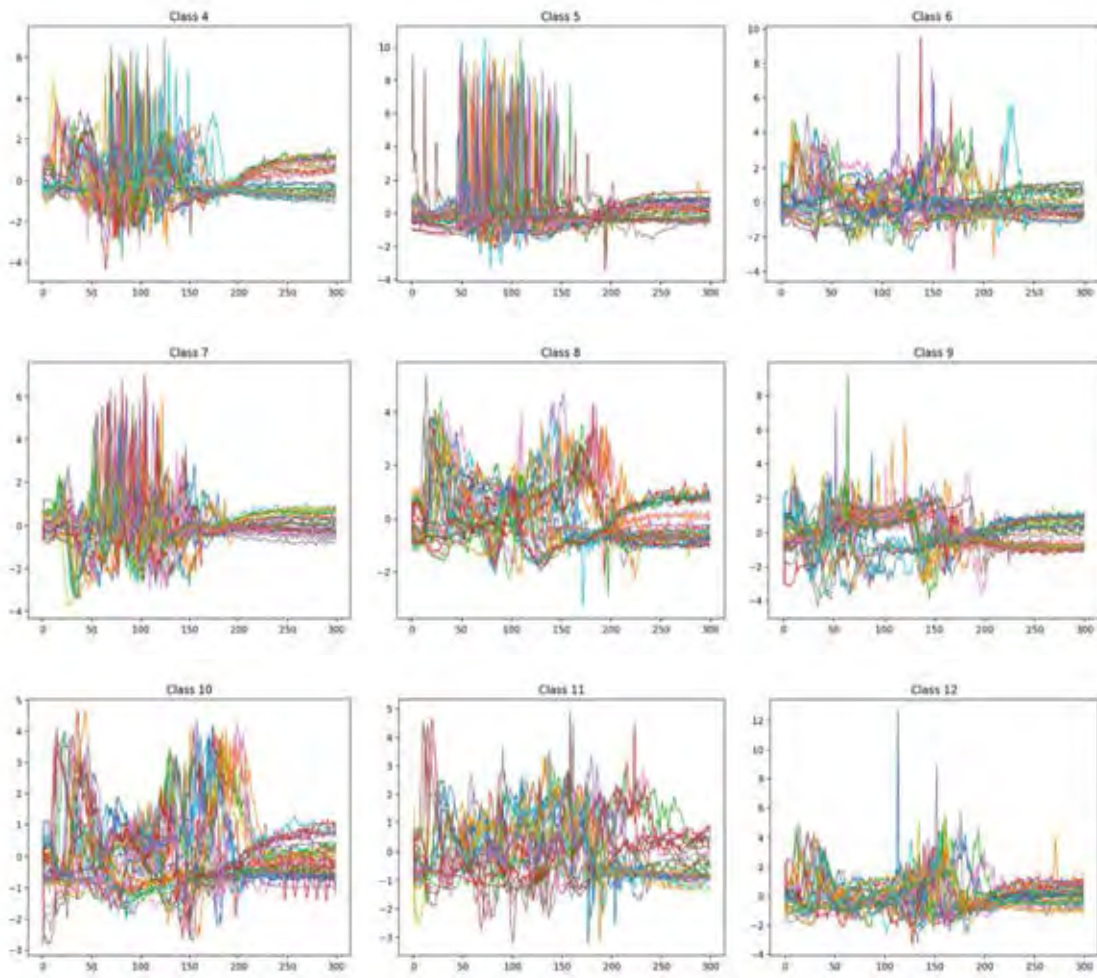
ตัวอย่างชุดข้อมูล Computers มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.12 กราฟแสดงคลาสของชุดข้อมูล Computers

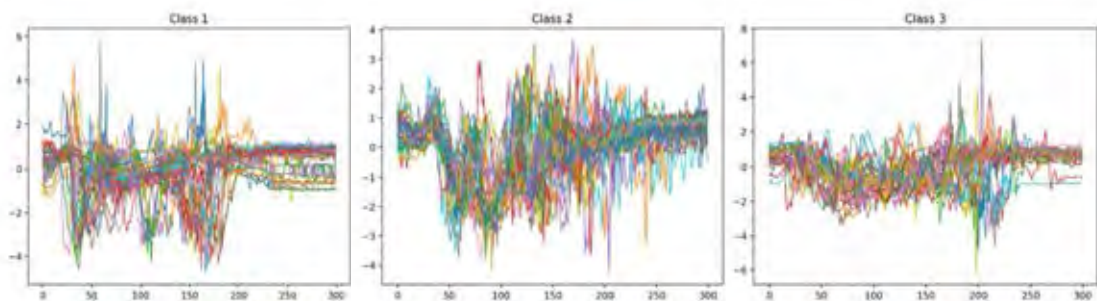
ตัวอย่างชุดข้อมูล Cricket_X มีข้อมูลทั้งหมด 12 คลาส

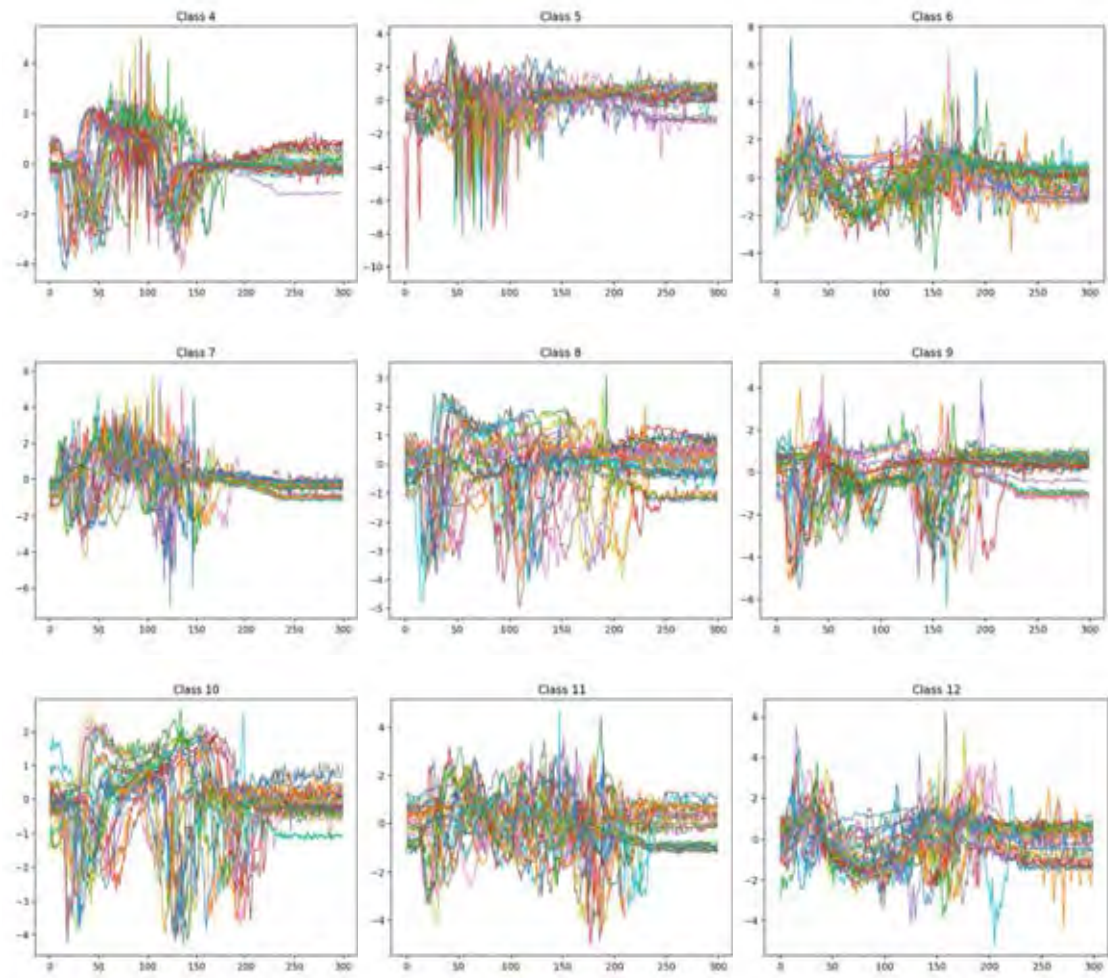




รูปที่ ก.13 กราฟแสดงคลาสของชุดข้อมูล Cricket_X

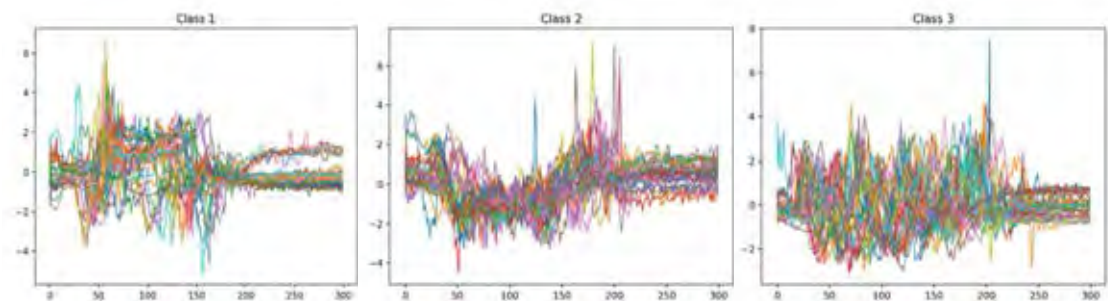
ตัวอย่างชุดข้อมูล Cricket_Y มีข้อมูลทั้งหมด 12 คลาส

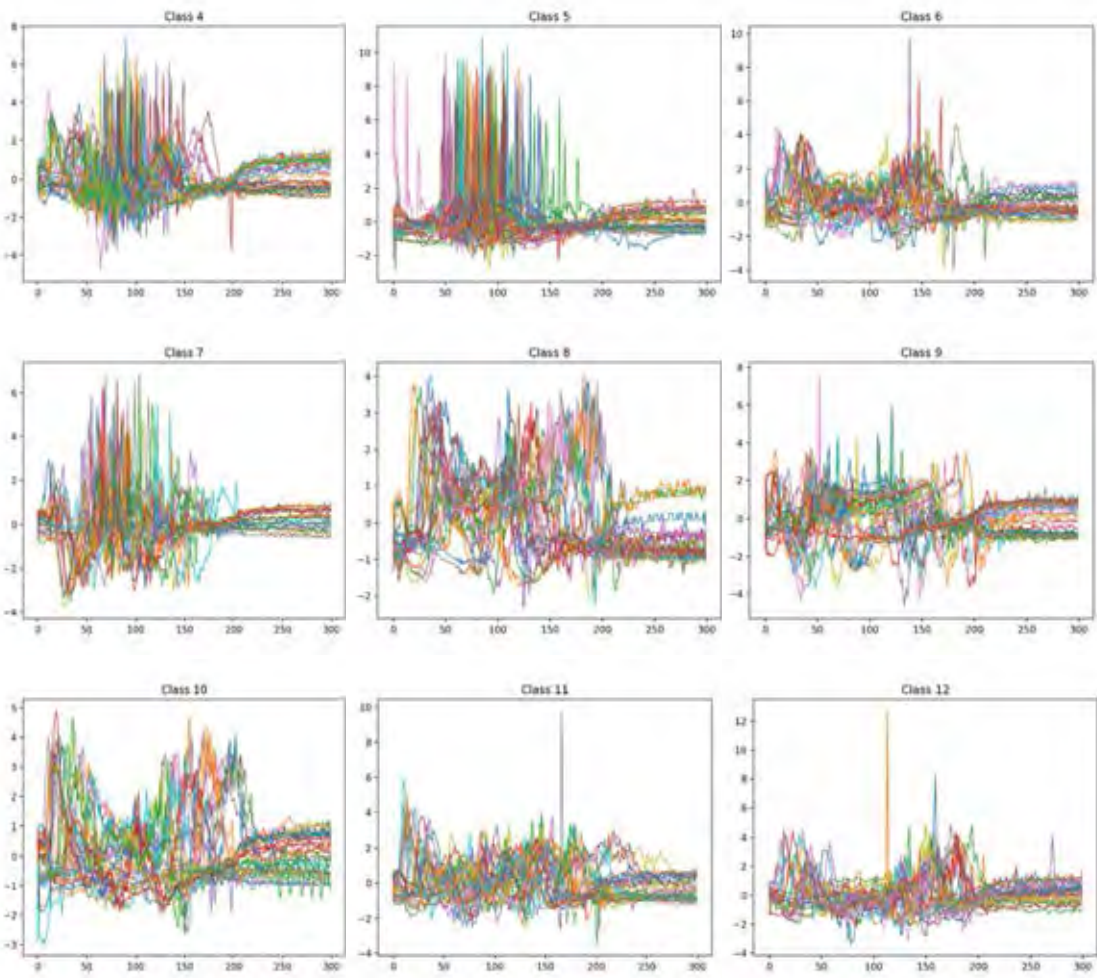




รูปที่ ก.14 กราฟแสดงคลาสของชุดข้อมูล Cricket_Y

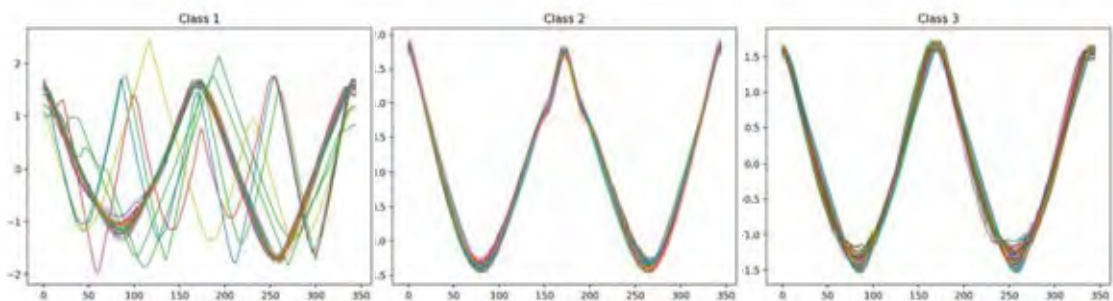
ตัวอย่างชุดข้อมูล Cricket_Z มีข้อมูลทั้งหมด 12 คลาส

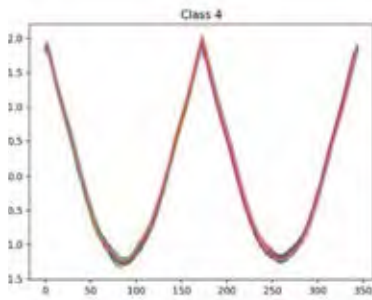




รูปที่ ก.15 กราฟแสดงคลาสของชุดข้อมูล Cricket_Z

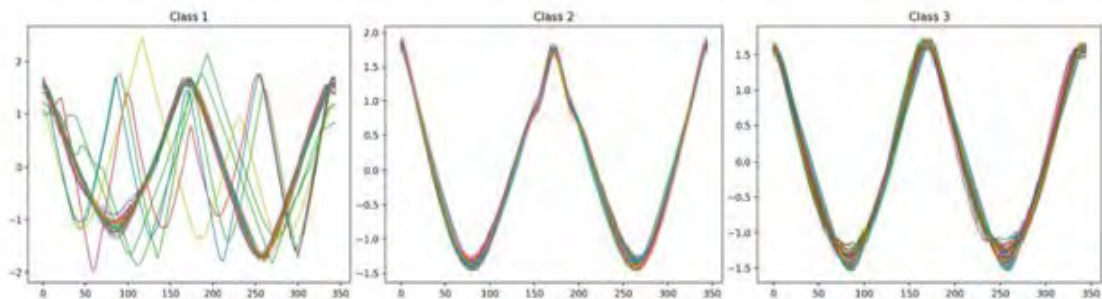
ตัวอย่างชุดข้อมูล DiatomSizeReduction มีข้อมูลทั้งหมด 4 ในแต่ละคลาส





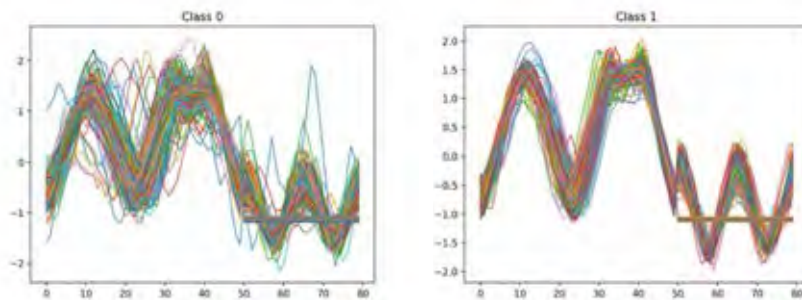
รูปที่ ก.16 กราฟแสดงคลาสของชุดข้อมูล DiatomSizeReduction

ตัวอย่างชุดข้อมูล DistalPhalanxOutlineAgeGroup มีข้อมูลทั้งหมด 3 คลาส



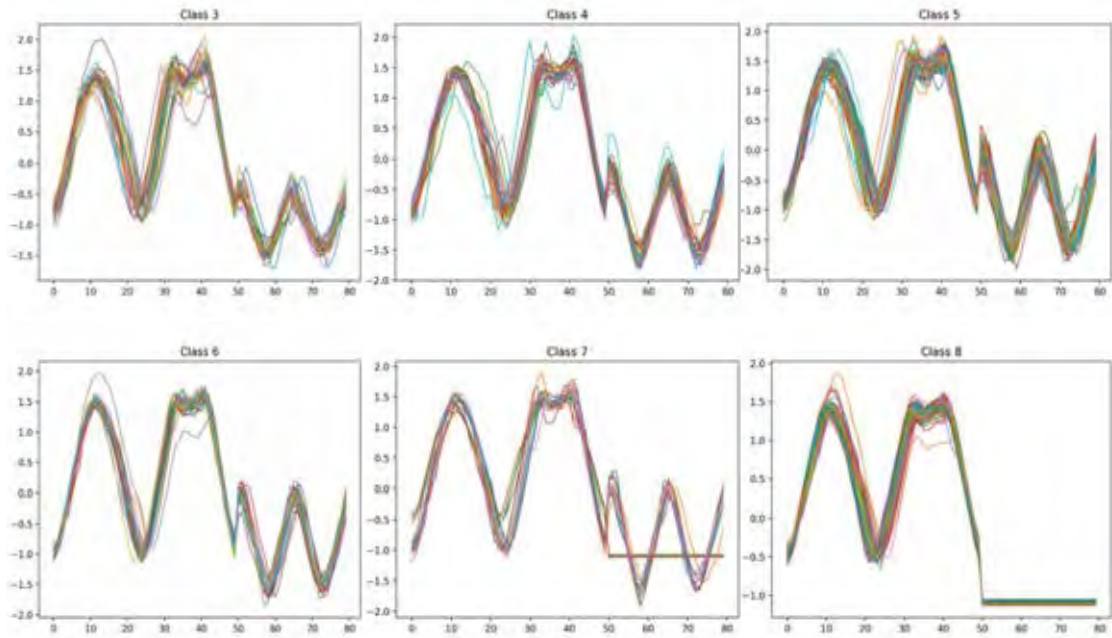
รูปที่ ก.17 กราฟแสดงคลาสของชุดข้อมูล DistalPhalanxOutlineAgeGroup

ตัวอย่างชุดข้อมูล DistalPhalanxOutlineCorrect มีข้อมูลทั้งหมด 2 คลาส



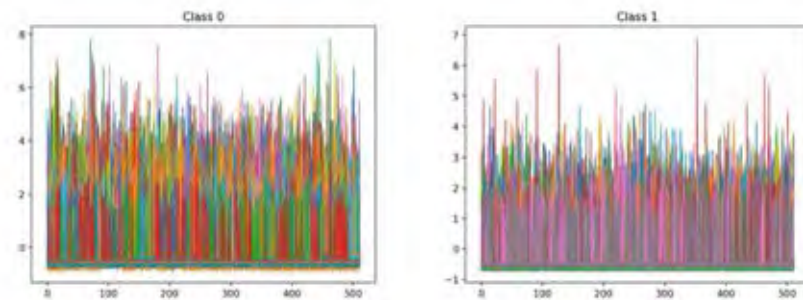
รูปที่ ก.18 กราฟแสดงคลาสของชุดข้อมูล DistalPhalanxOutlineCorrect

ตัวอย่างชุดข้อมูล DistalPhalanxTW มีข้อมูลทั้งหมด 6 คลาส



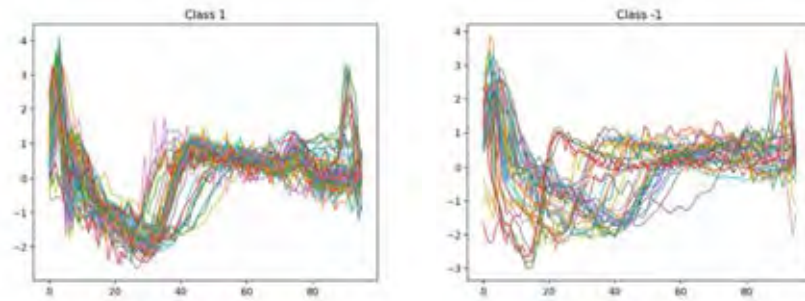
รูปที่ ก.19 กราฟแสดงคลาสของชุดข้อมูล DistalPhalanxTW

ตัวอย่างชุดข้อมูล Earthquakes มีข้อมูลทั้งหมด 2 คลาส



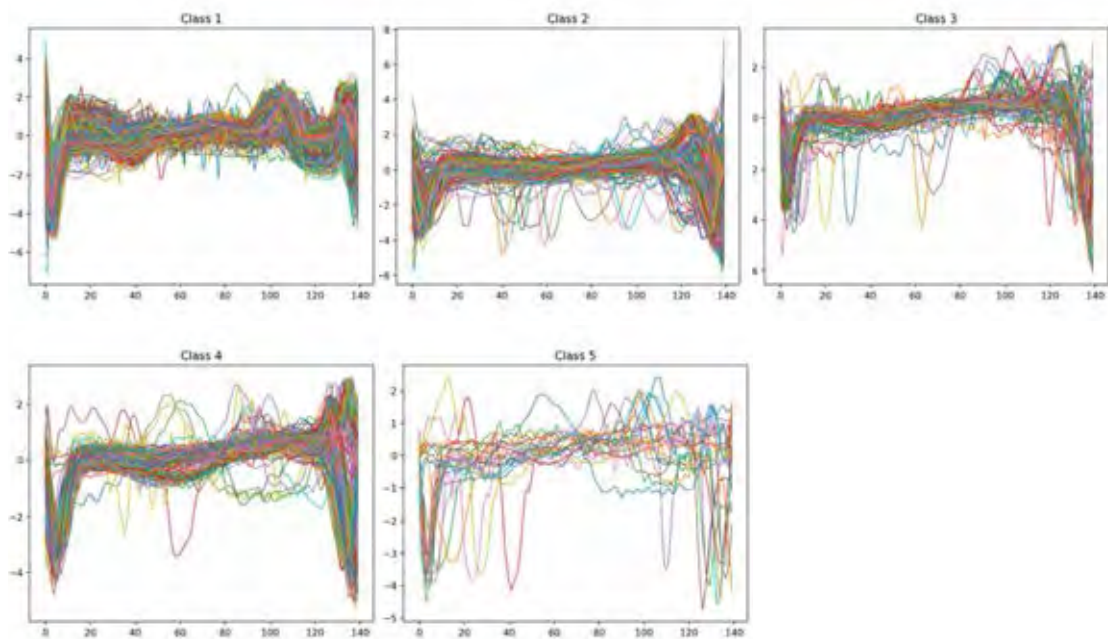
รูปที่ ก.20 กราฟแสดงคลาสของชุดข้อมูล Earthquakes

ตัวอย่างชุดข้อมูล ECG200 มีข้อมูลทั้งหมด 2 คลาส



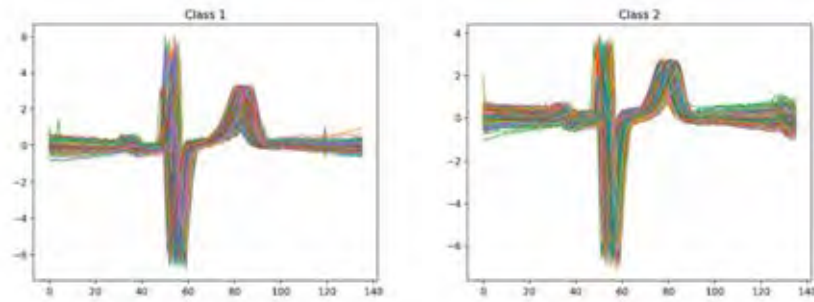
รูปที่ ก.21 กราฟแสดงคลาสของชุดข้อมูล ECG200

ตัวอย่างชุดข้อมูล ECG5000 มีข้อมูลทั้งหมด 5 คลาส



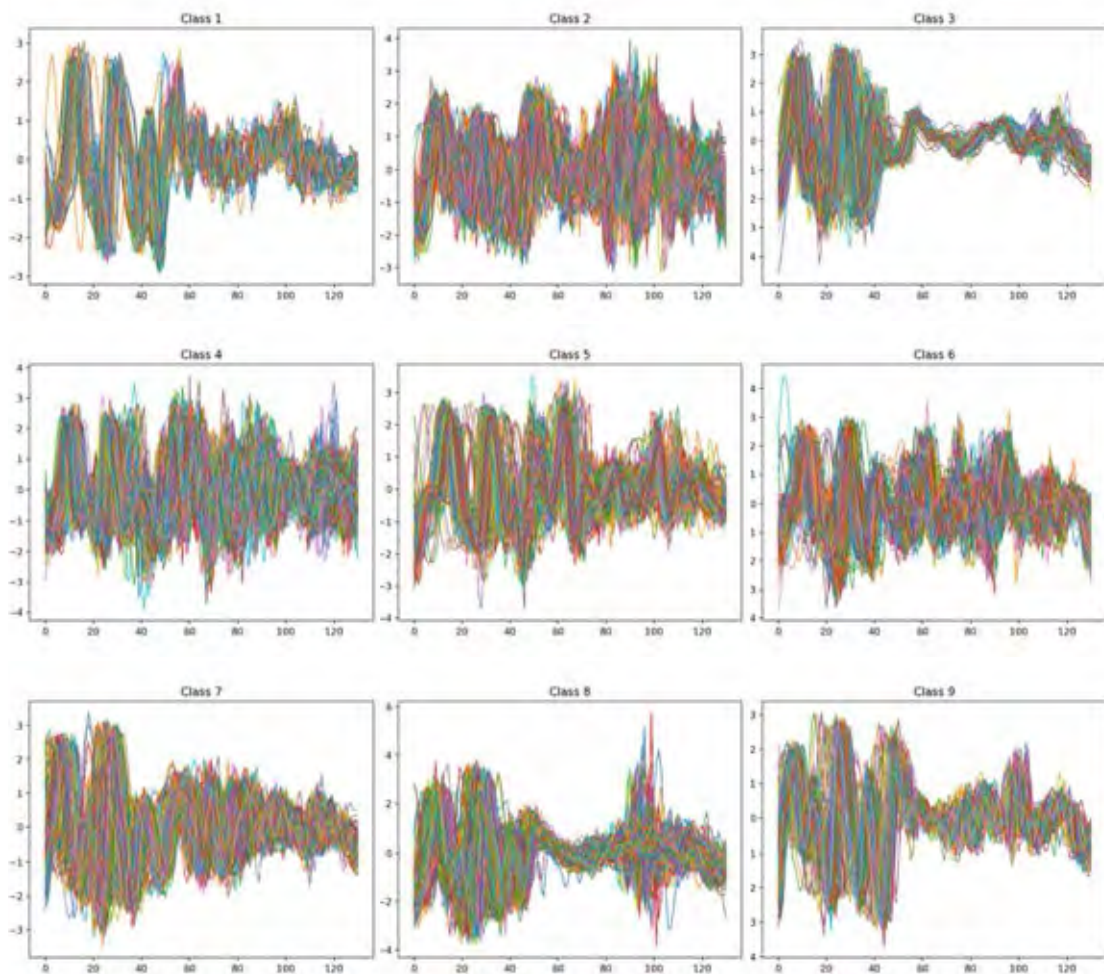
รูปที่ ก.22 กราฟแสดงคลาสของชุดข้อมูล ECG5000

ตัวอย่างชุดข้อมูล ECGFiveDays มีข้อมูลทั้งหมด 2 คลาส

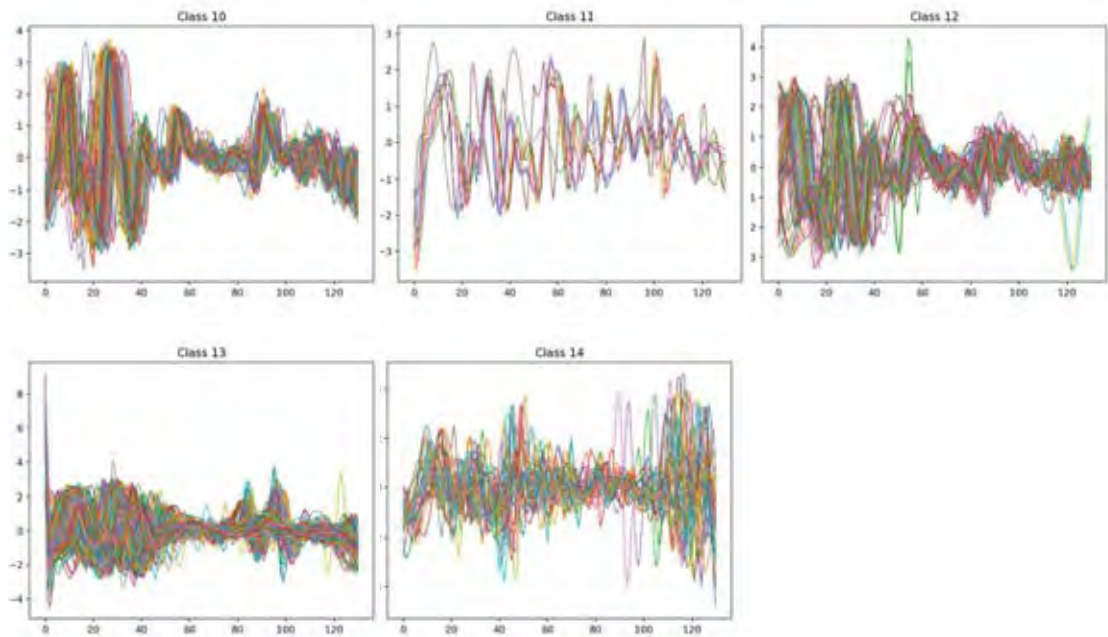


รูปที่ ก.23 กราฟแสดงคลาสของชุดข้อมูล ECGFiveDays

ตัวอย่างชุดข้อมูล Face (all) มีข้อมูลทั้งหมด 14 คลาส

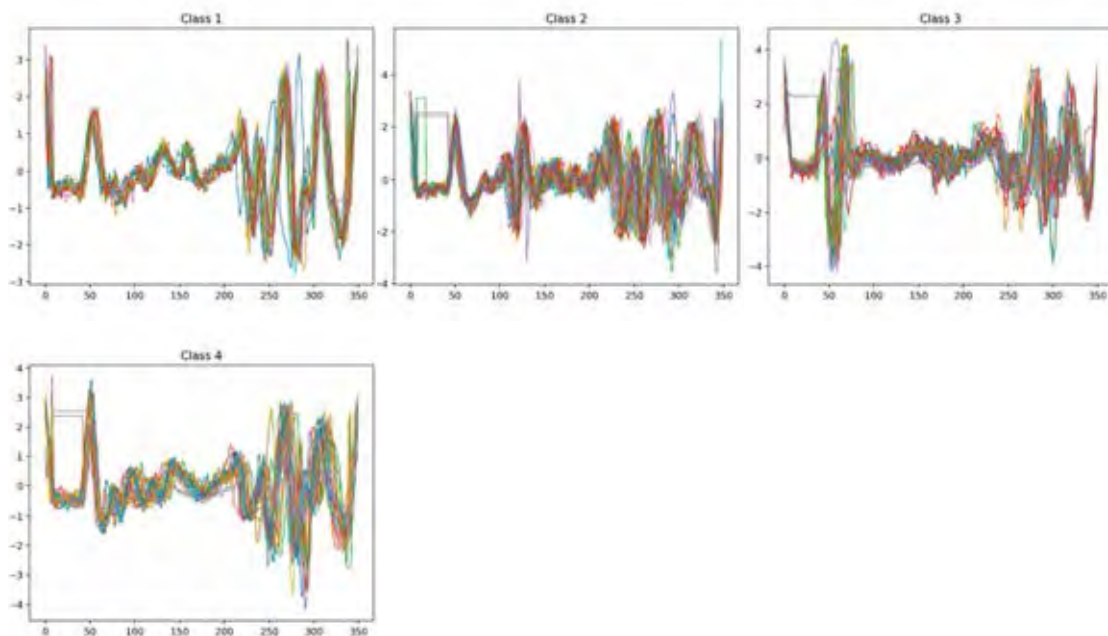


425352952



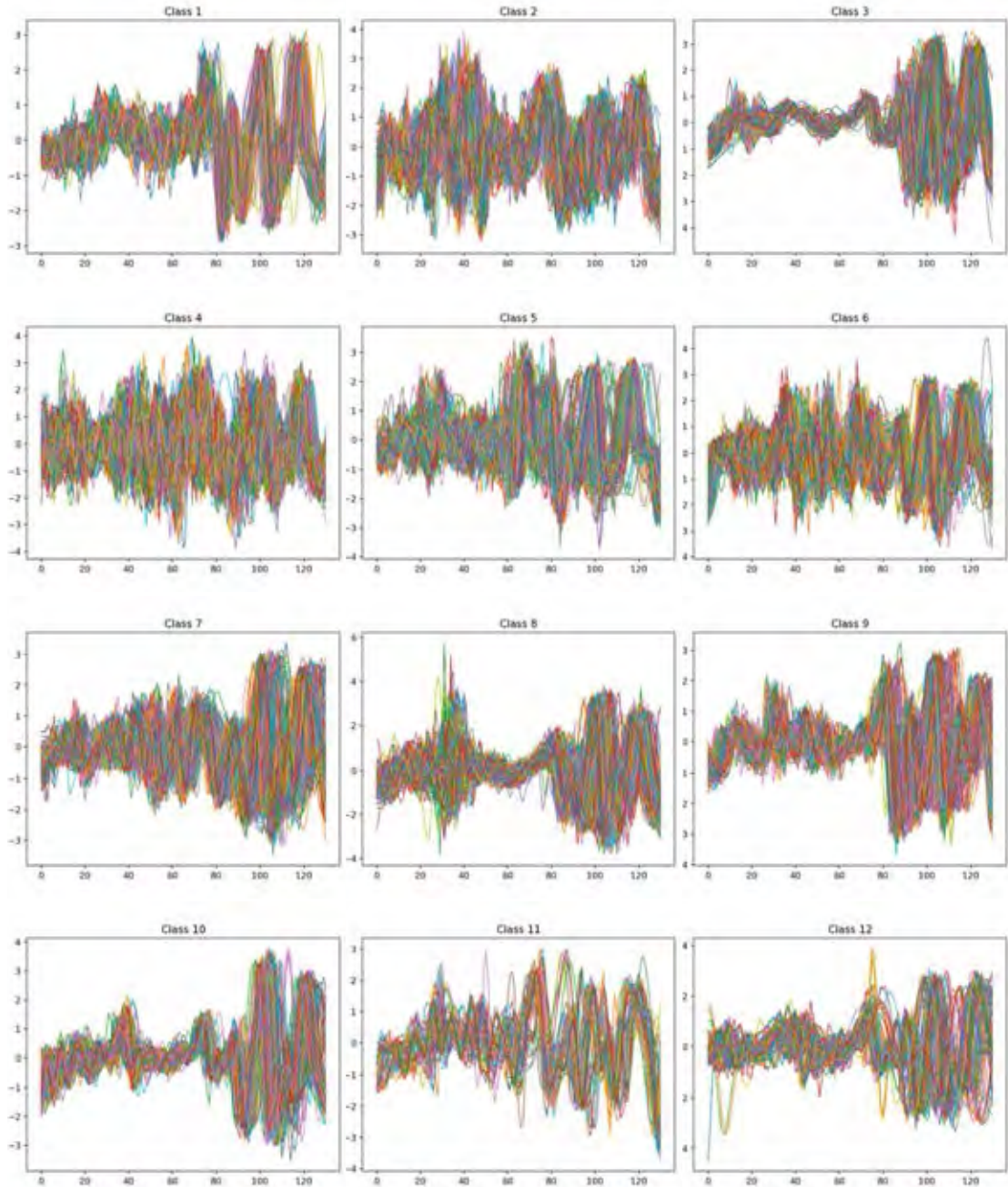
รูปที่ ก.24 กราฟแสดงคลาสของชุดข้อมูล Face (all)

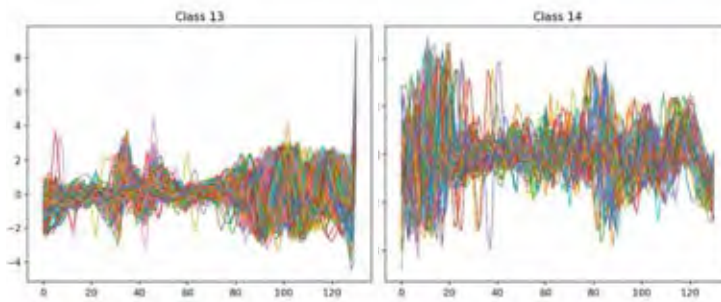
ตัวอย่างชุดข้อมูล Face (four) มีข้อมูลทั้งหมด 4 คลาส



รูปที่ ก.25 กราฟแสดงคลาสของชุดข้อมูล Face (four)

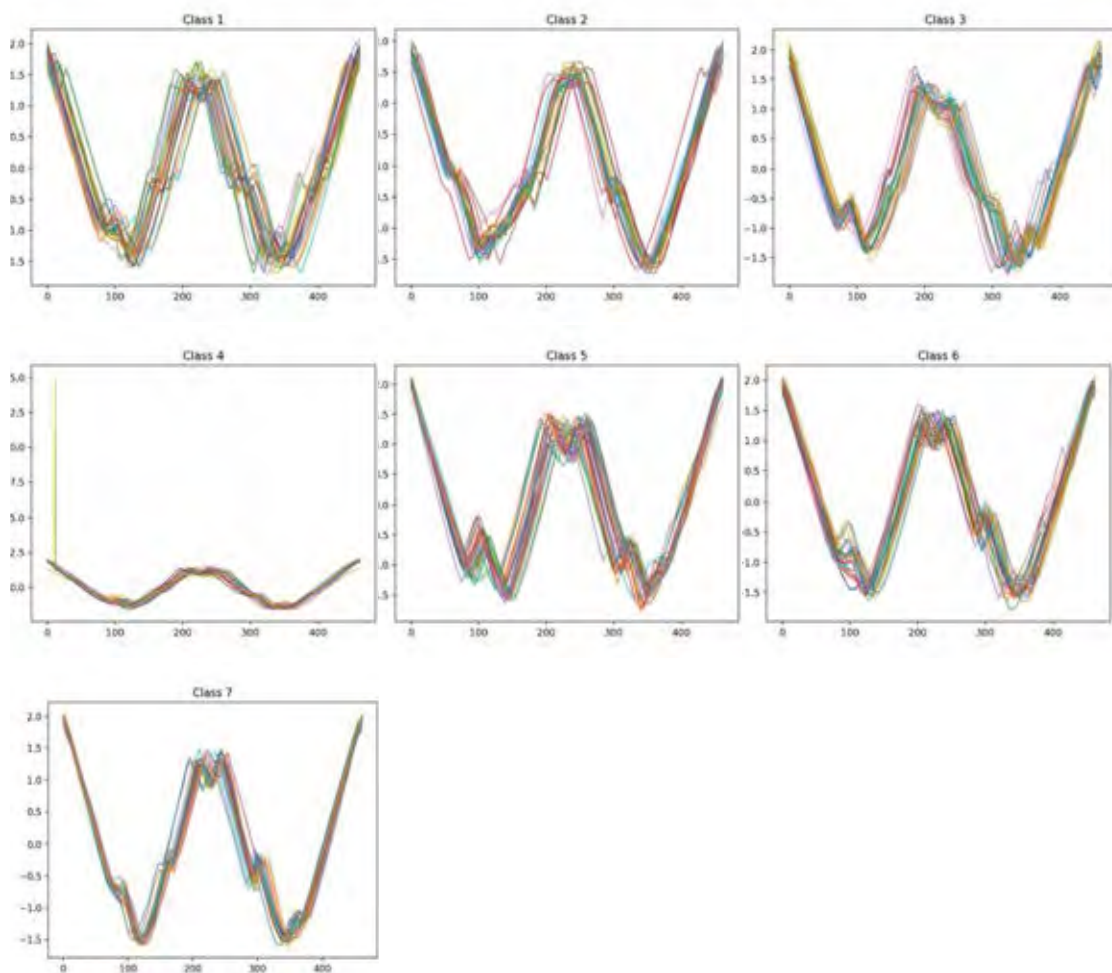
ตัวอย่างชุดข้อมูล FacesUCR มีข้อมูลทั้งหมด 14 คลาส





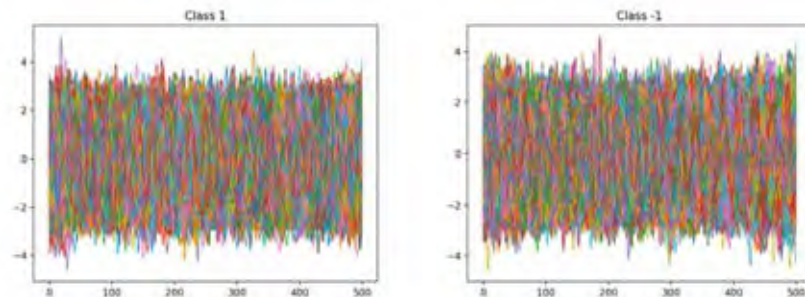
รูปที่ ก.26 กราฟแสดงคลาสของชุดข้อมูล FacesUCR

ตัวอย่างชุดข้อมูล Fish มีข้อมูลทั้งหมด 7 คลาส



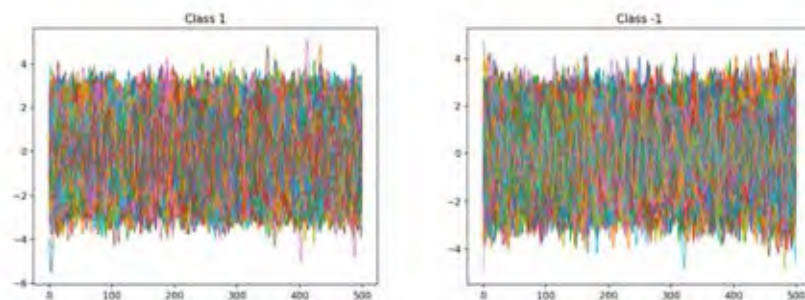
รูปที่ ก.27 กราฟแสดงคลาสของชุดข้อมูล Fish

ตัวอย่างชุดข้อมูล FordA มีข้อมูลทั้งหมด 2 คลาส



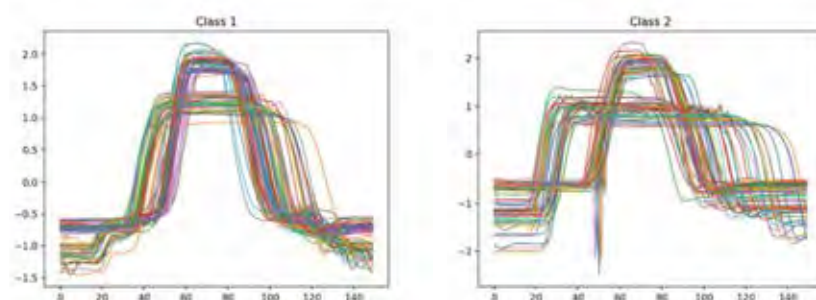
รูปที่ ก.28 กราฟแสดงคลาสของชุดข้อมูล FordA

ตัวอย่างชุดข้อมูล FordB มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.29 กราฟแสดงคลาสของชุดข้อมูล FordB

ตัวอย่างชุดข้อมูล Gun-Point มีข้อมูลทั้งหมด 2 คลาส

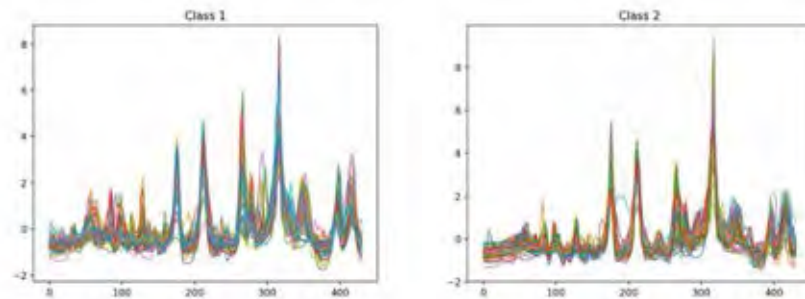


รูปที่ ก.30 กราฟแสดงคลาสของชุดข้อมูล Gun-Point



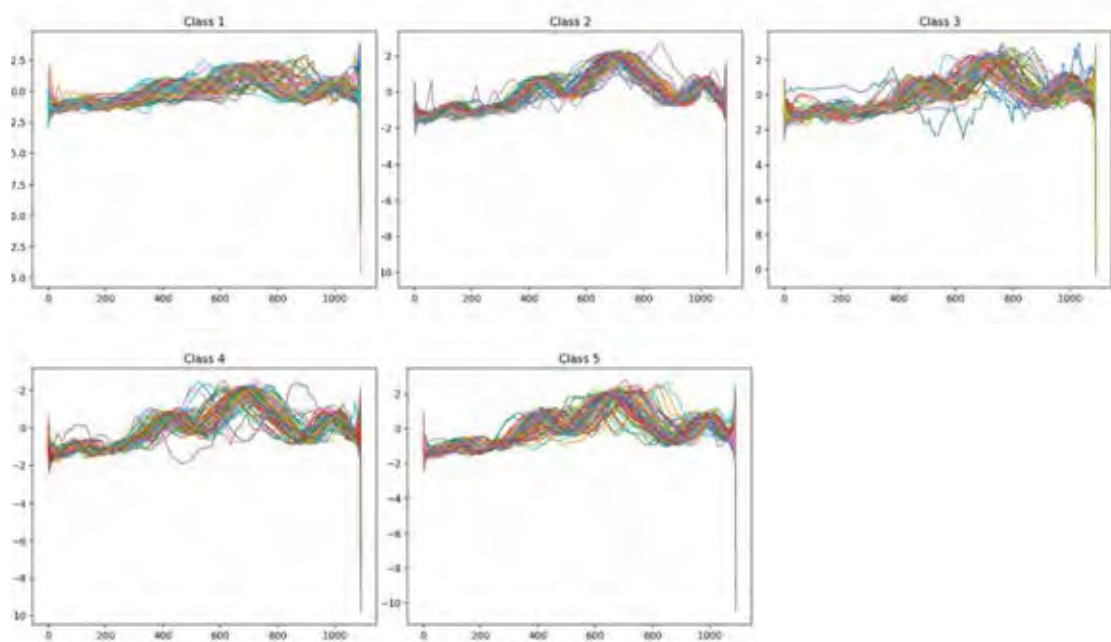
425352952

ตัวอย่างชุดข้อมูล Ham มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.31 กราฟแสดงคลาสของชุดข้อมูล Ham

ตัวอย่างชุดข้อมูล Haptics มีข้อมูลทั้งหมด 5 คลาส

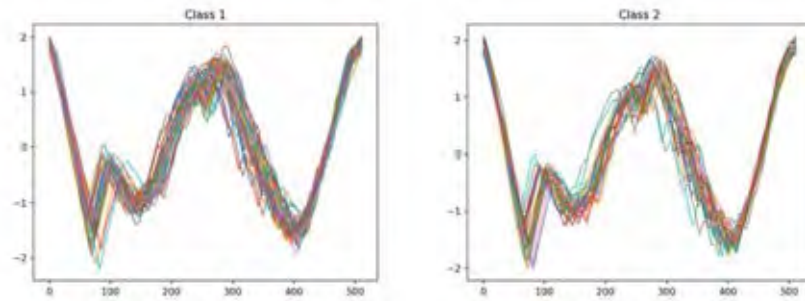


รูปที่ ก.32 กราฟแสดงคลาสของชุดข้อมูล Haptics



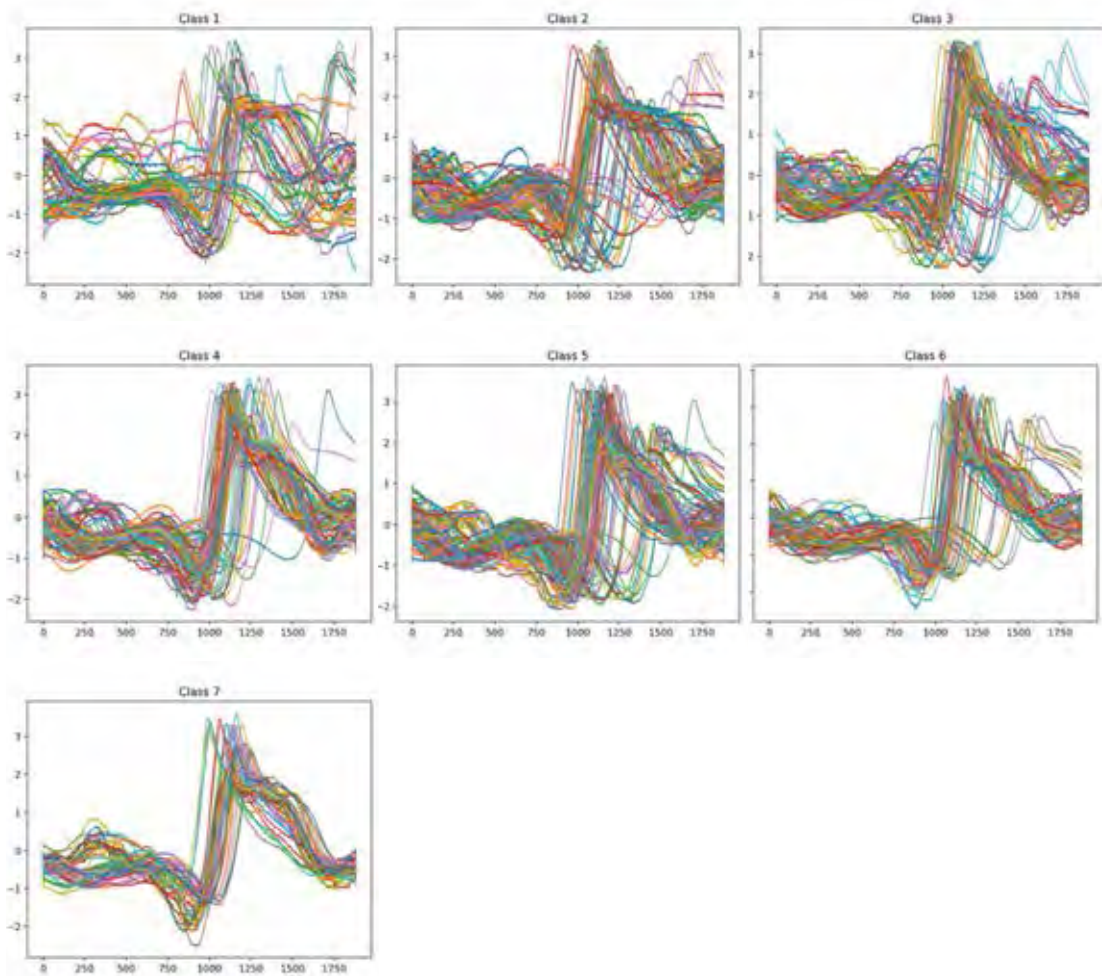
4253529252

ตัวอย่างชุดข้อมูล Herring มีข้อมูลทั้งหมด 2 คลาส



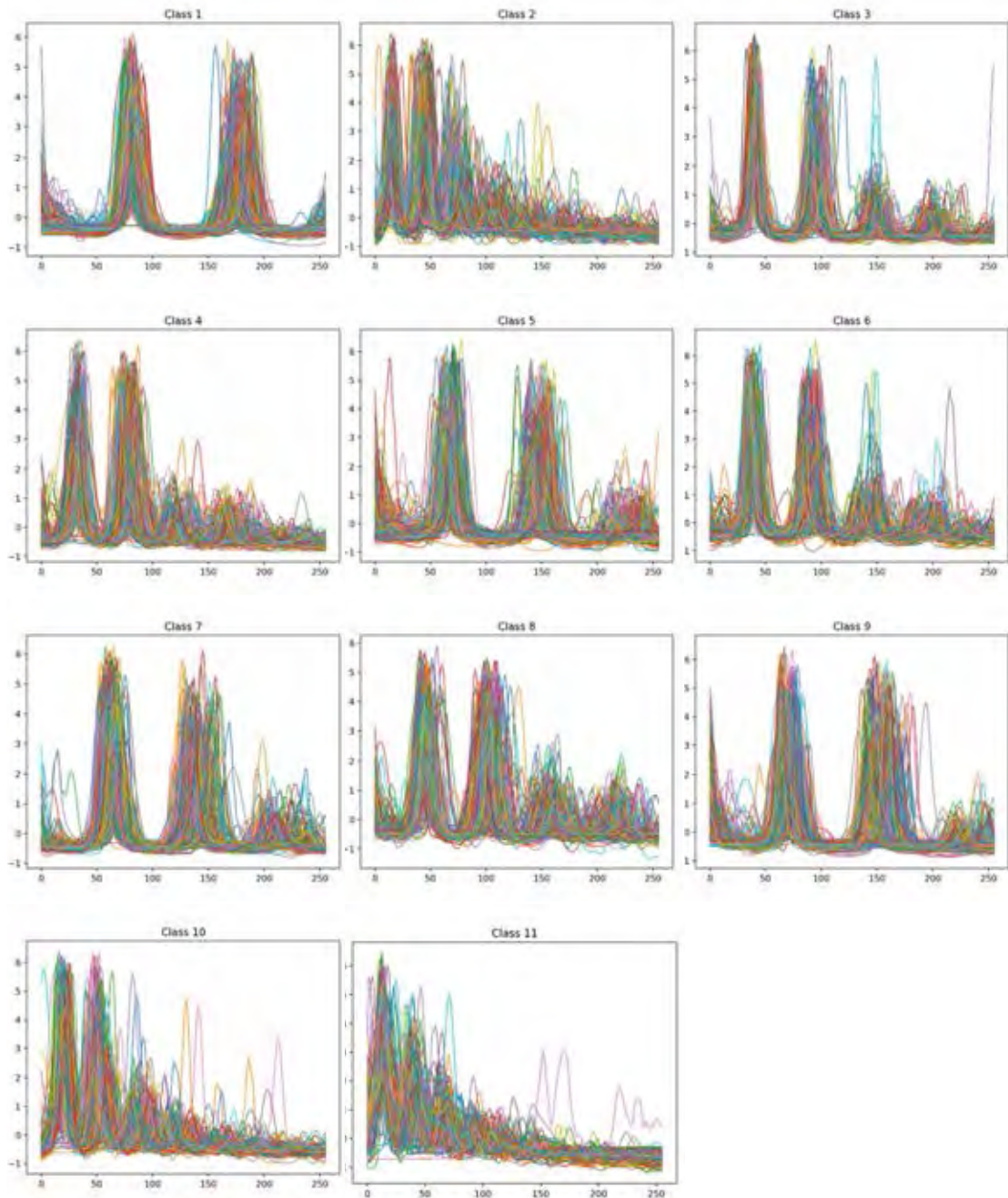
รูปที่ ก.33 กราฟแสดงคลาสของชุดข้อมูล Herring

ตัวอย่างชุดข้อมูล InlineSkate มีข้อมูลทั้งหมด 7 คลาส



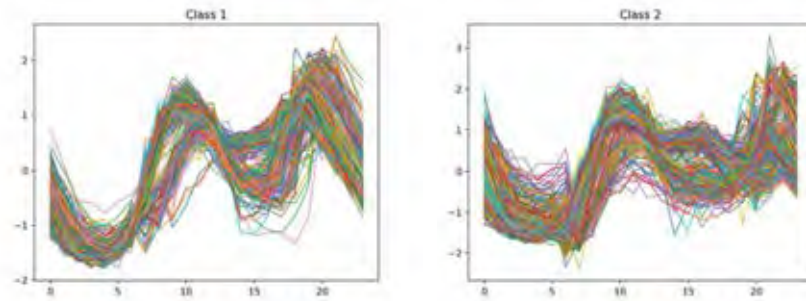
รูปที่ ก.34 กราฟแสดงคลาสของชุดข้อมูล InlineSkate

ตัวอย่างชุดข้อมูล InsectWingbeatSound มีข้อมูลทั้งหมด 11 ในแต่ละคลาส



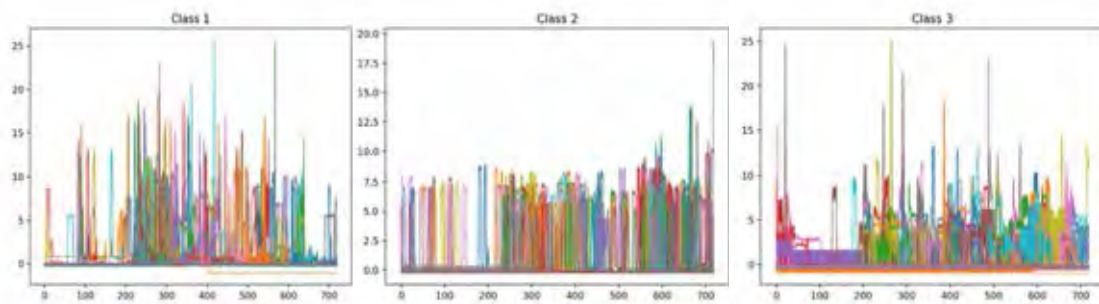
รูปที่ ก.35 กราฟแสดงคลาสของชุดข้อมูล InsectWingbeatSound

ตัวอย่างชุดข้อมูล ItalyPowerDemand มีข้อมูลทั้งหมด 2 คลาส



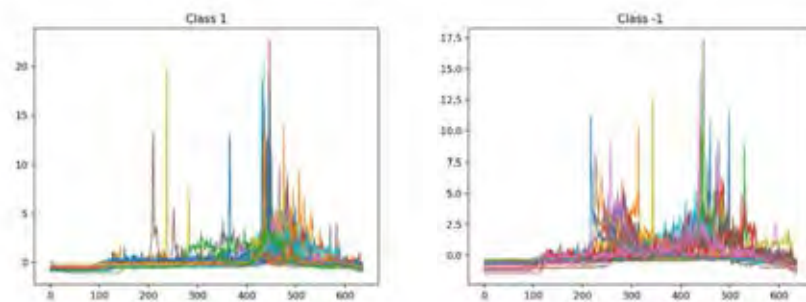
รูปที่ ก.36 กราฟแสดงคลาสของชุดข้อมูล ItalyPowerDemand

ตัวอย่างชุดข้อมูล LargeKitchenAppliances มีข้อมูลทั้งหมด 3 คลาส



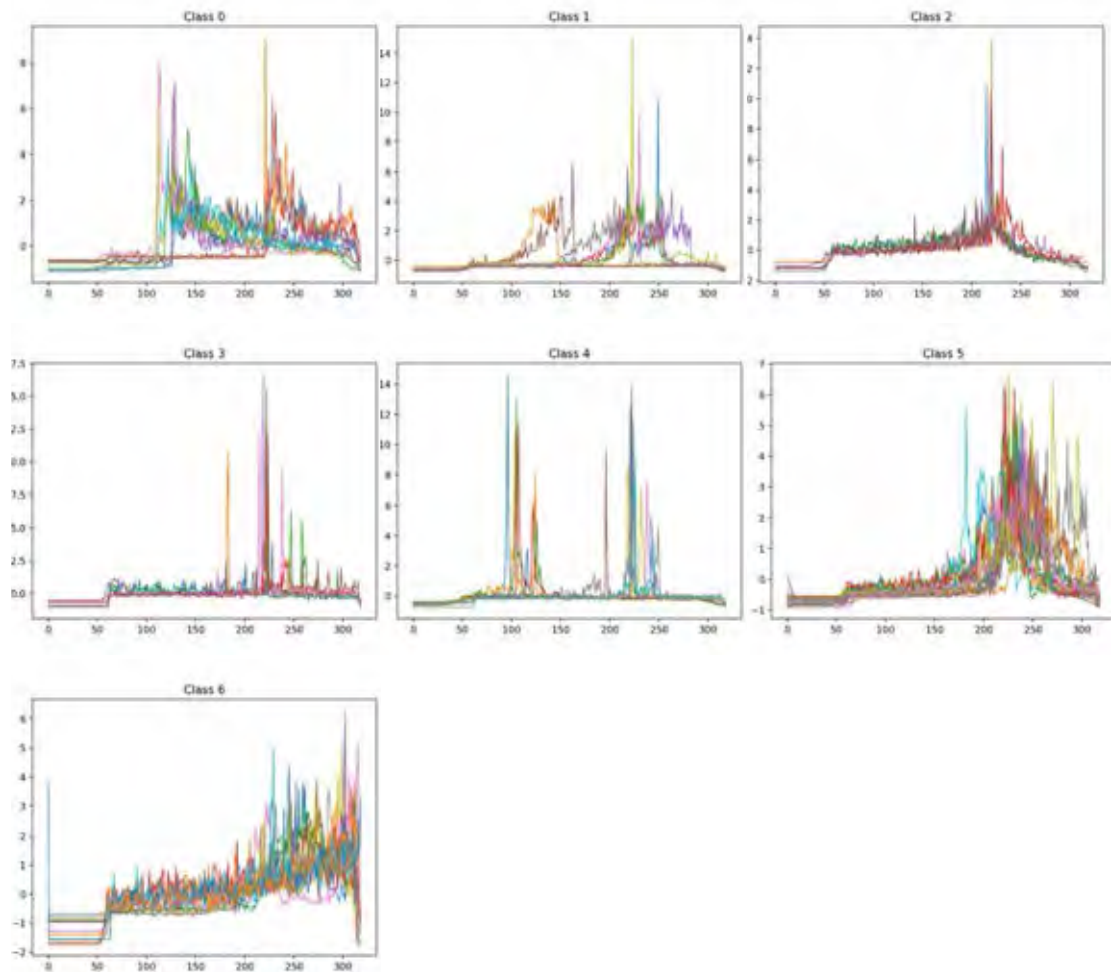
รูปที่ ก.37 กราฟแสดงคลาสของชุดข้อมูล LargeKitchenAppliances

ตัวอย่างชุดข้อมูล Lighting2 มีข้อมูลทั้งหมด 2 คลาส



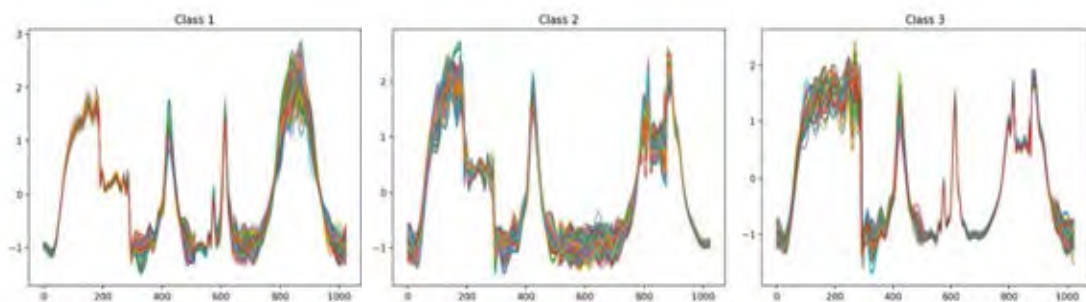
รูปที่ ก.38 กราฟแสดงคลาสของชุดข้อมูล Lighting2

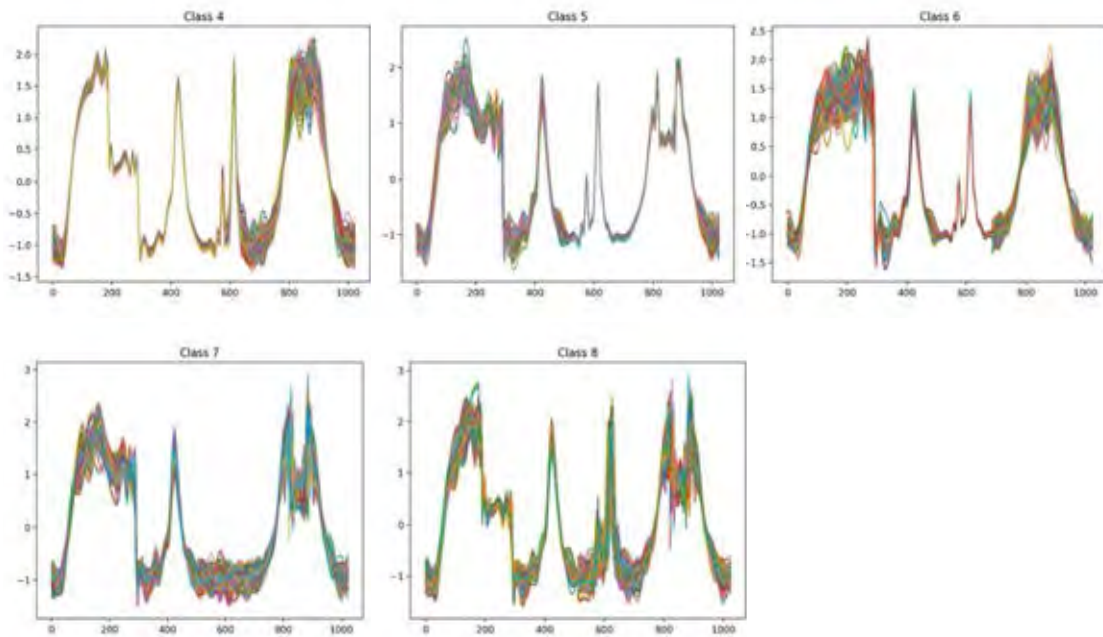
ตัวอย่างชุดข้อมูล Lighting7 มีข้อมูลทั้งหมด 7 คลาส



รูปที่ ก.39 กราฟแสดงคลาสของชุดข้อมูล Lighting7

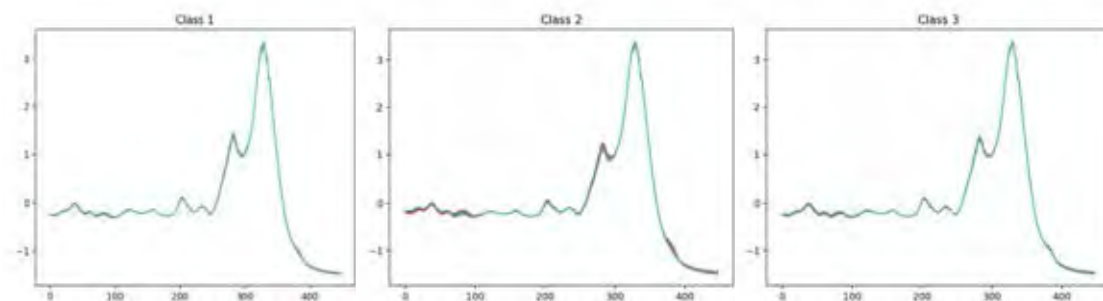
ตัวอย่างชุดข้อมูล MALLAT มีข้อมูลทั้งหมด 8 คลาส





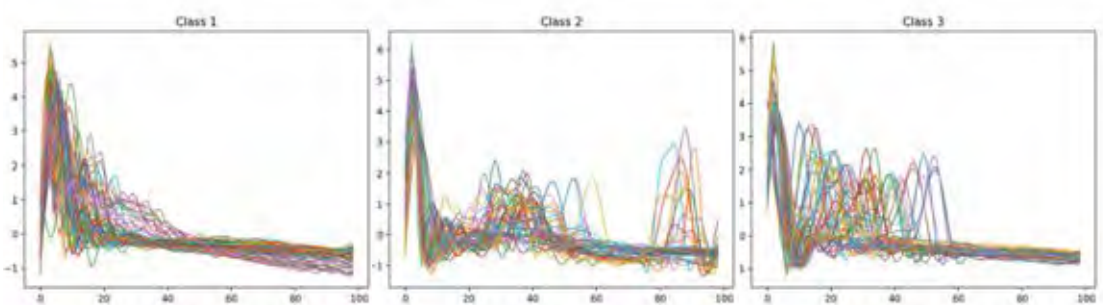
รูปที่ ก.40 กราฟแสดงคลาสของชุดข้อมูล MALLAT

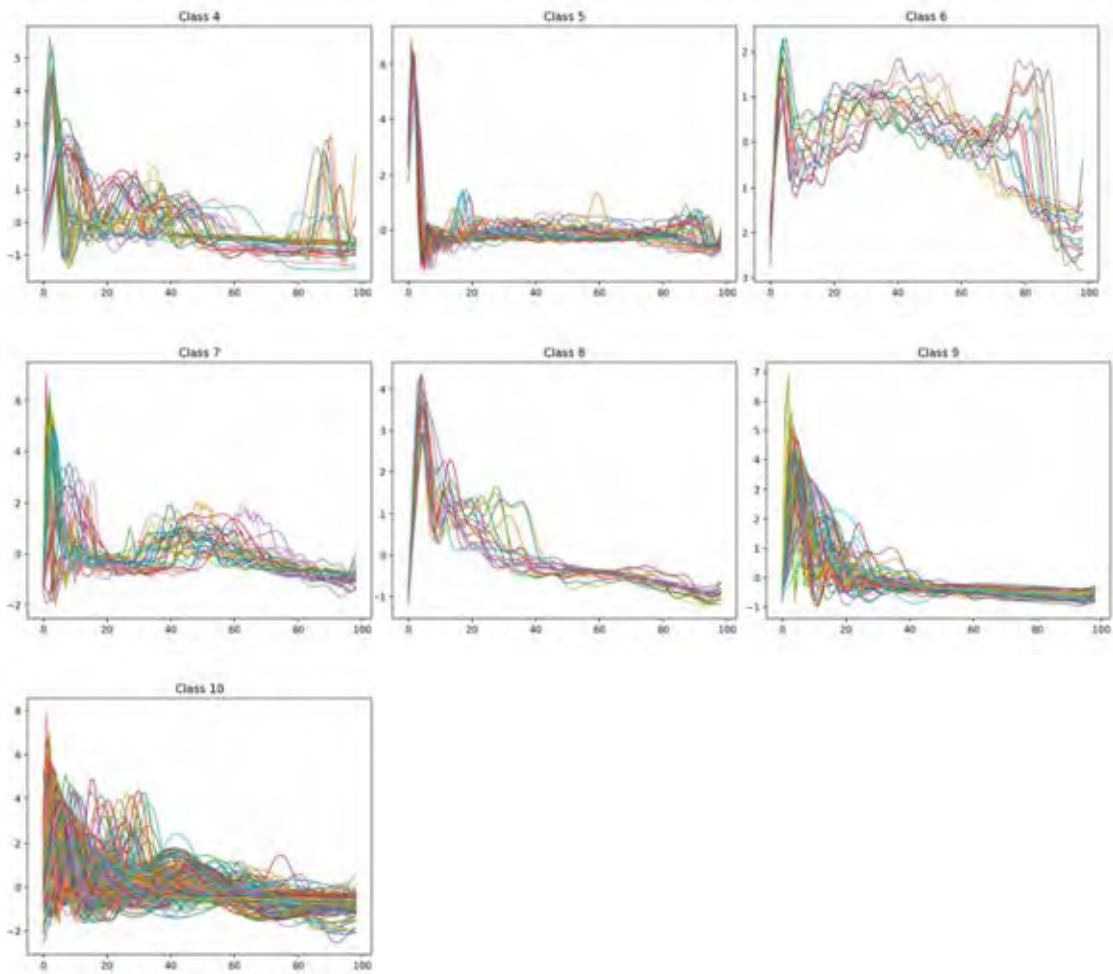
ตัวอย่างชุดข้อมูล Meat มีข้อมูลทั้งหมด 3 คลาส



รูปที่ ก.41 กราฟแสดงคลาสของชุดข้อมูล Meat

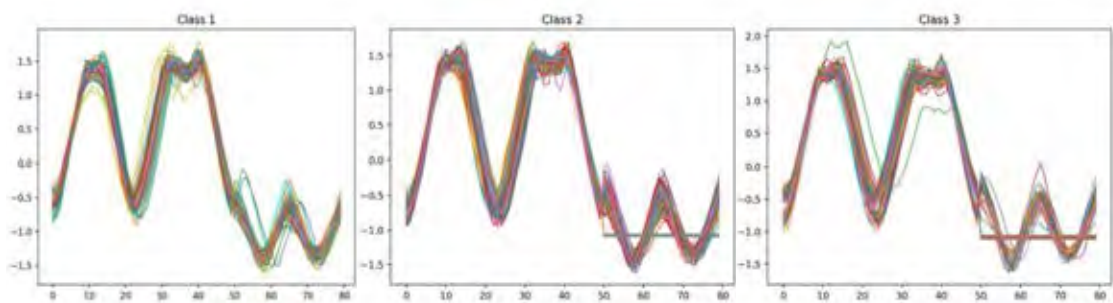
ตัวอย่างชุดข้อมูล MedicalImages มีข้อมูลทั้งหมด 10 ประเภท





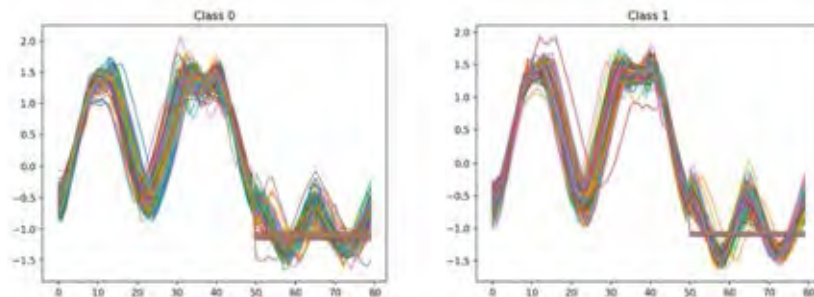
รูปที่ ก.42 กราฟแสดงคลาสของชุดข้อมูล MedicalImages

ตัวอย่างชุดข้อมูล MiddlePhalanxOutlineAgeGroup มีข้อมูลทั้งหมด 3 คลาส



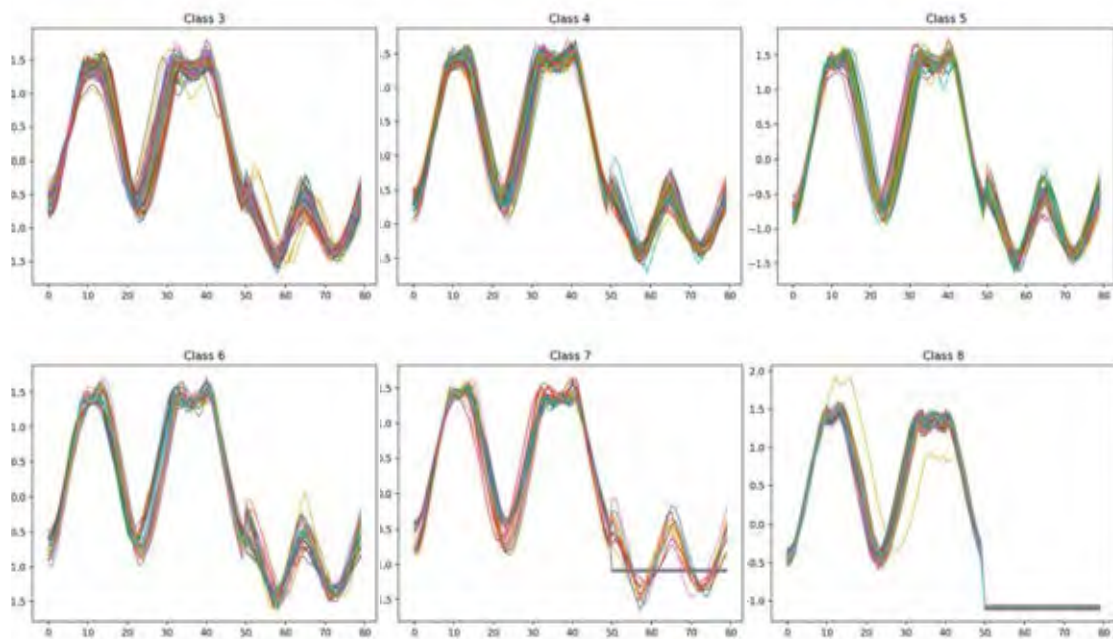
รูปที่ ก.43 กราฟแสดงคลาสของชุดข้อมูล MiddlePhalanxOutlineAgeGroup

ตัวอย่างชุดข้อมูล MiddlePhalanxOutlineCorrect มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.44 กราฟแสดงคลาสของชุดข้อมูล MiddlePhalanxOutlineCorrect

ตัวอย่างชุดข้อมูล MiddlePhalanxTW มีข้อมูลทั้งหมด 6 คลาส

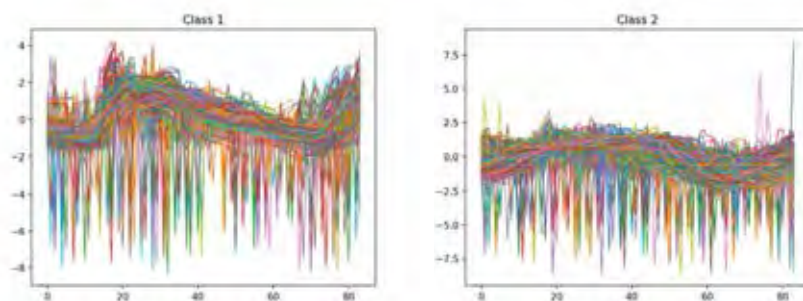


รูปที่ ก.45 กราฟแสดงคลาสของชุดข้อมูล MiddlePhalanxTW



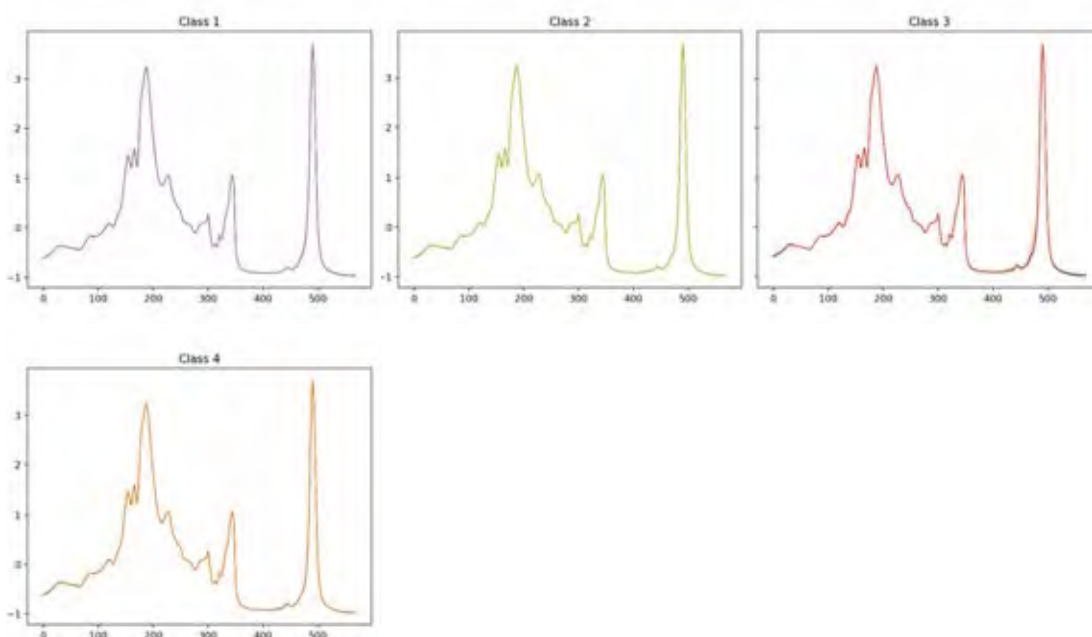
425352952

ตัวอย่างชุดข้อมูล MoteStrain มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.46 กราฟแสดงคลาสของชุดข้อมูล MoteStrain

ตัวอย่างชุดข้อมูล OliveOil มีข้อมูลทั้งหมด 4 คลาส

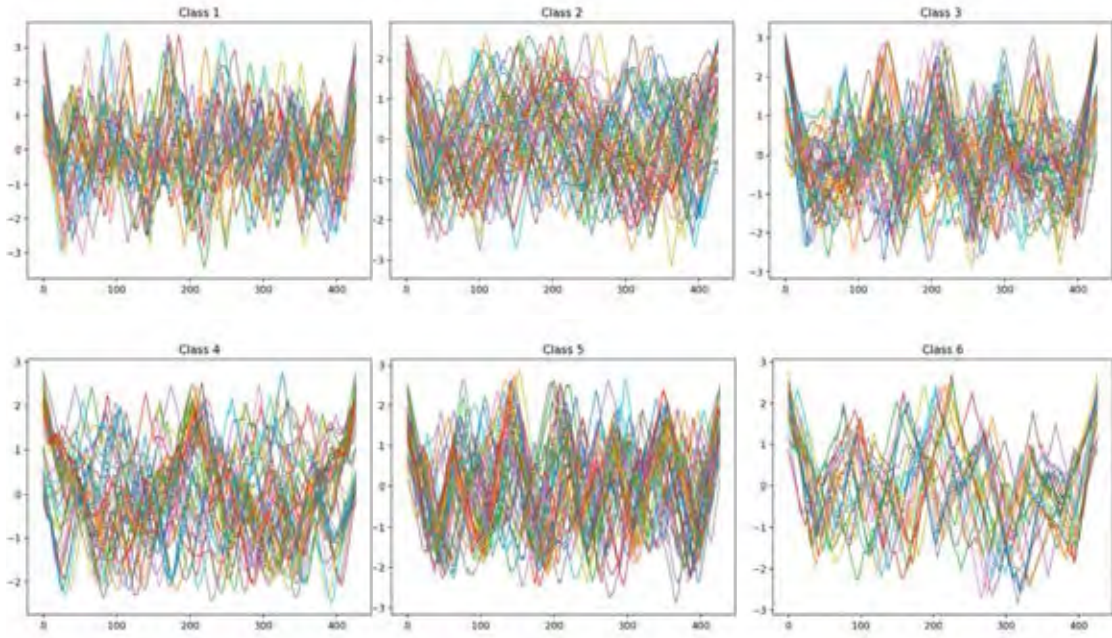


รูปที่ ก.47 กราฟแสดงคลาสของชุดข้อมูล OliveOil



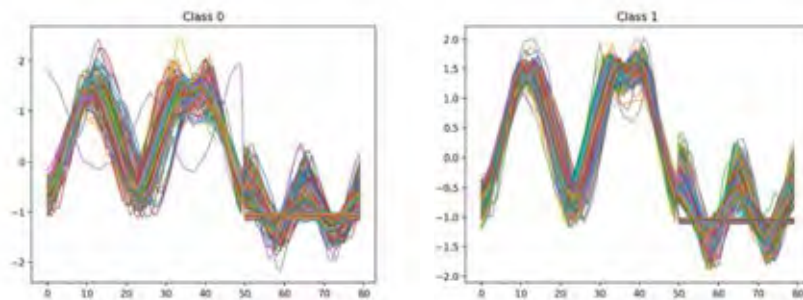
425352952

ตัวอย่างชุดข้อมูล OSU Leaf มีข้อมูลทั้งหมด 6 คลาส



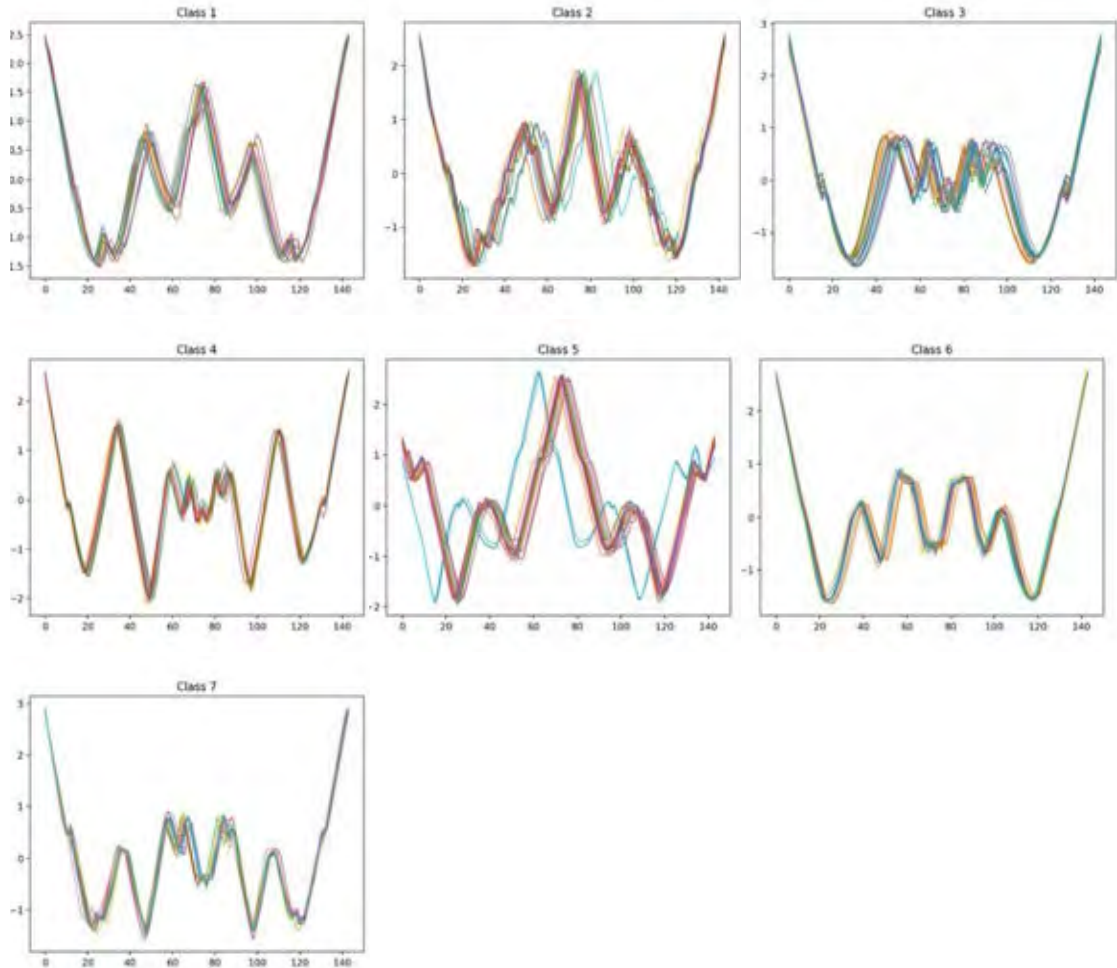
รูปที่ ก.48 กราฟแสดงคลาสของชุดข้อมูล OSU Leaf

ตัวอย่างชุดข้อมูล PhalangesOutlinesCorrect มีข้อมูลทั้งหมด 2 คลาส



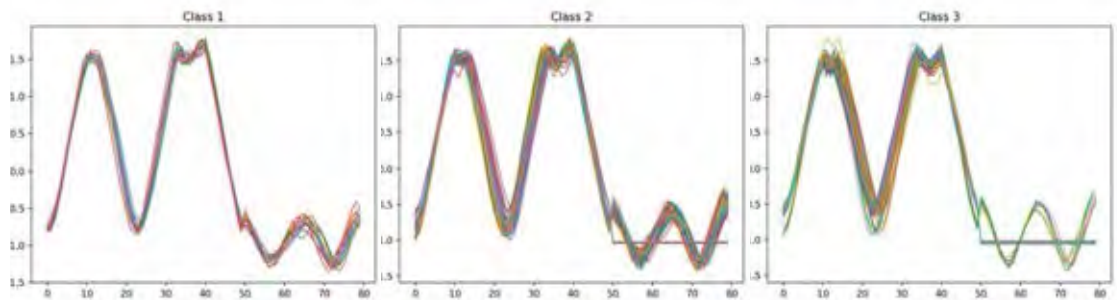
รูปที่ ก.49 กราฟแสดงคลาสของชุดข้อมูล PhalangesOutlinesCorrect

ตัวอย่างชุดข้อมูล Plane มีข้อมูลทั้งหมด 7 คลาส



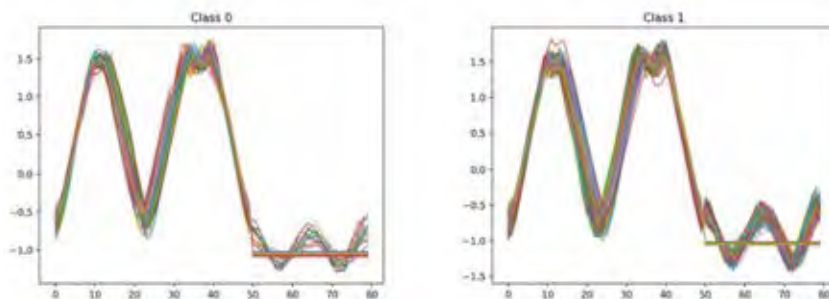
รูปที่ ก.50 กราฟแสดงคลาสของชุดข้อมูล Plane

ตัวอย่างชุดข้อมูล ProximalPhalanxOutlineAgeGroup มีข้อมูลทั้งหมด 3 คลาส



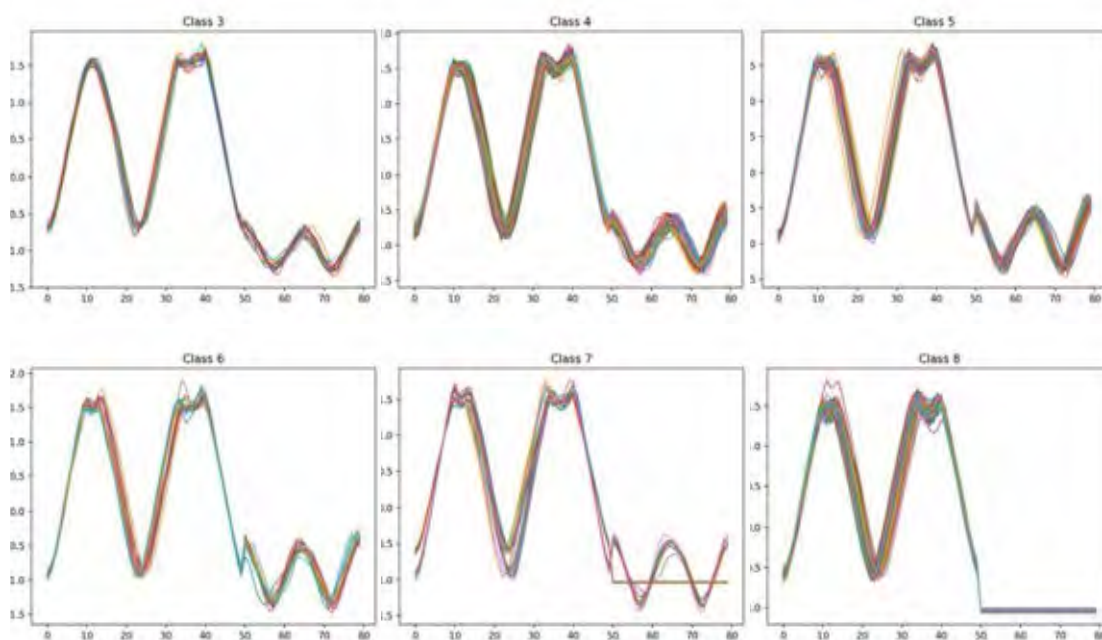
รูปที่ ก.51 กราฟแสดงคลาสของชุดข้อมูล ProximalPhalanxOutlineAgeGroup

ตัวอย่างชุดข้อมูล ProximalPhalanxOutlineCorrect มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.52 กราฟแสดงคลาสของชุดข้อมูล ProximalPhalanxOutlineCorrect

ตัวอย่างชุดข้อมูล ProximalPhalanxTW มีข้อมูลทั้งหมด 6 คลาส

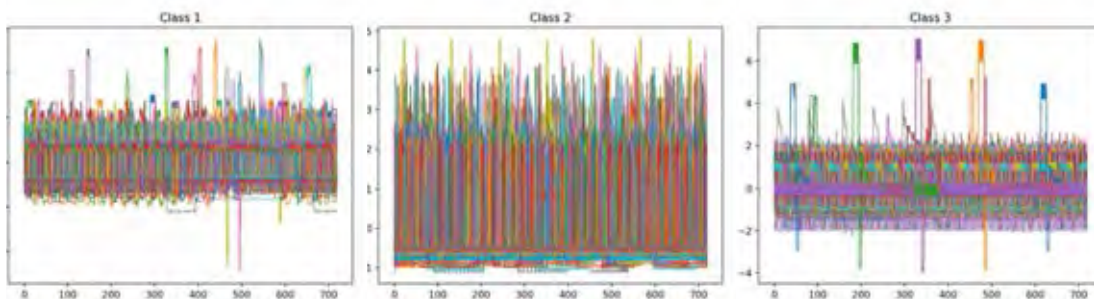


รูปที่ ก.53 กราฟแสดงคลาสของชุดข้อมูล ProximalPhalanxTW



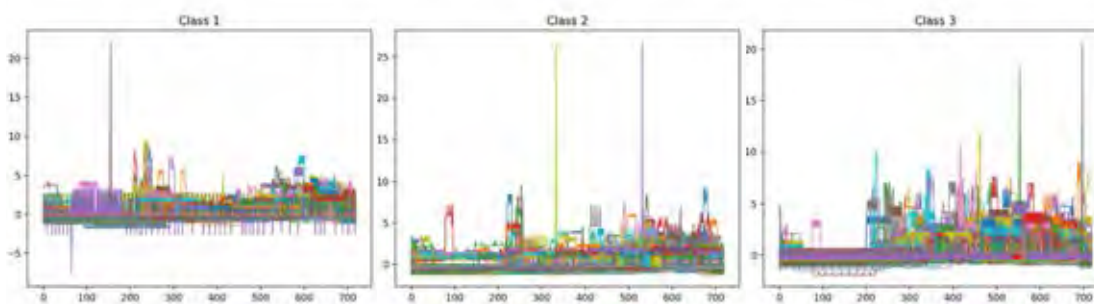
4253529252

ตัวอย่างชุดข้อมูล RefrigerationDevices มีข้อมูลทั้งหมด 3 คลาส



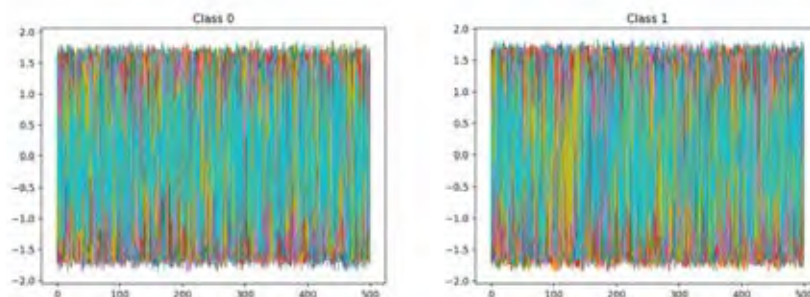
รูปที่ ก.54 กราฟแสดงคลาสของชุดข้อมูล RefrigerationDevices

ตัวอย่างชุดข้อมูล ScreenType มีข้อมูลทั้งหมด 3 คลาส



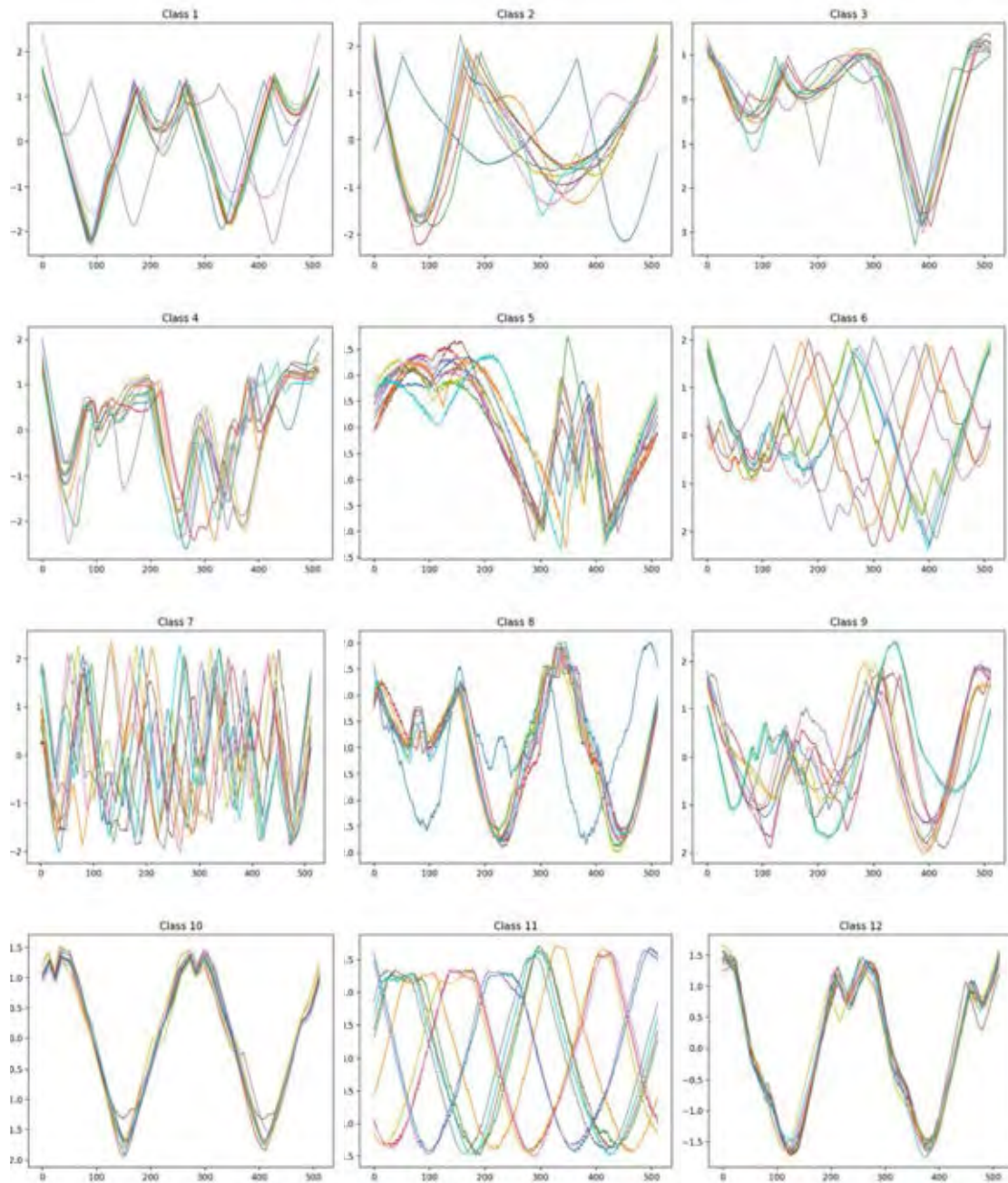
รูปที่ ก.55 กราฟแสดงคลาสของชุดข้อมูล ScreenType

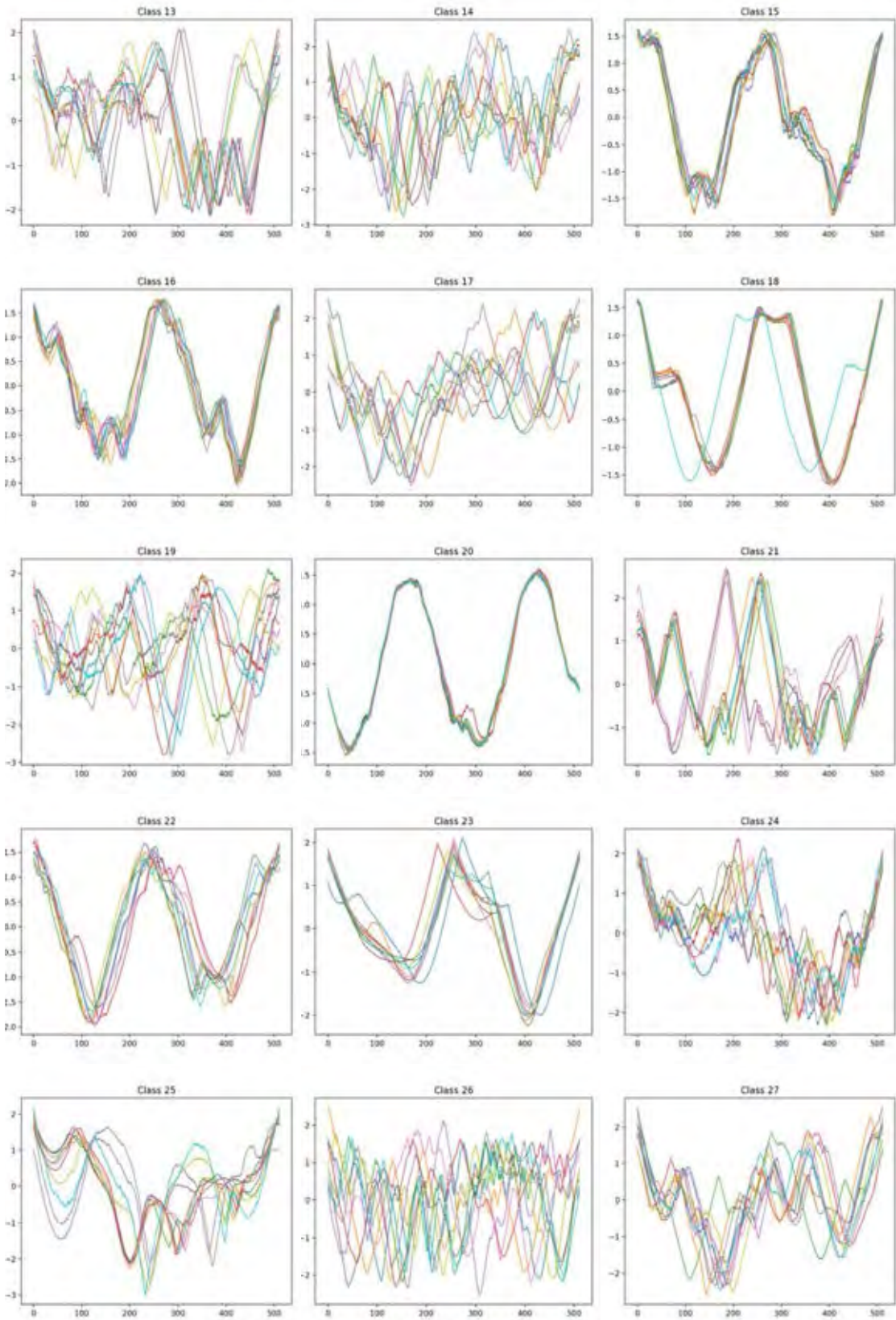
ตัวอย่างชุดข้อมูล ShapeletSim มีข้อมูลทั้งหมด 2 คลาส

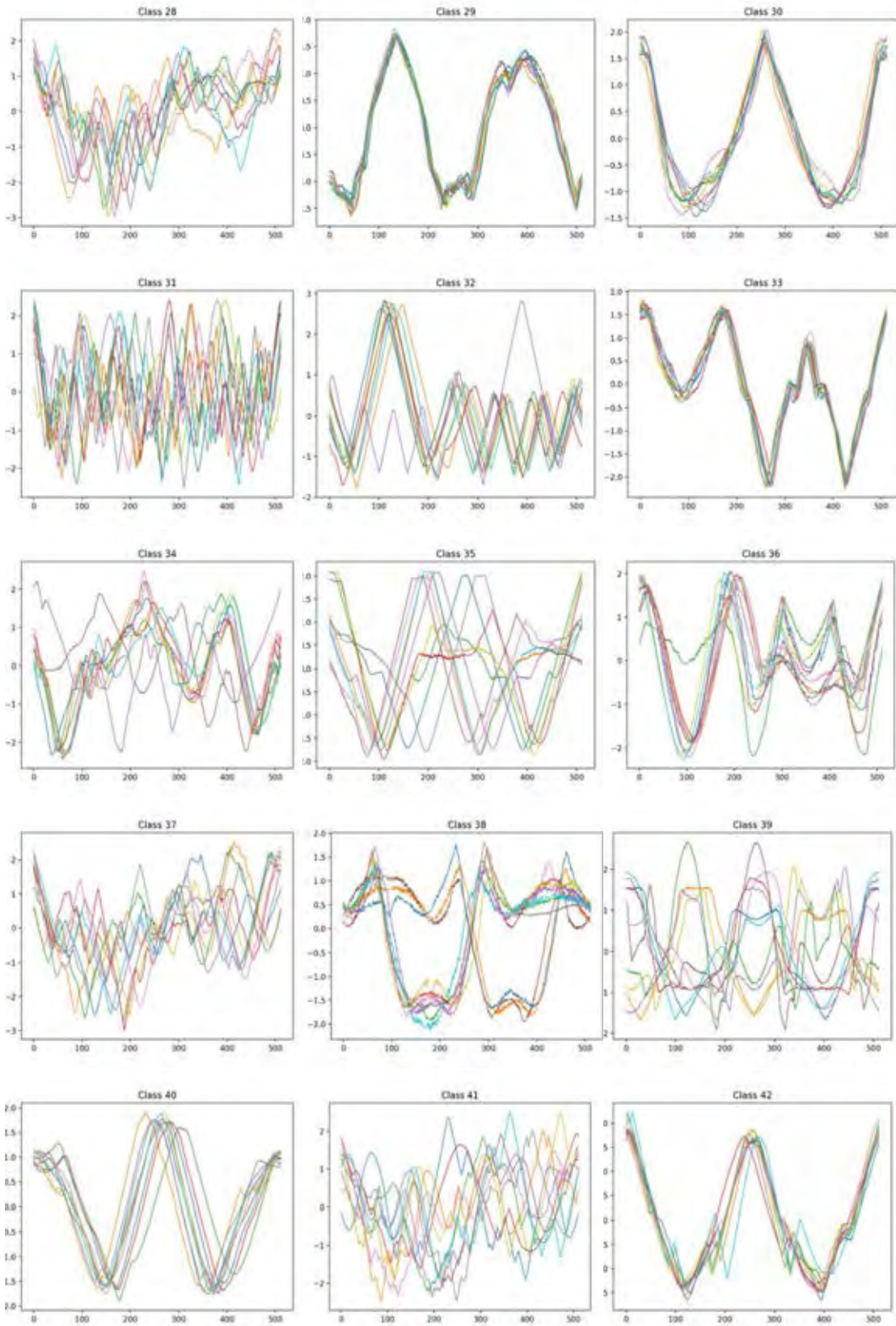


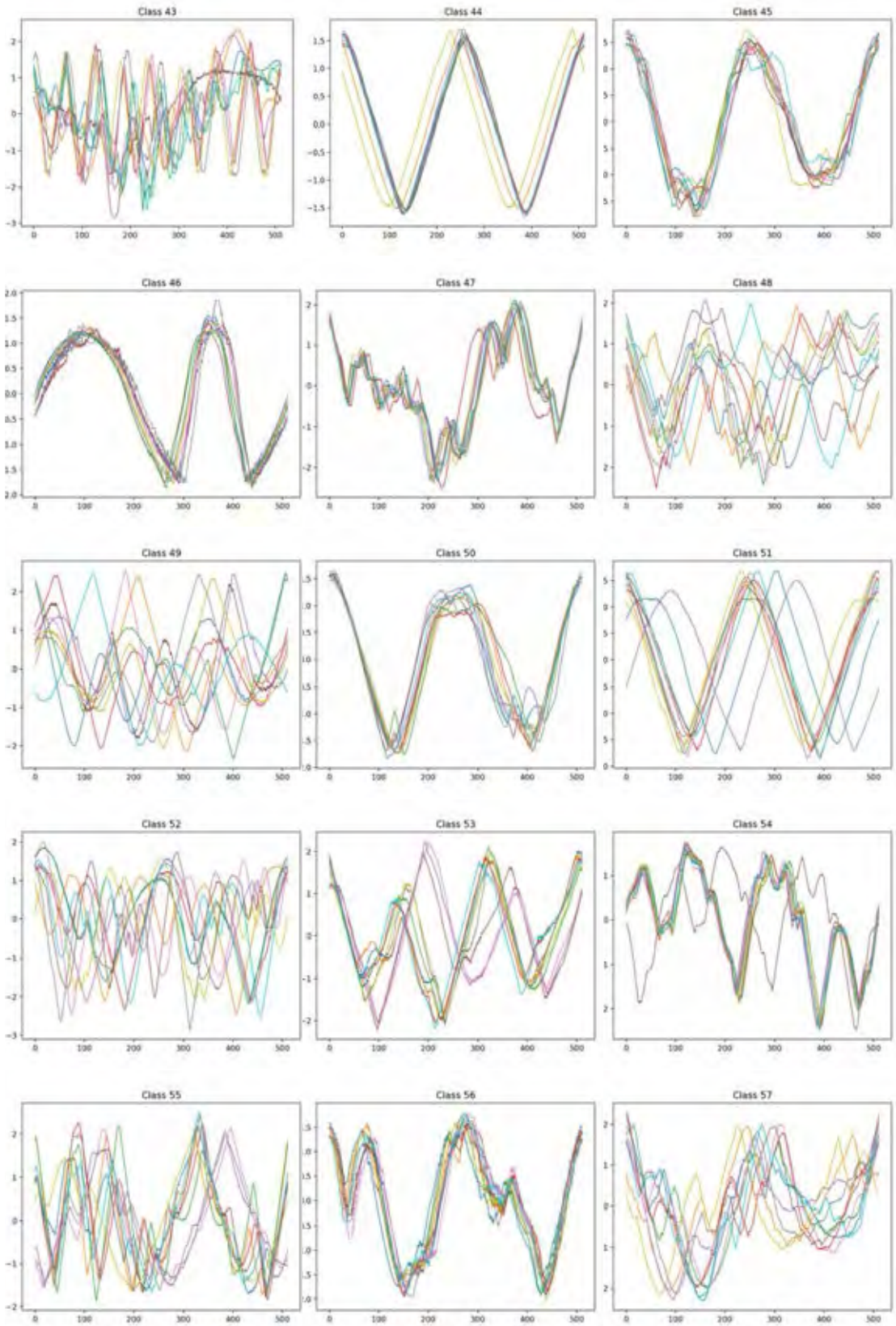
รูปที่ ก.56 กราฟแสดงคลาสของชุดข้อมูล ShapeletSim

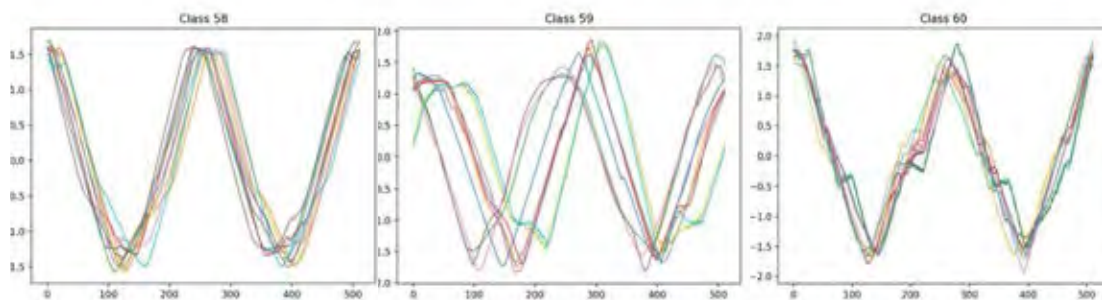
ตัวอย่างชุดข้อมูล ShapesAll มีข้อมูลทั้งหมด 60 คลาส





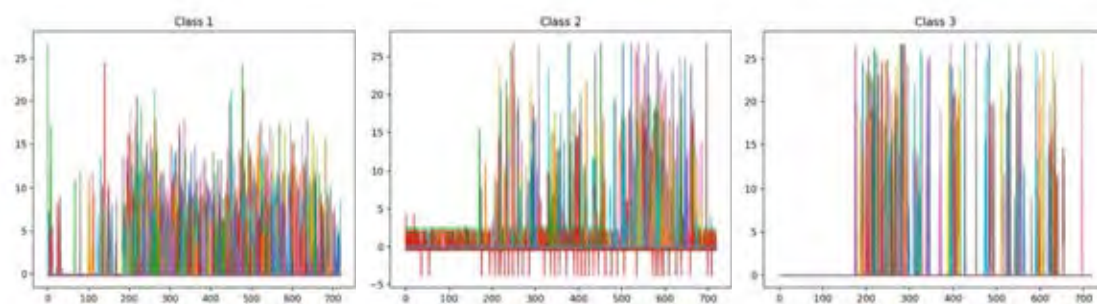






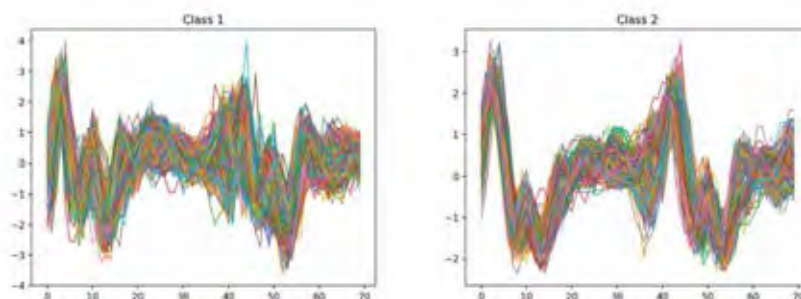
รูปที่ ก.57 กราฟแสดงตัวอย่างคลาสของชุดข้อมูล ShapesAll

ตัวอย่างชุดข้อมูล SmallKitchenAppliances มีข้อมูลทั้งหมด 3 คลาส



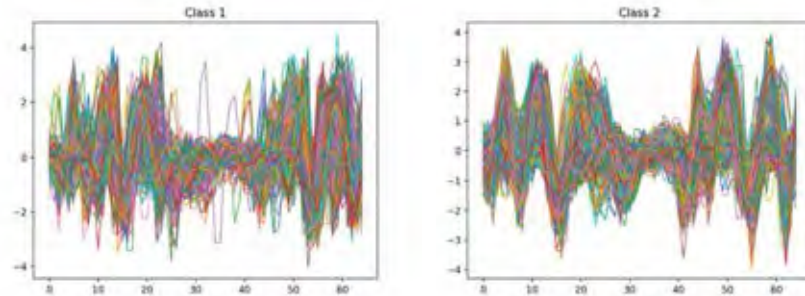
รูปที่ ก.58 กราฟแสดงคลาสของชุดข้อมูล SmallKitchenAppliances

ตัวอย่างชุดข้อมูล SonyAIBORobot Surface มีข้อมูลทั้งหมด 2 คลาส



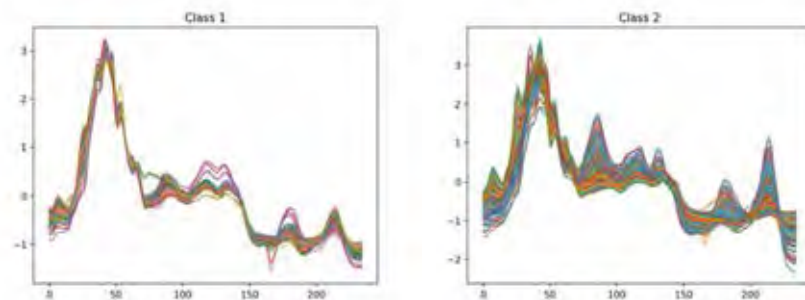
รูปที่ ก.59 กราฟแสดงคลาสของชุดข้อมูล SonyAIBORobot Surface

ตัวอย่างชุดข้อมูล SonyAIBORobot SurfaceII มีข้อมูลทั้งหมด 2 คลาส



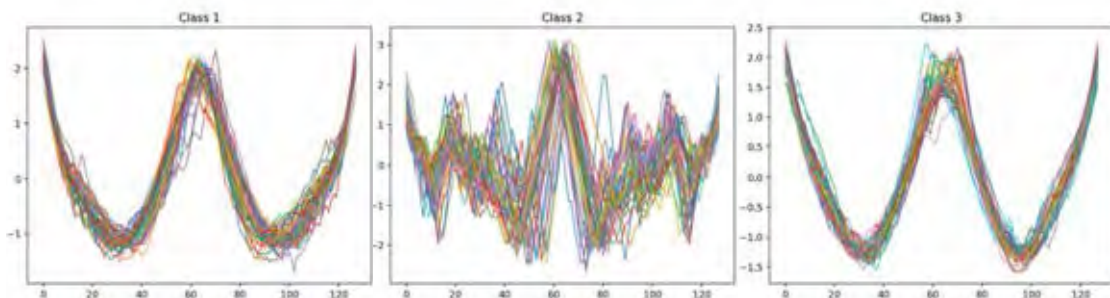
รูปที่ ก.60 กราฟแสดงคลาสของชุดข้อมูล SonyAIBORobot SurfaceII

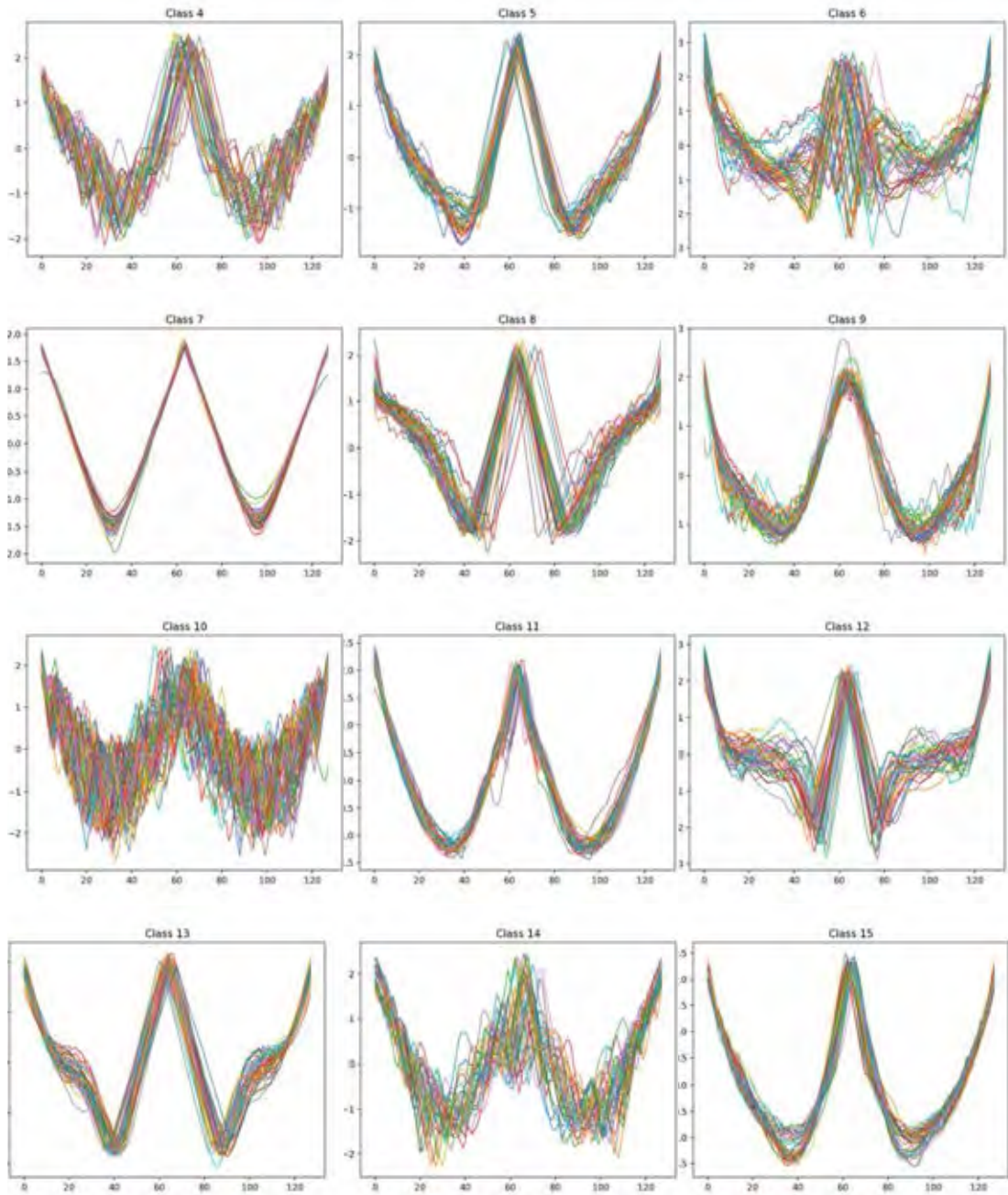
ตัวอย่างชุดข้อมูล Strawberry มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.61 กราฟแสดงคลาสของชุดข้อมูล Strawberry

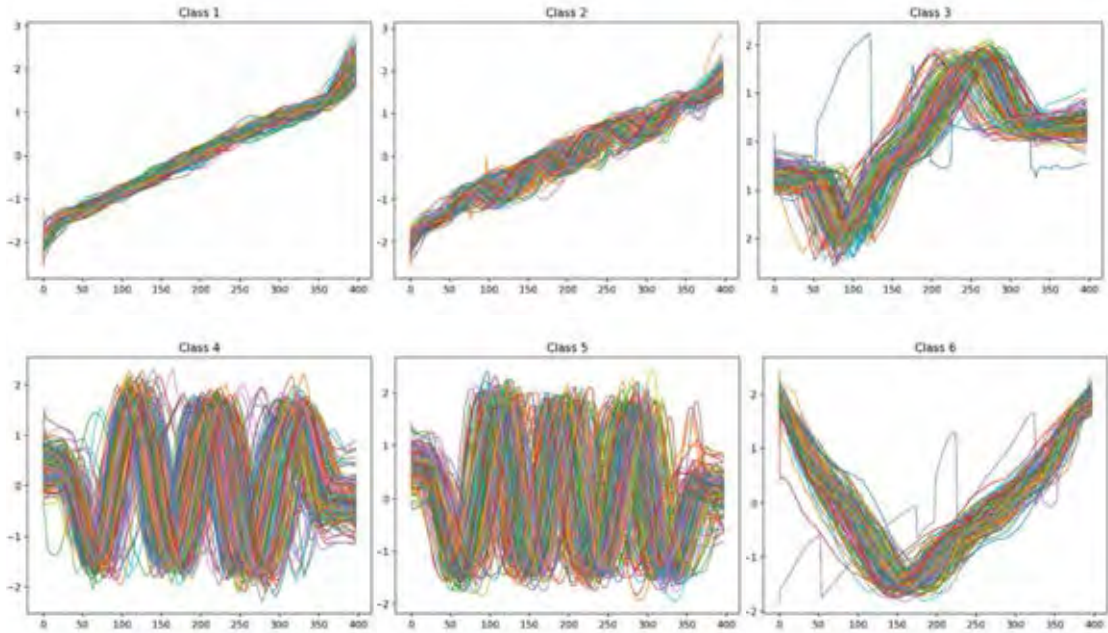
ตัวอย่างชุดข้อมูล Swedish Leaf มีข้อมูลทั้งหมด 15 คลาส





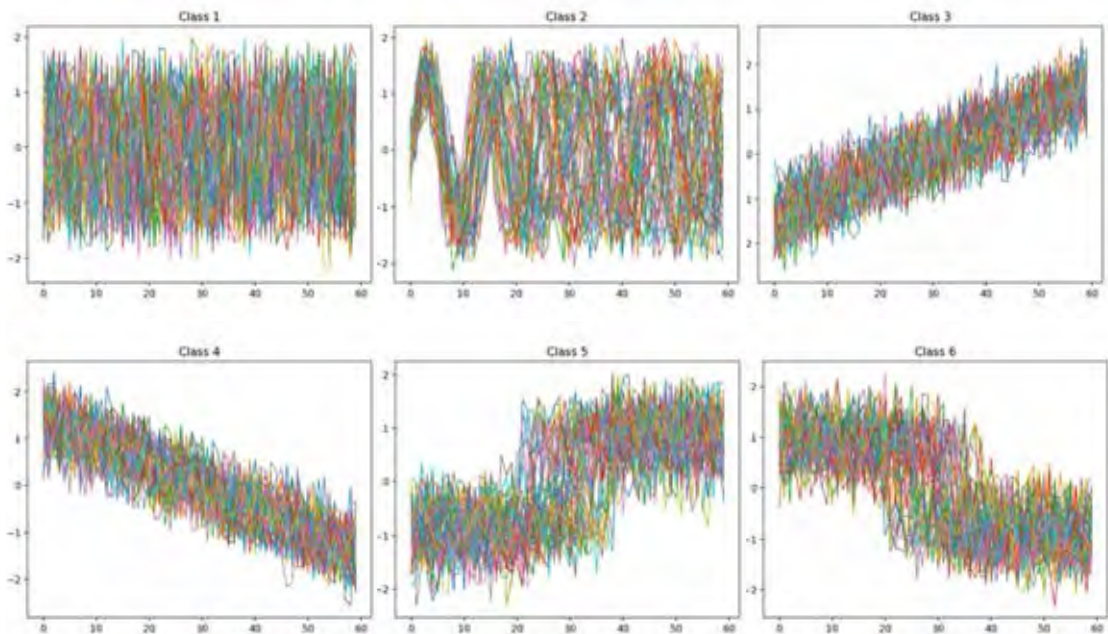
รูปที่ ก.62 กราฟแสดงคลาสของชุดข้อมูล Swedish Leaf

ตัวอย่างชุดข้อมูล Symbols มีข้อมูลทั้งหมด 6 คลาส



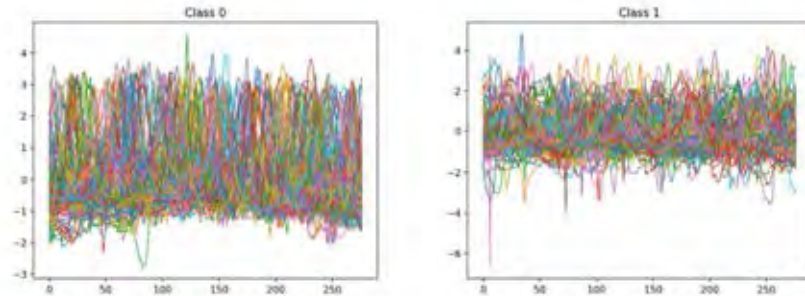
รูปที่ ก.63 กราฟแสดงคลาสของชุดข้อมูล Symbols

ตัวอย่างชุดข้อมูล Synthetic Control มีข้อมูลทั้งหมด 6 คลาส



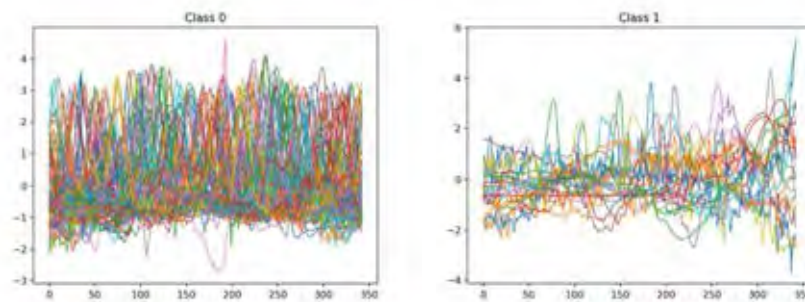
รูปที่ ก.64 กราฟแสดงคลาสของชุดข้อมูล Synthetic Control

ตัวอย่างชุดข้อมูล ToeSegmentation1 มีข้อมูลทั้งหมด 2 คลาส



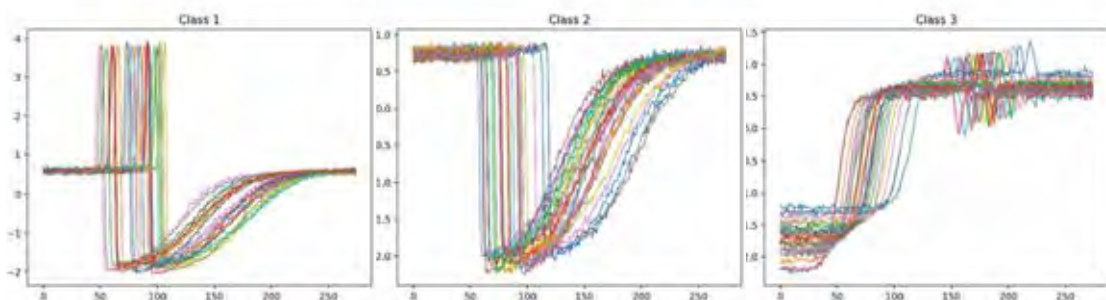
รูปที่ ก.65 กราฟแสดงคลาสของชุดข้อมูล ToeSegmentation1

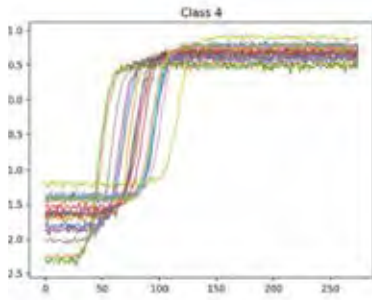
ตัวอย่างชุดข้อมูล ToeSegmentation 2 มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.66 กราฟแสดงคลาสของชุดข้อมูล ToeSegmentation2

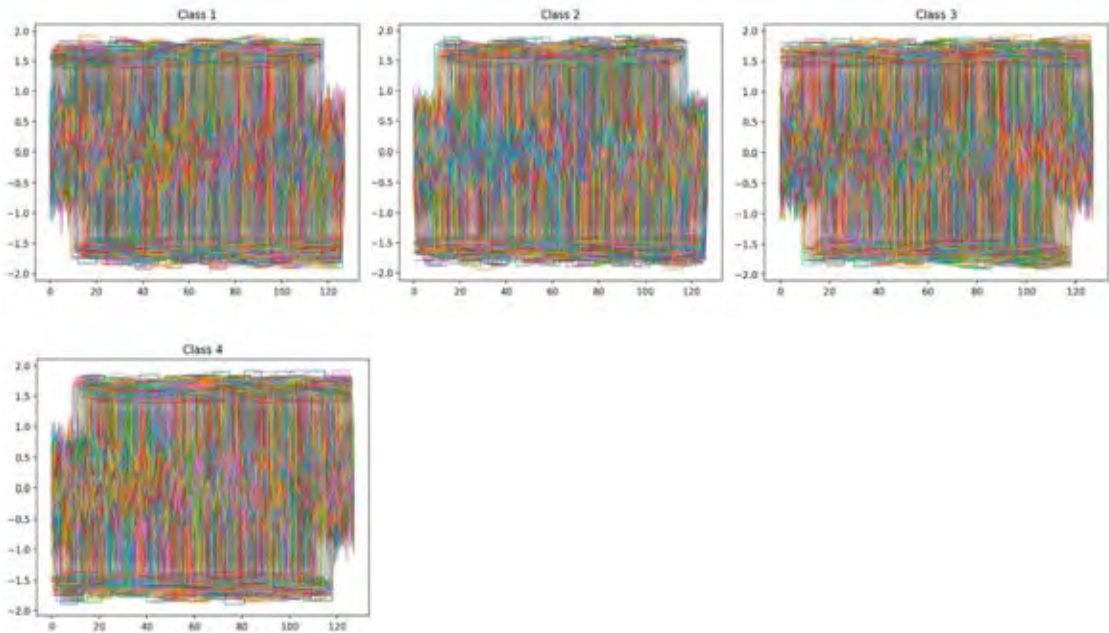
ตัวอย่างชุดข้อมูล Trace มีข้อมูลทั้งหมด 4 ในแต่ละคลาส





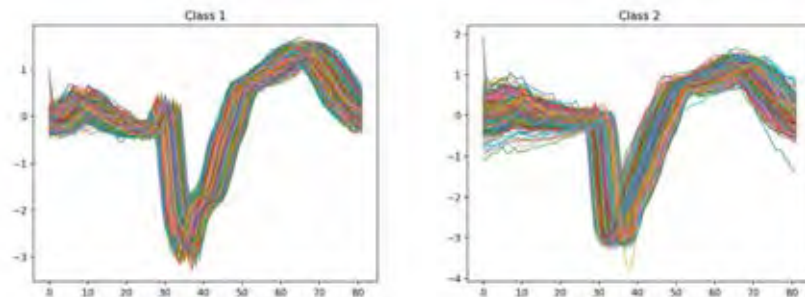
รูปที่ ก.67 กราฟแสดงคลาสของชุดข้อมูล Trace

ตัวอย่างชุดข้อมูล Two Patterns มีข้อมูลทั้งหมด 4 คลาส



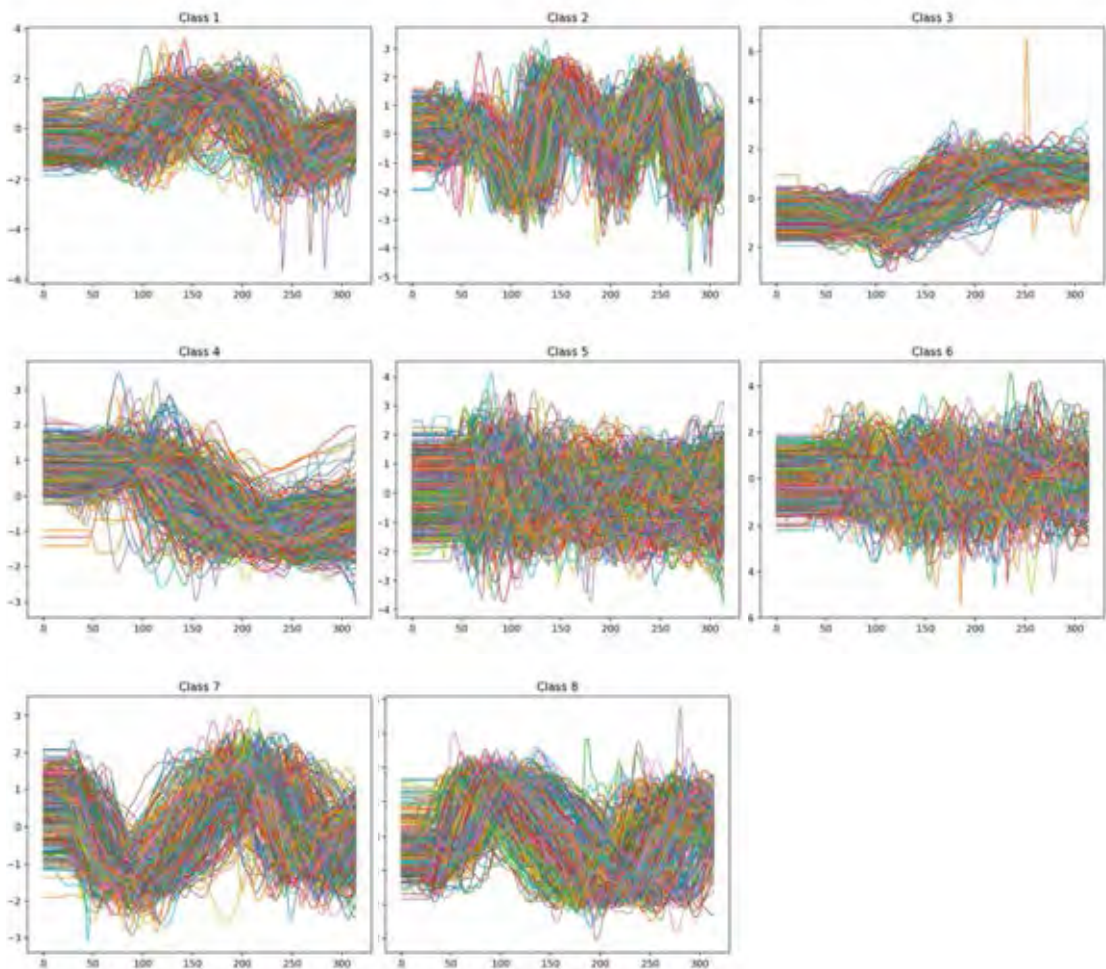
รูปที่ ก.68 กราฟแสดงคลาสของชุดข้อมูล Two Patterns

ตัวอย่างชุดข้อมูล TwoLeadECG มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.69 กราฟแสดงคลาสของชุดข้อมูล TwoLeadECG

ตัวอย่างชุดข้อมูล uWaveGestureLibrary_X มีข้อมูลทั้งหมด 8 คลาส

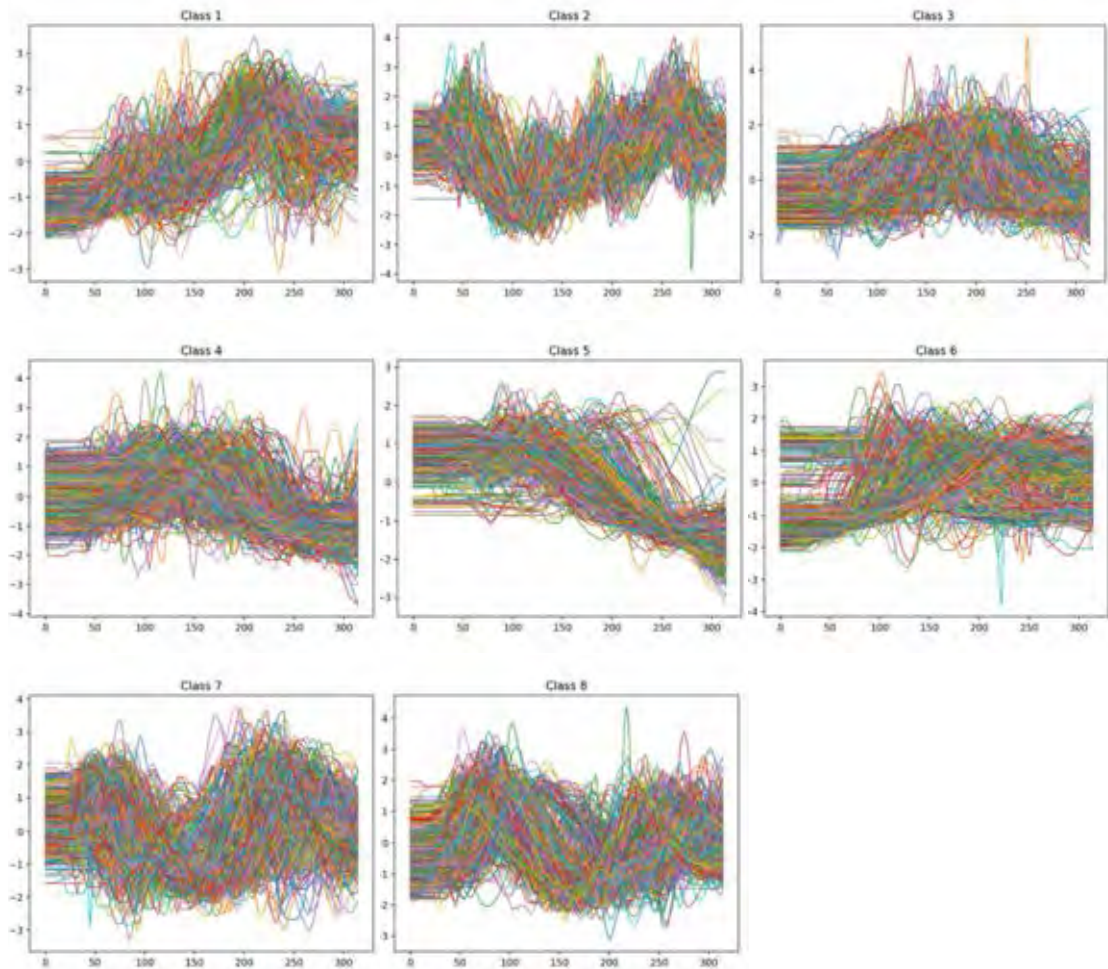


รูปที่ ก.70 กราฟแสดงคลาสของชุดข้อมูล uWaveGestureLibrary_X



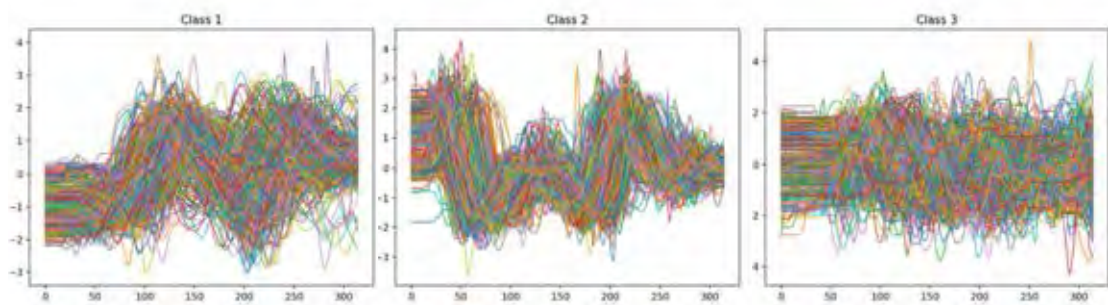
425352952

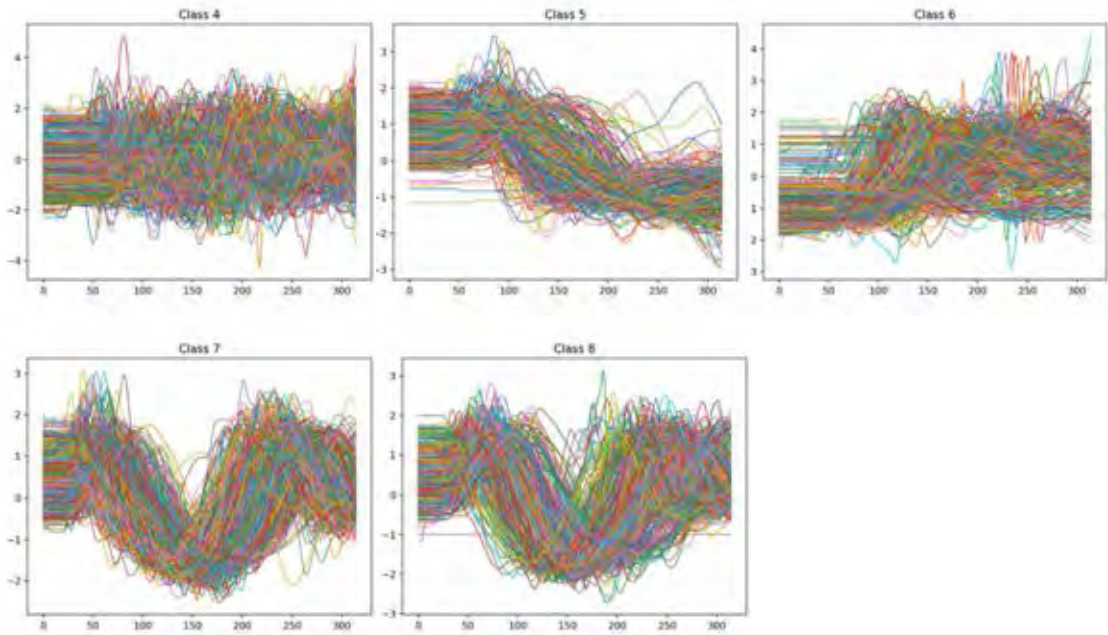
ตัวอย่างชุดข้อมูล uWaveGestureLibrary_Y มีข้อมูลทั้งหมด 8 ในแต่ละคลาส



รูปที่ ก.71 กราฟแสดงคลาสของชุดข้อมูล uWaveGestureLibrary_Y

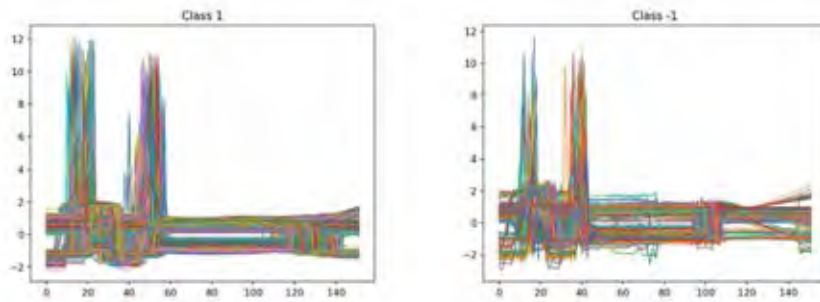
ตัวอย่างชุดข้อมูล uWaveGestureLibrary_Z มีข้อมูลทั้งหมด 8 คลาส





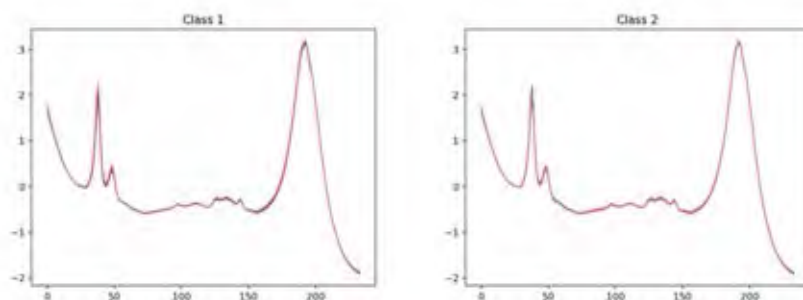
รูปที่ ก.72 กราฟแสดงคลาสของชุดข้อมูล uWaveGestureLibrary_Z

ตัวอย่างชุดข้อมูล Wafer มีข้อมูลทั้งหมด 2 คลาส



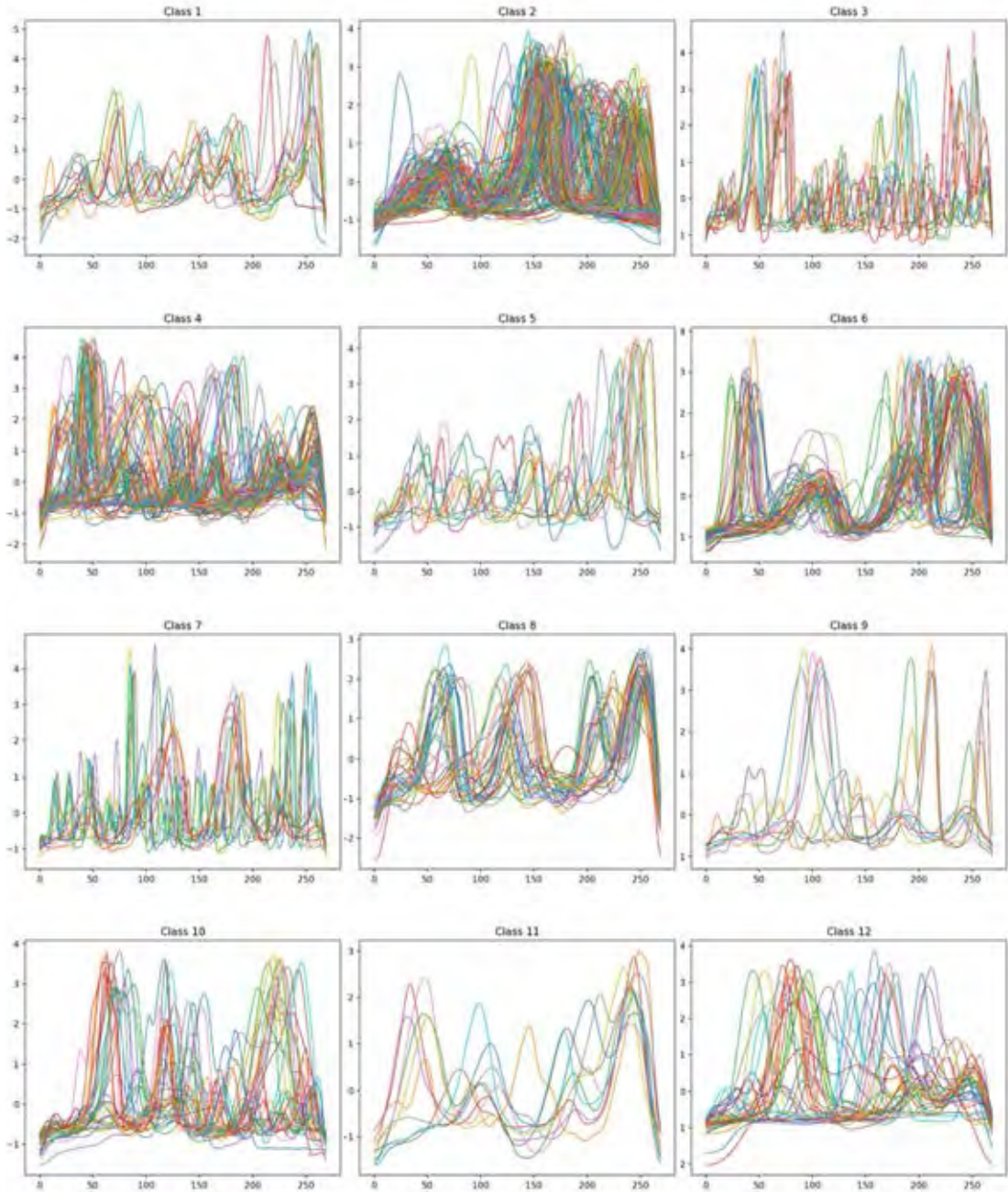
รูปที่ ก.73 กราฟแสดงคลาสของชุดข้อมูล Wafer

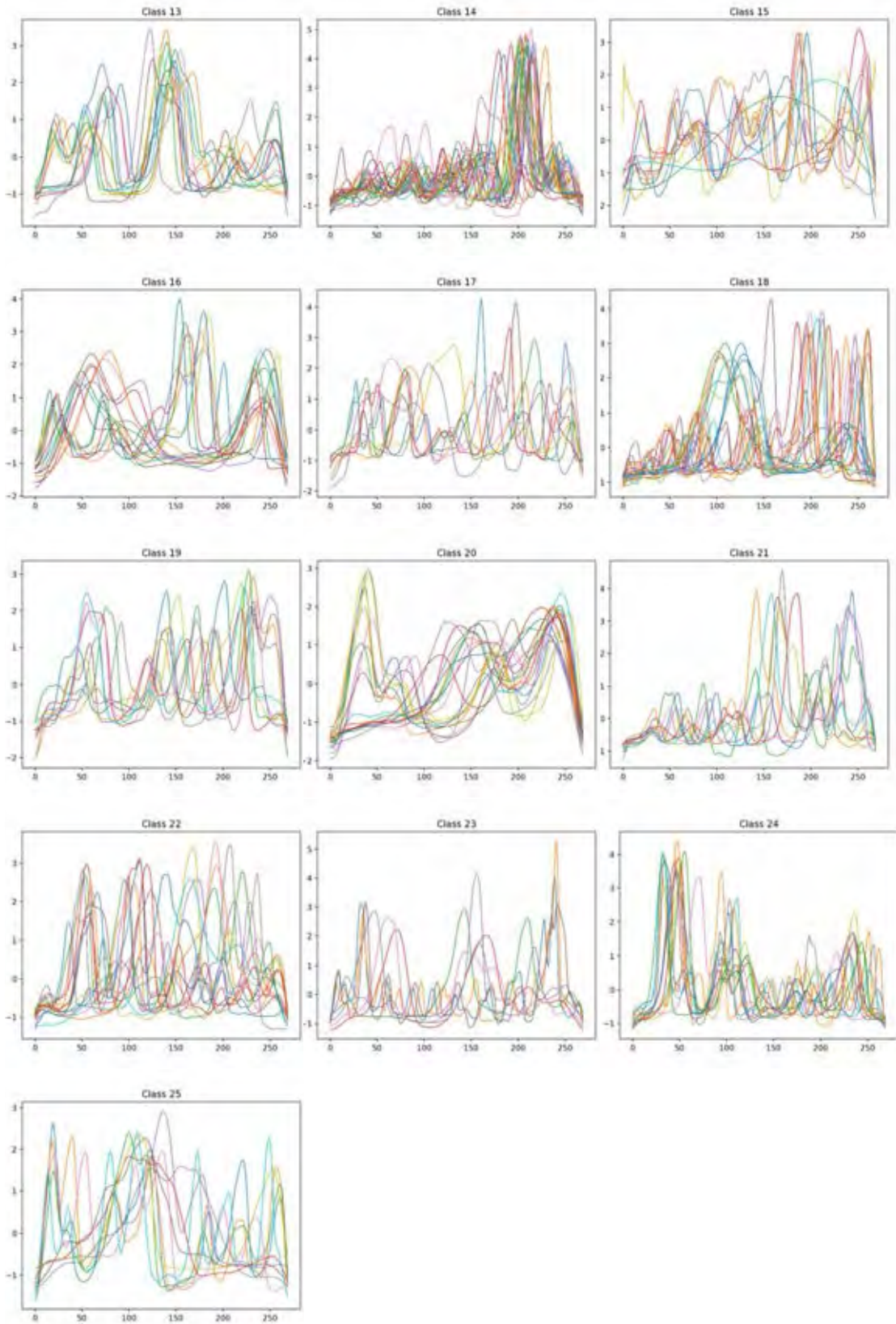
ตัวอย่างชุดข้อมูล Wine มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.74 กราฟแสดงคลาสของชุดข้อมูล Wine

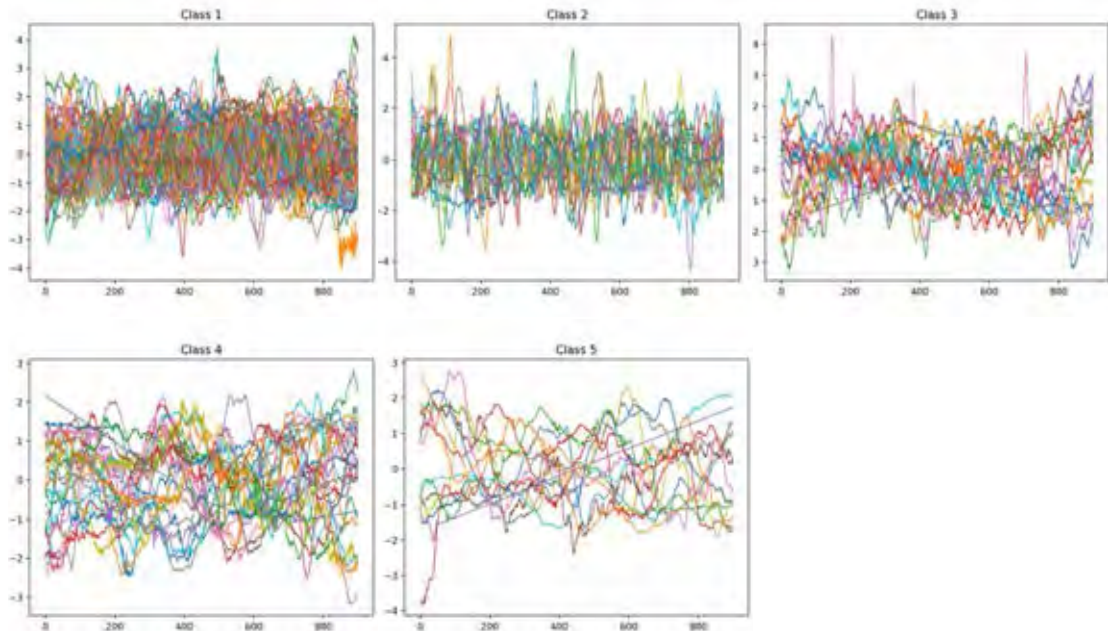
ตัวอย่างชุดข้อมูล WordSynonyms มีข้อมูลทั้งหมด 25 คลาส





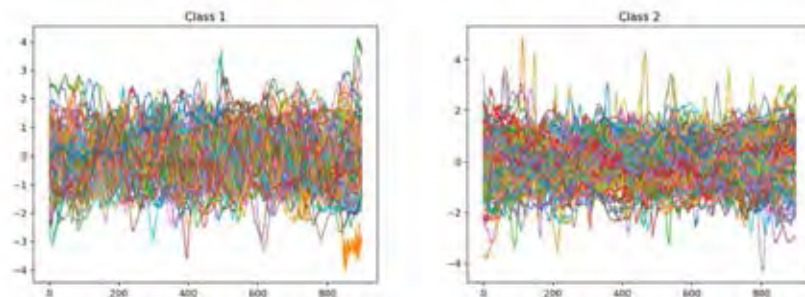
รูปที่ ก.75 กราฟแสดงคลาสของชุดข้อมูล WordSynonyms

ตัวอย่างชุดข้อมูล Worms มีข้อมูลทั้งหมด 5 คลาส



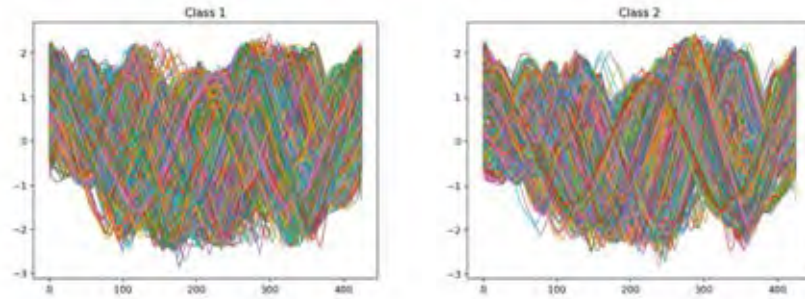
รูปที่ ก.76 กราฟแสดงคลาสของชุดข้อมูล Worms

ตัวอย่างชุดข้อมูล WormsTwoClass มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.77 กราฟแสดงคลาสของชุดข้อมูล WormsTwoClass

ตัวอย่างชุดข้อมูล Yoga มีข้อมูลทั้งหมด 2 คลาส



รูปที่ ก.78 กราฟแสดงคลาสของชุดข้อมูล Yoga



425352952

CU Thesais 6070323921 thesais / recv: 06082562 14:42:31 / seq: 83

ประวัติผู้เขียน

ชื่อ-สกุล	ศัลัมพา อานุภาพแสนยากร
วัน เดือน ปี เกิด	14 มีนาคม 2538
สถานที่เกิด	648/35 ถ.โนนม่วง ต. ในเมือง อ.เมือง จ.ชัยภูมิ 36000
วุฒิการศึกษา	สำเร็จการศึกษาระดับมัธยมจากโรงเรียนชัยภูมิภักดีชุมพล จากนั้นเข้าศึกษาต่อที่คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ในปีการศึกษา 2560 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต ภาควิชาวิทยาการคอมพิวเตอร์ ในปีการศึกษา และเข้าศึกษาในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมศาสตร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ในปี 2560
ที่อยู่ปัจจุบัน	บริษัท ปาป้ามาม้า จำกัด 12-12 ถ.สีพระยา แขวงมหาพฤฒาราม เขตบางรัก จ.กรุงเทพมหานคร 10500
ผลงานตีพิมพ์	S. Arnupapsanyakorn and C. A. Ratanamahatana (2019). An enhanced time series classification using linear regression based shape descriptor. in Second International Conference on Data Science and Information Technology (DSIT 2019). at Seoul, Korea during July 19-21, 2019.