การจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจเชิงวากยสัมพันธ์

นายสาธิต ตั้งติดธรรม

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

SYNTACTIC ELECTROCARDIOGRAPHY CLASSIFICATION

Mr.Satit Tangtidtham

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Computer Science
Department of Computer Engineering
Chulalongkorn University
Academic Year 2008
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจเชิงวากยสัมพันธ์
โดย	นายสาธิต ตั้งติดธรรม
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	รองศาสตราจารย์ นงลักษณ์ โควาวิสารัช

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

.... คณบดีคณะวิศวกรรมศาสตร์
(รองศาสตราจารย์ ดร.บุญสม เลิศหิรัญวงศ์)

คณะกรรมการสอบวิทยานิพนธ์

.... ประธานกรรมการ
(ศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา)

.... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์)

.... อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม
(รองศาสตราจารย์ นงลักษณ์ โควาวิสารัช)

.... กรรมการ
(ผู้ช่วยศาสตราจารย์ นายแพทย์ ทายาท ดีสุดจิต)

.... กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.อานนท์ รุ่งสว่าง)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สาธิต ตั้งติดธรรม : การจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจเชิงวากยสัมพันธ์ (SYNTACTIC ELECTROCARDIOGRAPHY CLASSIFICATION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร.อรรถสิทธิ์ สุรฤกษ์, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม: รศ. นงลักษณ์ โควาวิสารัช, 36 หน้า

การตรวจวัดภาพคลื่นไฟฟ้าหัวใจซึ่งเกิดจากการไหลเวียนของกระแสไฟฟ้าภายในหัวใจเพื่อช่วยในการวินิจฉัยและระบุปัญหาของหัวใจที่อาจเกิดขึ้นนั้น กำลังเป็นปัญหาที่สำคัญและน่าเบื่อ เนื่องจากระยะเวลาของภาพคลื่นไฟฟ้าหัวใจที่ถูกบันทึกและต้องนำมาวินิจฉัย มีความยาว 8 ถึง 24 ชั่วโมง ซึ่งการวินิจฉัยแบบนี้ต้องใช้เวลาอย่างมากในการตรวจสอบ การทำให้ภาพคลื่นไฟฟ้าหัวใจรวมกันอยู่เป็นกลุ่มพวกเดียวกันเป็นสิ่งที่จะช่วยในการแก้ปัญหานี้อย่างมาก ในงานวิจัยนี้เราสนใจที่จะจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจโดยใช้เทคนิคในการหารูปแบบของภาพคลื่นไฟฟ้าหัวใจ และแสดงออกมาเป็นรูปแบบของออโตมาตาแบบจำกัด โดยมีการปรับปรุงแก้ไขการสร้างออโตมาตาในแบบเดิมของการสร้างออโตมาตาตรวจสอบได้เค ให้มีความยืดหยุ่นมากขึ้นกว่ารูปแบบเดิม โดยผลการทดลองที่ได้แสดงให้เห็นถึงความสามารถที่เปลี่ยนไปอย่างเห็นได้ชัดด้วยเทคนิคนี้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมคอมพิวเตอร์ ลายมือชื่อนิสิต ....
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออ.ที่ปรึกษาวิทยานิพนธ์ ....
ปีการศึกษา 2551

4871454821: MAJOR COMPUTER SCIENCE

KEYWORDS: SYNTACTIC ECG, EKG / ELECTROCARDIOGRAPHY CLASSIFICATION

SATIT TANGTIDTHAM: SYNTACTIC ELECTROCARDIOGRAPHY CLASSIFICATION. ADVISOR: ASST. PROF ATHASIT SURARERKS Ph.D., CO-ADVISOR: ASSC. PROF. NONGLUK COVAVISARUCH, 36 pp.

Electrocardiography (EKG) measures heart's electrical activity to help physicians evaluate its function and identify any problems that might exist. EKG investigation is an important and tedious problem since EKG graph is needed to be recorded and monitored for a long period (8-24 hours). This requires a lot of time for identifying problems. EKG clustering is certainly useful for simplifying the problem. In this research, we aim to automatically cluster EKG. Our technique is to recognize EKG patterns and to express them by finite automata. We have proposed a modified k-Testable method which is more flexible than the classical one. The experimental results show that EKG clustering can be significantly performed using the introduced technique.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Department: Computer Engineering .... Student's Signature ....

Field of Study: Computer Science .... Advisor's Signature ....

Academic Year: 2008 ....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยความช่วยเหลืออย่างดียิ่งจาก ผู้ช่วยศาสตราจารย์ ดร.อรรถสิทธิ์ สุรฤกษ์ และรองศาสตราจารย์ นงลักษณ์ โควาวิสารัช อาจารย์ที่ปรึกษา ซึ่งเป็นผู้ให้คำปรึกษา ให้ข้อเสนอแนะที่เป็นประโยชน์อย่างยิ่งต่อการวิจัย และช่วยตรวจแก้ไขในส่วนที่บกพร่องต่างๆ จนกระทั่งบรรลุผลสำเร็จเป็นอย่างดี

ขอขอบพระคุณในความเอื้อเฟื้อของ ศาสตราจารย์ ดร.ประภาส จงสถิตย์วัฒนา ประธานกรรมการ ผู้ช่วยศาสตราจารย์ นายแพทย์.ทายาท ดีสุดจิต ผู้ช่วยศาสตราจารย์ ดร. อานนท์ รุ่งสว่าง กรรมการสอบวิทยานิพนธ์ ที่ได้กรุณาให้คำแนะนำในการแก้ไขวิทยานิพนธ์ให้มีคุณภาพยิ่งขึ้น และขอขอบพระคุณคณาจารย์ในภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัยทุกท่านที่ประสิทธิประสาทความรู้อันมีค่ายิ่งแก่ผู้วิจัย

กราบขอบพระคุณ บิดา มารดา ที่มีความเมตตา กรุณา และเป็นผู้คอยให้กำลังใจตลอดมา ขอขอบคุณ พี่ๆ เพื่อนๆ และน้องๆ ทุกคน โดยเฉพาะสมาชิกห้องปฏิบัติการ ELITE ที่ได้ให้ความช่วยเหลือและแก้ไขเอกสาร จนกระทั่งวิทยานิพนธ์สำเร็จเป็นรูปเล่ม และขอขอบคุณแรงสนับสนุนและกำลังใจทุกท่านที่มิได้กล่าวนามไว้ ณ ที่นี้

สุดท้ายนี้ ผู้วิจัยหวังเป็นอย่างยิ่งว่างานวิจัยนี้จะเป็นประโยชน์ต่อผู้ที่สนใจหรือผู้ที่เกี่ยวข้องทั่วไป และหากมีข้อผิดพลาดประการใด ผู้วิจัยขออภัยมา ณ ที่นี้ด้วย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย .... ง
บทคัดย่อภาษาอังกฤษ .... จ
กิตติกรรมประกาศ .... ฉ
สารบัญ .... ช
สารบัญตาราง .... ฌ
สารบัญภาพ .... ญ

บทที่

1 บทนำ 1
1.1 ความเป็นมาและความสำคัญของปัญหา 1
1.2 วัตถุประสงค์ของการวิจัย 2
1.3 ขอบเขตของการวิจัย 2
1.4 ขั้นตอนและวิธีดำเนินงานวิจัย 2
1.5 ประโยชน์ที่คาดว่าจะได้รับจากงานวิจัย 2

2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 3
2.1 ทฤษฎีทางการแพทย์ด้านคลื่นไฟฟ้าหัวใจ 3
2.2 ภาษารูปนัย 4
2.3 ภาษาสม่ำเสมอ 5
2.4 เครื่องจักรแบบจำกัดสถานะ 6
2.5 การเรียนรู้ด้วยการอุปนัย 7
2.6 งานวิจัยที่เกี่ยวข้อง 7

3 การจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจ 11
3.1 การหาจุดตัด หรือจุดแยกแยะคลื่นไฟฟ้าหัวใจ 12
3.2 การแปลงข้อมูลให้อยู่ในรูปของสายอักขระ 13
3.3 การเปรียบเทียบและจัดกลุ่มสายอักขระโดยใช้เครื่องจักรตรวจสอบได้เค 19

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ หน้า

4 การทดลอง 25
4.1 วัตถุประสงค์การทดลอง 25
4.2 เครื่องมือที่ใช้ในการทดลอง 25
4.3 ข้อมูลที่ใช้ในการทดลอง 26
4.4 ขั้นตอนการทดลอง 26

5 ผลการทดลอง...27
5.1 การทดลองภาพคลื่นไฟฟ้าหัวใจแบบไม่มีความผิดปกติ.....................................27
5.2 การทดลองภาพคลื่นไฟฟ้าหัวใจแบบพบความผิดปกติ 28

6 สรุปผลการวิจัยและข้อเสนอแนะ 33
6.1 สรุปผลการวิจัย...33
6.2 ปัญหาที่พบจากการวิจัย...33
6.3 ข้อเสนอแนะ 34
รายการอ้างอิง 35
ประวัติผู้เขียนวิทยานิพนธ์ 36

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญตาราง

ตารางที่ หน้า

2.1 ค่าพื้นที่ใต้กราฟของการกระจายมาตรฐาน......9

2.2 ค่าคงที่ของจุดเส้นแบ่งการเปลี่ยนแปลงในแต่ละอักขระที่การแบ่งตั้งแต่ 3 จนถึง 10 ขั้น.10

5.1 ข้อมูลของภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติ......25

5.2 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติที่ค่า k เท่ากับ 2 26

5.3 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติที่ค่า k เท่ากับ 3 26

5.4 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติที่ค่า k เท่ากับ 4 26

5.5 ข้อมูลของชุดภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติ 27

5.6 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติที่ค่า k เท่ากับ 2..........27

5.7 ผลผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติที่ค่า k เท่ากับ 3.....28

5.8 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติที่ค่า k เท่ากับ 4..........28

5.9 จำนวนภาพคลื่นไฟฟ้าหัวใจที่ผิดพลาดในแต่ละกลุ่ม ที่พบรูปแบบที่แตกต่างปะปนกับภาพคลื่นไฟฟ้าหัวใจรูปแบบปกติ 30

5.10 เปอร์เซ็นต์ความผิดพลาดที่เกิดขึ้นจากตารางที่ 5.9 30

6.1 สรุปเปอร์เซ็นต์ความผิดพลาดในแต่ละชุดและค่าเฉลี่ยความผิดพลาดทั้งหมด 31

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญภาพ

ภาพที่ หน้า

2.1 รอบการเต้นของหัวใจ 1 รอบปกติ 3
2.2 เครื่องจักรจำกัดสถานะที่ได้จาก ภาษาตรวจสอบได้แบบเค 8
2.3 การแปลงค่าอักขระโดยใช้ความสัมพันธ์กับกฎการกระจายมาตรฐาน 9
3.1 คลื่นไฟฟ้าหัวใจปกติ และภาพที่ผ่านการหาค่าความชันและยกกำลังสอง 13
3.2 กระบวนการทำงานของการตัดแบ่งคลื่นไฟฟ้าหัวใจ 14
3.3 ภาพความสัมพันธ์ของการกระจายแบบมาตรฐาน 15
3.4 การปรับค่าอักขระด้วยความสัมพันธ์ของพื้นที่ใต้กราฟการกระจายมาตรฐาน 16
3.5 ตัวอย่างการหาระดับชั้นทั้ง 10 และการแปลงอักษร 17
3.6 ค่าคะแนนดิบ x และ y ที่ไม่ผ่านการหาค่าคะแนนซี 18
3.7 กราฟรูปที่ 3.5 เมื่อผ่านการหาคะแนนซี เส้นลำดับขั้นและอักขระแทน 18
3.8 กระบวนการแบ่งนับ และแปลงค่าอักขระ 19
3.9 เครื่องจักรจำกัดสถานะ 21
3.10 ปัญหาการเกิดการวนซ้ำของเครื่องจักรตรวจสอบได้เค 24
5.1 (ซ้าย) ภาพคลื่นไฟฟ้าหัวใจจากชุดที่ 1 ที่ผิดปกติ (ขวา) ที่ปกติ 30
5.2 (ซ้าย) ภาพคลื่นไฟฟ้าหัวใจจากชุดที่ 2 ที่ผิดปกติ (ขวา) ที่ปกติ 30
5.3 (ซ้าย) ภาพคลื่นไฟฟ้าหัวใจจากชุดที่ 3 ที่ผิดปกติ (ขวา) ที่ปกติ 30

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การใช้คอมพิวเตอร์เข้ามาช่วยในทางการแพทย์มีความสำคัญ และถูกนำมาพัฒนาอย่างต่อเนื่อง เนื่องจากในการวินิจฉัยความผิดปกติของผู้ป่วยนั้นจำเป็นต้องใช้ข้อมูลหลาย ๆ ด้านประกอบกันเพื่อความถูกต้องและแม่นยำ จึงเป็นเหตุให้ต้องใช้เวลานานในการตรวจวินิจฉัยข้อมูลผู้ป่วยจำนวนมาก หนึ่งในงานที่ต้องการพิจารณานั้นคือพิจารณาสัญญาณคลื่นไฟฟ้าหัวใจ (Electrocardiography) โดยในงานวิจัยนี้ได้ศึกษาตัวอย่างของสัญญาณคลื่นไฟฟ้าหัวใจแบบไม่เชื่อมตรง (Off-line) ซึ่งเป็นการเก็บตัวอย่างข้อมูลนาน 8 - 24 ชม. ในการตรวจสอบสัญญาณเพื่อหาความผิดปกติของการเต้นของหัวใจนั้นเป็นเรื่องละเอียดอ่อนทั้งนี้เนื่องจากในช่วงเวลายาวนาน ผู้ป่วยอาจมีพฤติกรรมหรือกิจกรรมหลายอย่างซึ่งมีผลต่อการเปลี่ยนแปลงของสัญญาณคลื่นไฟฟ้าหัวใจ ซึ่งไม่ใช่ความผิดปกติของหัวใจ ข้อมูลเหล่านี้ปะปนอยู่ในสัญญาณคลื่นไฟฟ้าหัวใจทั้งหมด จึงไม่ใช่เรื่องง่ายที่มนุษย์จะระบุว่าช่วงเวลาใดเกิดอาการผิดปกติของคลื่นหัวใจ งานวิจัยนี้จึงมุ่งเน้นที่การตรวจสอบคลื่นไฟฟ้าหัวใจจากสัญญาณ และจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจทั้งหมดโดยดูจากความเปลี่ยนแปลงในรูปแบบต่าง ๆ ของสัญญาณคลื่นไฟฟ้าหัวใจในแต่ละกิจกรรมของผู้ป่วย เพื่อเป็นเครื่องมือช่วยในการวินิจฉัยของแพทย์

การจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจในงานวิจัย ได้ศึกษาการนำทฤษฎีทางภาษารูปนัย (Formal Language Theory) เข้ามาจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจเพื่อลดความซ้ำซ้อนของข้อมูลที่มากมายเหล่านี้ให้ง่ายต่อการวินิจฉัยของแพทย์ ปัจจุบันการนำทฤษฎีภาษารูปนัยเข้ามาใช้ในการเรียนรู้เริ่มเป็นที่นิยมมากขึ้นและสามารถทำงานร่วมกับข้อมูลทางด้านชีววิทยาได้เป็นอย่างดี [1] ภาษารูปนัยถูกนำไปประยุกต์ใช้ในการตรวจจับความผิดปกติของสัญญาณคลื่นไฟฟ้าหัวใจ โดยได้ใส่การเรียนรู้รูปแบบความผิดปกติของสัญญาณคลื่นไฟฟ้าหัวใจไว้ก่อนแล้ว [2] กลุ่มข้อมูลที่ใช้ในการเรียนรู้มีลักษณะเป็นชุด มีทั้งชุดข้อมูลตัวอย่างที่ถูกต้องหรือเรียกอีกอย่างหนึ่งว่ากลุ่มข้อมูลเชิงบวก (Positive examples) และข้อมูลที่ไม่อยู่ในกลุ่มที่ถูกต้องหรือเรียกอีกอย่างหนึ่งว่าข้อมูลเชิงลบ (Negative examples) เมื่อใช้ข้อมูลทั้งสองชนิดเรียนรู้ด้วยกันกับขั้นตอนวิธีชนิด ใด ๆ ก็จะได้เป็น ออโตมาตา (Automata) [3-4] หรือเป็นไวยากรณ์ (Grammar) ซึ่งจะสามารถนำมาใช้เป็นบรรทัดฐานในการเปรียบเทียบกับชุดข้อมูลถัดไปได้ว่าเป็นข้อมูลที่เหมือนหรือต่างกันหรือไม่ แต่ลักษณะของข้อมูลในงานวิจัยนี้เป็นข้อมูลหลายแบบที่ รวมอยู่ชุดเดียว โดยทราบรูปแบบปกติหรือผิดปกติ

จากปัญหา ดังกล่าวงานวิจัยนี้จึงได้คิดหาวิธีที่จะนำทฤษฎีเชิงวากยสัมพันธ์ (Syntactic Theory) [2][5][6][7] มาประยุกต์ใช้เพื่อสร้างไวยากรณ์ [8] หรือออโตมาตา เพื่อให้ได้มาซึ่งข้อมูลทางด้านสถิติที่เพียงพอ และนำมาวิเคราะห์ จัดกลุ่มคลื่นไฟฟ้าหัวใจต่อไป

1.2 วัตถุประสงค์ของการวิจัย

เพื่อหาวิธีการจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจโดยใช้ทฤษฎีการเรียนรู้เชิงวากยสัมพันธ์ และพัฒนาเครื่องมือในการจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจเพื่อให้ง่ายต่อบุคลากรทางการแพทย์ในการวิเคราะห์ข้อมูลสัญญาณคลื่นไฟฟ้าหัวใจจำนวนมาก

1.3 ขอบเขตของการวิจัย

1.3.1 ไวยากรณ์หรือออโตมาตาอย่างหนึ่งอย่างใดที่สร้างขึ้นมาได้นั้น ใช้เพื่อจัดกลุ่มชุดของข้อมูลเพียงข้อมูลตัวอย่างนั้น ๆ หรือเฉพาะคนผู้ป่วยคนเดียวกันนั้น

1.3.2 เครื่องมือที่ได้ไม่สามารถบอกความผิดปกติของหัวใจในทางการแพทย์

1.3.3 ข้อมูลที่นำมาจัดกลุ่ม มีความยาวโดยประมาณ 8 – 24 ชั่วโมง

1.4 ขั้นตอนและวิธีดำเนินการวิจัย

1.4.1 ศึกษาข้อมูลและแบบรูป (Pattern) ที่เกี่ยวข้องของสัญญาณคลื่นไฟฟ้าหัวใจ

1.4.2 ศึกษาวิธีการที่เกี่ยวข้องในการแปรสภาพข้อมูลหรือควอนไตซ์ (Quantize) ที่มีประสิทธิภาพเพื่อนำมาใช้ในการควอนไตซ์ข้อมูลสัญญาณคลื่นไฟฟ้าหัวใจ

1.4.3 ศึกษาวิธีการที่เกี่ยวข้องในการนำข้อมูลที่ผ่านกระบวนการแปรสภาพมาสร้างออโตมาตา (Automata) แบบจำกัด

1.4.4 ศึกษาวิธีการทางทฤษฎีวากยสัมพันธ์เพื่อนำมาประยุกต์ใช้ในการสร้างไวยากรณ์ทางข้อมูล

1.4.5 พัฒนา และออกแบบระบบจากที่ได้ศึกษาเพื่อทดลองผล

1.4.6 ทดสอบความถูกต้องของงานวิจัยจากข้อมูลจำลองและข้อมูลจริง

1.4.7 วิเคราะห์ สรุปผล และจัดทำวิทยานิพนธ์

1.5 ประโยชน์ที่ได้รับจากงานวิจัย

1.5.1 ได้วิธีอีกรูปแบบหนึ่งในการจัดกลุ่มของสัญญาณคลื่นไฟฟ้าหัวใจ

1.5.2 ช่วยให้แพทย์หรือบุคลากรทางการแพทย์ง่ายต่อการหาความเปลี่ยนแปลงของสัญญาณคลื่นไฟฟ้าหัวใจ

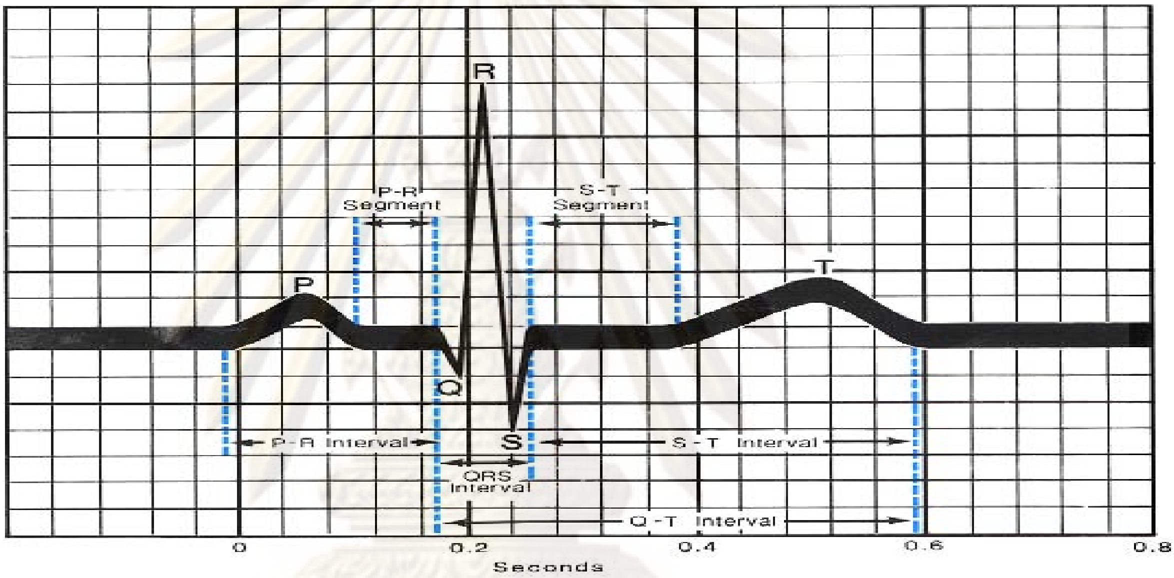
ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีทางการแพทย์ด้านภาพคลื่นไฟฟ้าหัวใจ

คลื่นไฟฟ้าหัวใจทางคลินิคเป็นทฤษฎีที่กล่าวถึงลักษณะของภาพคลื่นไฟฟ้าหัวใจ การเต้นของหัวใจที่ควรต้องพิจารณาของภาพคลื่นไฟฟ้าหัวใจในแต่ละแบบเพื่อนำมาประกอบการพิจารณาในการจัดกลุ่มของภาพคลื่นไฟฟ้าหัวใจ



รูปที่ 2.1 รอบการเต้นของหัวใจหนึ่งรอบปกติ

จากรูปที่ 2.1 ภาพคลื่นไฟฟ้าหัวใจประกอบด้วยสัญญาณ P, Q, R, S, T ซึ่งเป็นตัวแทนของคลื่นที่เกิดขึ้นในแต่ละรอบการเต้นของหัวใจโดยมีความหมายดังต่อไปนี้

P (Depolarization) คือช่วงที่กล้ามเนื้อหัวใจได้รับสัญญาณไฟฟ้าจากปุ่มไซนัสที่อยู่ในหัวใจห้องบนขวา และมีเลือดไหลมาที่หัวใจห้องบนขวา

QRS (Ventricular Depolarization) คือช่วงที่มีเลือดไหลเข้าใส่หัวใจห้องล่าง และบีบตัวส่งเลือดออกไปที่ร่างกายส่วนต่าง ๆ รวมถึงเลือดเสียที่ไหลไปสู่ปอด

T (Ventricular Repolarization) คือช่วงที่หัวใจกลับสู่ภาวะปกติก่อนที่จะเริ่มรับเลือดเข้าสู่หัวใจห้องบนในรอบถัดไป

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

คุณสมบัติทั่วไปของภาพคลื่นไฟฟ้าหัวใจ

ปกติใน 1 นาทีหัวใจจะเต้นประมาณ 80 ครั้ง
คลื่นไฟฟ้าหัวใจ P จะยาว 0.12 – 0.14 วินาที
คลื่นไฟฟ้าหัวใจ Q จะยาวไม่เกิน 0.04 วินาที
คลื่นไฟฟ้าหัวใจ QRS ของคนปกติจะยาวไม่เกิน 0.10 วินาที
คลื่นไฟฟ้าหัวใจ T นั้นไม่มีมาตรฐานที่แน่นอนของความกว้าง

2.2 ภาษารูปนัย (Formal language)

ภาษารูปนัยคือกลุ่มหรือชุดของภาษา ซึ่งคำต่าง ๆ ในภาษามีกฎหรือกติกาที่ชัดเจนในการพิจารณาว่าเป็นสมาชิกในภาษาโดยปราศจากความกำกวม นอกจากนี้ จำนวนของกฎหรือกติกาต้องมีเป็นจำนวนจำกัดด้วย

เราให้คำจำกัดความของชุดตัวอักษร ตัวอักษรหรืออักขระ คำ และภาษาไว้ดังนี้

นิยามที่ 2.1 ชุดตัวอักษร (Alphabet) หมายถึงเซตจำกัดของสัญลักษณ์ที่เป็นหน่วยย่อยสุดไม่สามารถแบ่งแยกได้ นิยมใช้สัญลักษณ์แทนชุดตัวอักษรด้วยซิกมา Σ และเรียกสมาชิกในชุดตัวอักษรว่าอักขระหรือตัวอักษร (Character)

นิยามที่ 2.2 สายอักขระ (String) หมายถึงลำดับของชุดตัวอักษร ถ้าลำดับมีจำนวนอักขระเป็นจำนวนจำกัด สายอักขระนั้นจะถูกเรียกว่าสายอักขระจำกัด (Finite string) แต่ถ้าลำดับมีจำนวนอักขระเป็นอนันต์สายอักขระนั้นจะถูกเรียกว่าสายอักขระอนันต์ (Infinite string)

นิยามที่ 2.3 ภาษา (Language) หมายถึงชุดของสายอักขระจำกัดที่มีจำนวนของอักขระเป็นจำนวนจำกัดโดยสมาชิกในภาษาจะถูกเรียกว่าคำ (Word)

นิยามที่ 2.4 สายอักขระที่ไม่มีอักขระเลยจะถูกเรียกว่าเป็น สายอักขระว่าง (Null-string or Empty string) นิยมใช้สัญลักษณ์แทนด้วยแลมด้า (λ)

ในกรณีที่ภาษามีจำนวนคำเป็นอนันต์ ภาษานั้นจะถูกเรียกว่าเป็น ภาษาอนันต์ (Infinite language) ไม่เช่นนั้นจะเป็น ภาษาจำกัด (Finite language)

นิยามที่ 2.5 ความยาวของสายอักขระ (String length) หมายถึงจำนวนของอักขระที่ประกอบอยู่ในสายนั้น สำหรับสายอักขระว่าง เราจะถือว่า ความยาวจะมีค่าเป็น 0 เสมอ

2.3 ภาษาสม่ำเสมอ (Regular Language)

ภาษาสม่ำเสมอเป็นการอธิบายภาษารูปแบบหนึ่งซึ่งเราสามารถนิยามตัวดำเนินการในการอธิบายว่าภาษานั้น เป็นภาษาสม่ำเสมอได้ดังนี้

นิยามที่ 2.6 ตัวดำเนินการคลีนสตาร์ (Kleene's star) ที่บนเซตของสายอักขระ ใดๆ นิยามโดย

$$S^* = \bigcup_{i=0}^{\infty} S^i \qquad (2.1)$$

โดยที่ $S^i = \{x_1 x_2 x_3 \ldots x_i \mid \forall x_j \in S\}$

ตัวดำเนินการคลีนสตาร์ หรือบางทีเรียกว่า ตัวดำเนินการปิดของคลีน (Kleene's closure) เป็นตัวดำเนินการหนึ่งที่นิยมใช้กันในการกำหนดภาษา และจะเห็นว่า ภาษาที่เกิดจากตัวดำเนินการนี้จะสามารถเป็นเซตอนันต์ได้

นิยามที่ 2.7 ตัวดำเนินการบวก (Positive closure) นิยามบนเซต S ใด ๆ หมายถึง

$$S^+ = \bigcup_{i=1}^{\infty} S^i \qquad (2.2)$$

โดยที่ $S^i = \{x_1 x_2 x_3 \ldots x_i \mid \forall x_j \in S\}$

จากนิยามเราพอสังเกตได้ว่าสำหรับเซต S ใด ๆ จะได้ว่า $S^+ \subseteq S^*$

ภาษาที่มีขนาดใหญ่ที่สุดคือ ภาษาที่ประกอบด้วยทุกคำที่สามารถสร้างได้จาก $\sum$ สามารถเขียนแทนด้วย $\sum^*$ เช่น $\sum = \{0,1\}$ เขียนแทนได้ด้วย $\sum = \{0,1\}^* = \{\lambda, 0, 1, 00, 01, 10, 11, 000, 001, \ldots\}$

ดังนั้นทุก L ที่เกิดขึ้นจาก $\sum$ จะได้ว่า $L \subseteq \sum^*$

นิยามที่ 2.8 การบรรยายสม่ำเสมอ (Regular expression) กำหนดให้ $\sum$ เป็นชุดตัวอักษร และการบรรยายสม่ำเสมอหมายถึงข้อกำหนดต่อไปนี้

1. ทุกสมาชิกใน $\sum$ เป็น การบรรยายสม่ำเสมอ
2. อักขระว่างเป็นการบรรยายสม่ำเสมอ

3. สำหรับ x และ y ที่เป็น การบรรยายสม่ำเสมอ แล้วจะได้ว่า $(x), x+y, xy$ และ x^* เป็น การบรรยายสม่ำเสมอด้วย
4. ใช้เฉพาะกฎ 3 ข้อข้างต้นเท่านั้นในการสร้าง การบรรยายสม่ำเสมอ

นิยามที่ 2.12 ภาษาใดที่สามารถเขียนบรรยายได้ด้วย การบรรยายสม่ำเสมอ จะเรียกภาษานั้นว่าเป็น ภาษาสม่ำเสมอ (Regular language)

2.4 เครื่องจักรแบบจำกัดสถานะ (Finite State Machine)

การตรวจสอบการเป็นสมาชิกภายในภาษานั้น สามารถใช้ตัวแบบทางคณิตศาสตร์และยังเป็นรูปแบบอย่างง่ายในการตรวจสอบเรียกว่า เครื่องจักรจำกัดสถานะ หรือ ออโตมาตา โดยการสร้างตัวแบบที่เหมาะสมกับภาษาที่กำหนดมาให้ และตัวแบบนี้จะสามารถตอบคำถามได้เพียง สองรูปแบบคือ เป็นสมาชิก หรือไม่เป็นสมาชิกเท่านั้น

นิยามที่ 2.13 เครื่องจักรจำกัดสถานะ M ประกอบด้วยส่วนสำคัญ 5 ส่วน คือ

$$M = (Q, \Sigma, q_0, A, \delta)$$

โดยที่

Q เป็นเซตจำกัดของสถานะ (State)

Σ เป็นเซตจำกัดของชุดข้อมูลนำเข้า เรียกว่าอักขระ

q_0 เป็นสถานะเริ่มต้น (Initial state) และ $q_0 \in Q$

A เป็นสถานะของสถานะยอมรับ (Accepted state) และ $A \subseteq Q$

δ เป็นฟังก์ชันการเปลี่ยนที่นิยามโดย $\delta : Q \times \Sigma \to Q$

การทำงานของออโตมาตานั้น จะทำงานโดยการรับข้อมูลนำเข้าที่เป็นสายอักขระโดย การอ่านข้อมูลทีละอักขระ ตามลำดับ ทั้งนี้เพราะออโตมาตาเป็นเครื่องจักรที่มีการทำงานแบบลำดับ (sequential machine) โดยในตอนเริ่มต้นก่อนการทำงาน ออโตมาตาจะอยู่ในสถานะเริ่มต้น เมื่อข้อมูลนำเข้าถูกอ่านทีละอักขระ ออโตมาตาจะเปลี่ยนสถานะของเครื่องจักรไปตามข้อมูลที่อ่านเข้ามานั้น ซึ่งการเปลี่ยนสถานะนี้ได้ถูกกำหนดไว้แล้วโดยฟังก์ชันการเปลี่ยนสถานะ ออโตมาตาจะดำเนินการอ่านต่อไปทีละอักขระพร้อมกับการพิจารณาเปลี่ยนสถานะไปตามฟังก์ชัน จนกระทั่งอักขระสุดท้ายถูกอ่านไปเรียบร้อยแล้วออโตมาตาจะหยุดทำงาน สถานะสุดท้ายที่ออโตมาตาหยุดการทำงานจะมีผลต่อคำตอบ เพราะถ้าออโตมาตาหยุดที่สถานะที่เป็นสถานะยอมรับ เราจะเรียกว่าออโตมาตานั้นยอมรับข้อมูลนำเข้า ไม่เช่นนั้นจะเรียกว่าออโตมาตานั้นปฏิเสธข้อมูลนำเข้า

2.5 การเรียนรู้ด้วยการอุปนัย (Inductive learning)

การเรียนรู้ด้วยการอุปนัย หมายถึง การเรียนรู้จากข้อมูลตัวอย่างเพื่อให้ได้มาซึ่งกฎเกณฑ์ทั่วไปของชุดข้อมูลตัวอย่างนั้น [6] ตัวอย่างเช่น การเรียนรู้จากตัวอย่างชุดข้อมูล aaabb, aabb, aabbbb, aaaabbbb และ aaabbbb อาจจะสรุปได้ว่า กฎเกณฑ์ที่อธิบายสภาพข้อมูลนี้คือ สายอักขระที่ขึ้นต้นด้วย a ไม่จำกัดจำนวน และตามด้วย b ไม่จำกัดจำนวนหรือเป็นอักขระที่ลงท้ายด้วย b ในการเรียนรู้ด้วยอุปนัยนั้นมีองค์ประกอบของการเรียนรู้

1. คลาสของภาษาที่พิจารณา ตัวอย่างเช่น กฎเกณฑ์ที่ได้จากการ เรียนรู้อยู่ในคลาสของภาษาไวยากรณ์สม่ำเสมอ
2. คลาสของไวยากรณ์ หมายถึง ไวยากรณ์ที่ได้จากการเรียนรู้ ตัวอย่างเช่น ในการเรียนรู้ภาษาสม่ำเสมอ เราจะออกแบบกฎให้อยู่ในรูปของออโตมาตา หรือ ไวยากรณ์สม่ำเสมอโดยตัวอย่างของภาษา
3. ตัวอย่างของภาษา (Sample) เป็นส่วนหนึ่งในภาษาที่พิจารณามี 2 ประการคือ ตัวอย่างที่ประกอบด้วยสายอักขระที่อยู่ในภาษา และตัวอย่างที่ประกอบด้วยสายอักขระที่ไม่อยู่ในภาษา
4. ขั้นตอนวิธีการเรียนรู้
5. เกณฑ์ในการวัดความถูกต้องของการเรียนรู้

2.6 งานวิจัยที่เกี่ยวข้อง

2.6.1 การอนุมานภาษาตรวจสอบได้เค (k-Testable Languages Inference)

โกลด์ [9] เสนอทฤษฎีไว้ว่าภาษาสม่ำเสมอ สามารถที่จะเรียนรู้ในขอบเขตจำกัด (Identification in the Limit) โดยใช้ทั้งชุดข้อมูลตัวอย่างบวก และชุดข้อมูลตัวอย่างลบ ต่อมาในปี 1990 กาเซียร์ และวิเดล [10] ได้เสนอการเรียนรู้ด้วยการอนุมาน โดยใช้ชุดข้อมูลบวก (Positive data) เพียงอย่างเดียวและได้ผลลัพธ์เครื่องจักรตรวจสอบได้ k (k-Testable Machine) ที่สามารถตรวจสอบการเป็นสมาชิกภายในภาษาได้

ศูนย์วิทยทรัพยากร

นิยามที่ 2.14 กำหนดให้ k มากกว่าหรือเท่ากับ 0 และเครื่องจักรตรวจสอบได้เคหรือ K-TSS ประกอบด้วยสี่ส่วนสำคัญ

$$Z_k = (\Sigma, I, F, T)$$

จุฬาลงกรณ์มหาวิทยาลัย

Σ เป็นเซตจำกัดของอักขระ

I เป็นเซตโดยที่ $I \subseteq \Sigma^{<k}$ โดยที่ I เป็นส่วนเติมหน้า (Prefix) เท่านั้น

F เป็นเซตโดยที่ $F \subseteq \Sigma^{<k}$ โดยที่ F เป็นส่วนเติมท้าย (Suffix) เท่านั้น

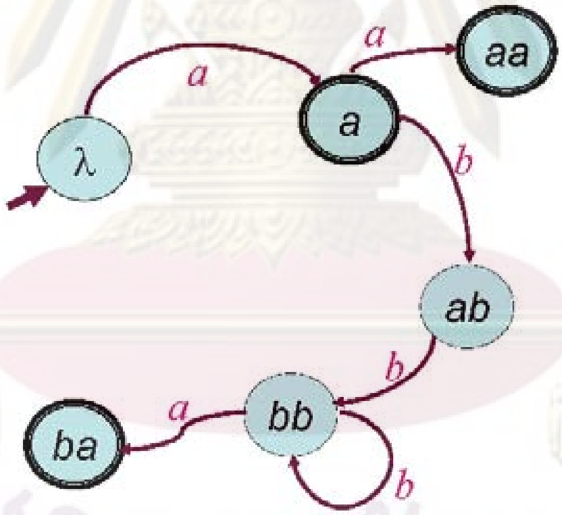
T เป็นเซตโดยที่ $T \subseteq \Sigma^{k}$

จากเครื่องจักรตรวจสอบได้เคสามารถที่จะแปลงให้อยู่ในรูปแบบของเครื่องจักรแบบจำกัดสถานะ หรือออโตมาตาได้โดย

แต่ละสายอักขระใน I เป็นสถานะ
แต่ละสายอักขระที่มีความยาวเท่ากับ k-1 ใน T เป็นสถานะ
ค่าว่าง (Null) เป็นสถานะเริ่มต้น
เพิ่มการเปลี่ยนสถานะโดยมีป้าย b จาก u ไปที่ ub ในแต่ละสายอักขระใน I
เพิ่มการเปลี่ยนสถานะโดยมีป้าย b จาก au ไปที่ ub ในแต่ละสายอักขระใน T
แต่ละสายอักขระใน F เป็นสถานะสุดท้าย หรือสถานะยอมรับ

ตัวอย่างที่ 2.6.1 ให้ X = {'a', 'aa', 'abba', 'abbbba'} และกำหนดให้ k=3 จะได้เครื่องจักรตรวจสอบได้เค เป็น

I = {'Null', 'a', 'aa', 'ab'}
F = {'a', 'aa', 'ba'}
T = {'abb', 'bbb', 'bba'}
และสามารถแปลงเป็นเครื่องจักรจำกัดตรวจสอบได้เค ได้เป็น



รูปที่ 2.2 เครื่องจักรตรวจสอบได้เคที่ได้จากตัวอย่างที่ 2.6.1

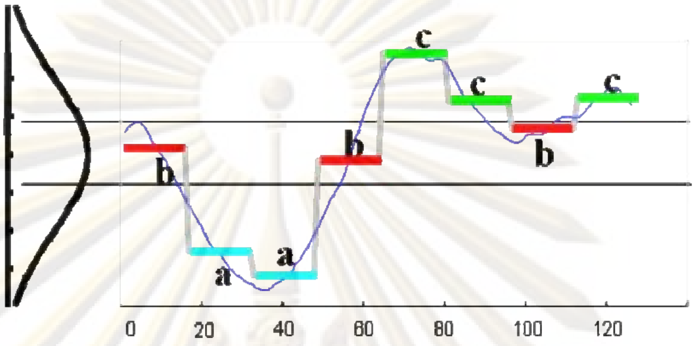
2.6.2 A Symbolic Representation of Time Series, with Implications for Streaming Algorithms.

ในปี 2003 ลิน, เคิร์ธ, โลนาดิ และชุย [11] ได้เสนอการใช้อักขระในการเปรียบเทียบหาความสัมพันธ์ของชุดข้อมูลเวลา สองชุดโดยอาศัยกฎพื้นฐานทางคณิตศาสตร์ และสถิติ เข้ามาช่วยในการแปลงชุดข้อมูลเวลาให้เป็นอักขระ และนำมาเปรียบเทียบกัน โดยมีขั้นตอนหลักดังนี้

1. แบ่งนับ (Quantize) ข้อมูลออกเป็นส่วน ๆ เพื่อหาข้อมูลร่วม และยังเป็นการปรับลดขนาดของข้อมูลที่มีความละเอียดสูง

2. ปรับค่าตัวเลขให้เป็นค่าคะแนนซี (Z-Score) สำหรับเปรียบเทียบกับชุดข้อมูลตัวอื่น

3. ค่าตัวเลขที่ได้จะสามารถแบ่งเป็นลำดับชั้นของอักขระที่นำมาแทนได้ โดยอิงลำดับชั้นจากการเปลี่ยนแปลงของข้อมูลตาม กฎการกระจายมาตรฐาน (Normal Distribution)



รูปที่ 2.3 การแปลงค่าอักขระโดยใช้ความสัมพันธ์กับกฎการกระจายมาตรฐาน

ซึ่งค่าความกว้างในแต่ละลำดับชั้นของอักขระนั้น สามารถคำนวณหาได้จาก ตารางสถิติมาตรฐาน (Standard Statistical Tables) ตามตารางที่ 2.1 ได้ออกมาเป็นตารางค่าคงที่ 2.2 โดยจะกล่าวเพิ่มเติมในบทต่อไป

ตารางที่ 2.1 แสดงค่าพื้นที่ใต้กราฟของการกระจายมาตรฐาน

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177

1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.975	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.983	0.9834	0.9838	0.9842	0.9846	0.985	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.989
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.992	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.994	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.996	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.997	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.998	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.999	0.999

ตารางที่ 2.2 ค่าคงที่ของจุดเส้นแบ่งการเปลี่ยนแปลงในแต่ละอักขระ ที่การแบ่งตั้งแต่ 3 จนถึง 10 ขั้น

3	4	5	6	7	8	9	10
-0.43	-0.67	-0.84	-0.97	-1.07	-1.15	-1.22	-1.28
0.43	0	-0.25	-0.43	-0.57	-0.67	-0.76	-0.84
	0.67	0.25	0	-0.18	-0.32	-0.43	-0.52
		0.84	0.43	0.18	0	-0.14	-0.25
			0.97	0.57	0.32	0.14	0
				1.07	0.67	0.43	0.25
					1.15	0.76	0.52
						1.22	0.84
							1.28

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

การจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจ

จากบทที่แล้วได้กล่าวถึงงานวิจัยที่เกี่ยวข้องกับการจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจในงานวิจัยนี้ ซึ่งจะนำมาช่วยในการจัดกลุ่มสัญญาณคลื่นไฟฟ้าหัวใจ โดยสัญญาณคลื่นไฟฟ้าหัวใจ หรือข้อมูลที่นำมาใช้ในการจัดกลุ่มนั้นเป็นข้อมูลที่ผ่านการตรวจจับมาจากผู้ป่วยเรียบร้อยแล้วและมีระยะเวลาตรวจจับสัญญาณมาจากผู้ป่วยประมาณ 8 – 24 ชั่วโมงจึงนำมาจัดกลุ่มจึงไม่ใช่การตรวจจับจากผู้ป่วยและจัดกลุ่มในขณะเดียวกัน ข้อมูลสัญญาณคลื่นไฟฟ้าหัวใจจากที่กล่าวมาข้างต้นจึงเป็นข้อมูลที่ยังไม่สามารถที่จะบอกได้ว่ามีจำนวนสัญญาณคลื่นไฟฟ้าหัวใจอยู่เป็นจำนวนเท่าใด แต่การจัดกลุ่มนั้นคือการนำสัญญาณคลื่นไฟฟ้าหัวใจแต่ละรอบการเต้นของหัวใจมาเปรียบเทียบซึ่งกันและกันเพื่อหาว่ารอบการเต้นที่นำมาเปรียบเทียบเป็นสัญญาณที่เหมือนหรือต่างกัน งานวิจัยนี้จึงต้องมีกระบวนการก่อนการจัดกลุ่มขึ้นคือการตัดสัญญาณคลื่นไฟฟ้าหัวใจออกเป็นสัญญาณแต่ละสัญญาณคลื่นไฟฟ้าหัวใจ โดยแต่ละสัญญาณที่ตัดได้ประกอบไปด้วยสัญญาณคลื่นไฟฟ้าหัวใจเพียงคลื่นเดียวหรือหนึ่งรอบการเต้นของหัวใจ เพื่อนำสัญญาณคลื่นไฟฟ้าหัวใจแต่ละสัญญาณมาเปรียบเทียบหารูปแบบว่าเป็นรูปแบบเดียวกันและควรจัดให้อยู่ในกลุ่มเดียวกันหรือไม่ การจะนำสัญญาณคลื่นไฟฟ้ามาหารูปแบบได้นั้นในงานวิจัยนี้ใช้วิธีการแปลงชุดข้อมูลตัวเลขที่บอกถึงระดับการเต้นของหัวใจแต่ละจุดที่เครื่องมือตรวจจับสัญญาณคลื่นไฟฟ้าหัวใจได้มาเป็นอักขระแทน เทคนิคที่ใช้ในการแปลงชุดค่าตัวเลขคือ SAX [11] ซึ่งจะได้อธิบายถึงขั้นตอนวิธีในการแปลงในหัวข้อการแปลงข้อมูลให้อยู่ในรูปของสายอักขระต่อไป เมื่อได้สายอักขระแทนระดับสัญญาณมาแล้ว จึงนำสายอักขระนี้มาหารูปแบบโดยใช้วิธีสร้างเครื่องจักรตรวจสอบ เพื่อใช้ในการตรวจสอบว่ารูปแบบของสัญญาณคลื่นไฟฟ้าหัวใจสองตัวที่นำมาเปรียบเทียบมีความสัมพันธ์กันหรือไม่

จากที่กล่าวมาข้างต้นงานวิจัยนี้ต้องการหาจำนวนกลุ่มของรูปแบบที่เป็นไปได้ทั้งหมดที่เกิดขึ้นกับผู้ป่วยหนึ่งราย ไม่สามารถระบุได้ว่าผู้ป่วยรายใดมีรูปแบบของสัญญาณคลื่นไฟฟ้าหัวใจเป็นจำนวนกลุ่มเท่าใด เป็นเหตุให้ข้อมูลที่ใช้ในการหารูปแบบและสร้างเครื่องจักรจำกัดสถานะในการจัดกลุ่มไม่มีข้อมูลที่นำมาใช้ในการลดทอนสถานะของเครื่องจักรจำกัดสถานะ เหมือนวิธีการหารูปแบบอื่นซึ่งใช้ข้อมูลรูปแบบของกลุ่มอื่นมาลดทอนสถานะของเครื่องจักรจำกัดสถานะ การลดทอนสถานะของเครื่องจักรจำกัดสถานะให้ผลดีในด้านความรวดเร็วของเครื่องจักรที่ส่งคำตอบของความเป็นสมาชิกเพื่อจัดกลุ่มและทำให้ไม่เกิดความผิดพลาดในการปฏิเสธ (False Negative) งานวิจัยนี้ได้ใช้วิธีการของการหาความสัมพันธ์โดยใช้เครื่องจักรตรวจสอบได้เค ซึ่งเป็นหนึ่งในงานวิจัยที่สร้างเครื่องจักรจำกัดสถานะโดยมีความ

พิเศษคือการสร้างเครื่องจักรจำกัดสถานะโดยไม่จำเป็นต้องใช้ข้อมูลทดสอบเชิงลบเพื่อลดทอนสถานะของเครื่องจักรจำกัดสถานะ

แบ่งเป็นขั้นตอนการทำงานหลักสามขั้นตอน คือ

1. การแบ่งคาบหาหนึ่งรอบคลื่นไฟฟ้าหัวใจ
2. การแปลงข้อมูลให้อยู่ในรูปของสายอักขระ
3. การเปรียบเทียบสัญญาณคลื่นไฟฟ้าหัวใจและจัดกลุ่มสายอักขระ

3.1 การแบ่งคาบหาหนึ่งรอบคลื่นไฟฟ้าหัวใจ

สัญญาณคลื่นไฟฟ้าหัวใจที่ได้จากการตรวจวัดนั้นเป็นสัญญาณเรียงต่อกันโดยไม่มีจุดใดบอกถึงจุดของการเริ่มต้นรอบสัญญาณคลื่นไฟฟ้าหัวใจ โดยการแบ่งคาบนั้นมีความต้องการที่จะได้สัญญาณคลื่นไฟฟ้าหัวใจออกมาเพียงหนึ่งรอบสัญญาณคลื่นไฟฟ้าหัวใจที่ประกอบไปด้วย พี (P), คิว (Q), อาร์ (R), เอส (S) และที (T) เท่านั้นซึ่งถือเป็นหนึ่งลูกคลื่น

คลื่นไฟฟ้าหัวใจมีคุณสมบัติบางประการดังนี้

1. จุดสูงสุดเรียกว่าจุดอาร์ พบในคนทุกคน ไม่ว่าจะเป็นผู้ป่วยหรือไม่ก็ตาม มีค่าสูงเพียงใดนั้นขึ้นอยู่กับความแข็งแรงของแต่ละคนและตำแหน่งในการติดอุปกรณ์ตรวจวัดคลื่นไฟฟ้าหัวใจ ส่วนลูกคลื่นอื่น ๆ (พี, คิว, เอส, ที, ยู) มีความสูงไม่เกินไปกว่าจุดสูงสุดของจุดอาร์ [12]
2. คลื่นไฟฟ้าหัวใจตำแหน่งพี ของคนปกติ อยู่ในช่วง 0.12 – 0.14 วินาที ซึ่งถ้ามากกว่า 0.12 วินาทีถือว่าเริ่มผิดปกติ
3. คลื่นไฟฟ้าหัวใจตำแหน่งคิว ของคนปกติยาวไม่เกิน 0.04 วินาที
4. คลื่นไฟฟ้าหัวใจตำแหน่งคิวอาร์เอส ของคนปกติยาวไม่เกิน 0.10 วินาที
5. คลื่นไฟฟ้าหัวใจตำแหน่งที ของคนปกติยาวแตกต่างกันออกไป
6. เส้นมาตรฐานสัญญาณคลื่นไฟฟ้าหัวใจ เป็นเส้นตรงเดียวกันไม่โค้งขึ้นลง นอกจากจะเกิดจากการติดตั้งอุปกรณ์ตรวจวัดไม่ดีพอ ผิดตำแหน่ง หรือไม่มั่นคง

3.1.1 การหาจุดสูงสุดอาร์

ด้วยคุณสมบัติดังกล่าวข้างต้น การหาจุดตัดของแต่ละการเต้นของหัวใจโดยใช้การตรวจหาจุดสูงสุดอาร์ ซึ่งเกิดขึ้นหนึ่งครั้งในการเต้นของหัวใจหนึ่งรอบ โดยกำหนดค่าวัดระดับความเป็นไปได้ที่จะเป็นค่าสูงสุดอาร์ขึ้นหนึ่งค่า เมื่อค่าใดอยู่สูงกว่าค่าที่กำหนดจะถือว่าในตำแหน่งนั้นมีความเป็นไปได้สูงที่จะเป็นตำแหน่งของจุดสูงสุดอาร์ รอบการเต้นของหัวใจหนึ่ง

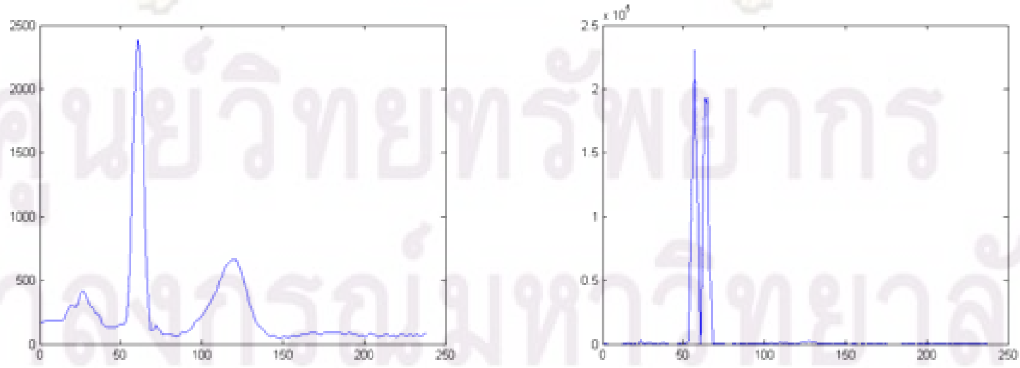
รอบเกิดอาร์ครั้งเดียวเพราะฉะนั้นงานวิจัยนี้จะไม่สนใจค่าสูงสุดที่อยู่ติดกันเป็นจำนวน 30 จุดจึงทำให้ระบบสามารถกำหนดจุดอาร์ได้ว่าเกิดขึ้น ตรงตำแหน่งใดบ้างแต่ความต้องการของงานวิจัยนี้คือการหาหนึ่งรอบการเต้นของหัวใจซึ่งจุดอาร์ที่หาได้เป็นตำแหน่งที่ผ่านจุดพีมาแล้วจึงต้องหาคาบเวลาที่ต้องย้อนกลับไปเพื่อรวมเอาจุดพีเข้ามาในรอบการเต้นของหัวใจหนึ่งรอบ

3.1.2 การหาคาบรวมพี คิว อาร์ เอส และที

เมื่อได้ค่าสูงสุดอาร์มาแล้วต่อไปต้องคำนวณหาความยาวของช่วงเวลาการเกิดคลื่นพีและคิวเพราะคลื่นพีและคิวนั้นจะเกิดขึ้นก่อนหน้าจุดอาร์ซึ่งสามารถระบุตำแหน่งมาได้แล้วในหัว 3.1.1 เพื่อรวมเอาคลื่นพีและคิวเข้ามาในหนึ่งคาบของรอบการเต้นของหัวใจ และต้องเป็นค่าปกติสูงสุดรวมกัน โดยใช้การคำนวณความเป็นไปได้นานที่สุดที่จุดพีจะเกิดในหนึ่งรอบการเต้นของหัวใจ (0.14 วินาที) บวกด้วยความเป็นไปได้นานที่สุดที่จุดคิวจะเกิดในหนึ่งรอบการเต้นของหัวใจ (0.04 วินาที) บวกด้วยความเป็นไปได้นานที่สุดที่จุดคิวอาร์และเอสจะเกิดในหนึ่งรอบการเต้นของหัวใจ (0.10 วินาที) แต่มีการบวกค่าคิวก่อนหน้านี้จึงต้องลบช่วงเวลาที่เกิดคิวออกจาก จุดคิวอาร์และเอส หารด้วยสองเพื่อแบ่งครึ่งจุดอาร์ซึ่งเป็นจุดสูงสุดออกมา ก็จะได้จำนวนเวลาที่จะต้องนับย้อนหลังไปเพื่อให้ได้จุดพีรวมเข้ามาในหนึ่งรอบสัญญาณการเต้นของหัวใจ

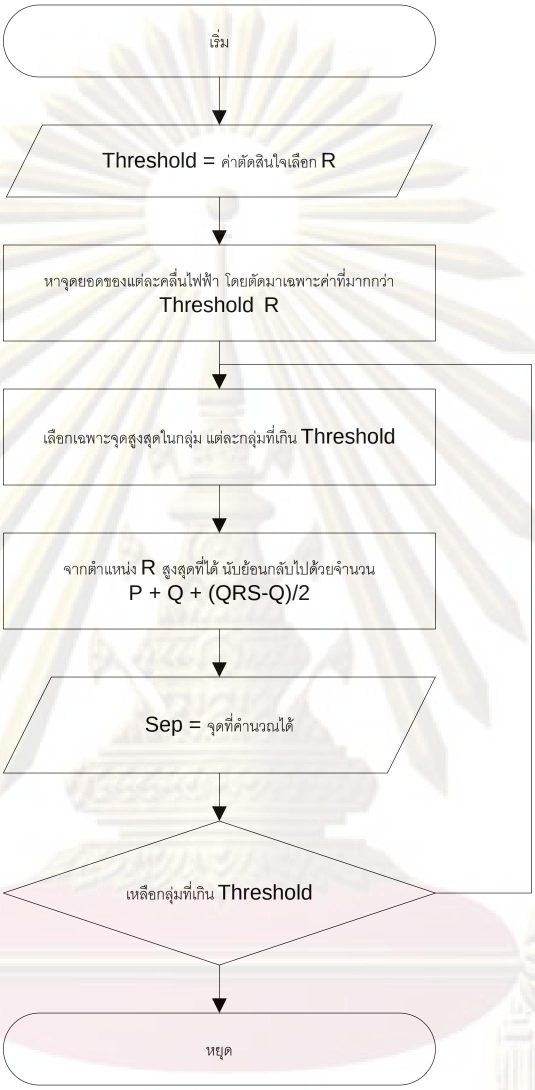
3.1.3 ปัญหาจุดสูงสุดอาร์มีค่าเป็นลบ

ด้วยสัญญาณคลื่นไฟฟ้าหัวใจบางประเภทนั้น จุดสูงสุดอาร์ไม่ได้มีค่าเป็นบวกหรือพุ่งสูงขึ้นอย่างสัญญาณคลื่นไฟฟ้าหัวใจปกติ แต่กลับเป็นค่าลบซึ่งทำให้เกิดคลื่นอาร์ ที่ดิ่งลงจึงไม่สามารถหาค่าอาร์ หรือทำให้ระบบเข้าใจว่าค่าสูงสุดที่ตำแหน่งอื่นเป็นค่าอาร์ แทนซึ่งทำให้เกิดการหาคาบสัญญาณคลื่นไฟฟ้าหัวใจที่ผิดพลาด จึงมีวิธีการหาจุดอาร์โดยใช้ค่าความชันแทนซึ่งจุดอาร์นั้นโดยปกติจะมีความชันที่สูงมากแต่เพื่อให้ความชันนั้นชัดเจนมากยิ่งขึ้นและทำให้ค่าอาร์ที่เป็นลบนั้นกลับเป็นค่าบวกด้วยจะมีการยกกำลังสองเพิ่มเข้าไปดังแสดงในรูปที่ 3.1 อีกทั้งทำให้การหาจุดอาร์มีความทนทานต่อสัญญาณรบกวนมากยิ่งขึ้น



รูปที่ 3.1 ภาพคลื่นไฟฟ้าหัวใจปกติ และสัญญาณที่ผ่านการหาค่าความชันและยกกำลังสอง

รูปที่ 3.2 แสดงถึงกระบวนการทำงานในการแบ่งคาบหาหนึ่งรอบคลื่นการเต้นของหัวใจโดยมีการกำหนดค่าเพื่อแบ่งระดับความสูงของค่าอาร์ เมื่อค่าใดมีค่ามากกว่าค่าที่กำหนดนี้ จึงจะถือว่ามีความเป็นไปที่จะเป็นค่าอาร์ หาจุดย้อนกลับดังที่กล่าวมาแล้วข้างต้น วนซ้ำจนหมดข้อมูล



รูปที่ 3.2 กระบวนการทำงานของการตัดแบ่งคลื่นไฟฟ้าหัวใจ

3.2 การแปลงข้อมูลให้อยู่ในรูปของสายอักขระ

การแปลงข้อมูลตัวเลขให้เป็นสายอักขระนั้น ทำโดยแบ่งค่าตัวเลขออกเป็นส่วน ๆ เช่นค่า 0 ถึง 50 กำหนดให้ใช้สัญลักษณ์ a และค่า 51 ถึง 60 กำหนดให้ใช้สัญลักษณ์ b เป็นต้น วิธีนี้เรียกว่าการแบ่งนับแบบเส้นตรง (Linear Quantization) ซึ่งเป็นการแบ่งโดยไม่สนใจความสำคัญของข้อมูล อีกทั้งจะทำให้การเปรียบเทียบสายอักขระของสองสายอักขระใด ๆ นั้นไม่ถูกต้องได้ กล่าวคือถึงแม้รูปร่างของสัญญาณคลื่นไฟฟ้าหัวใจจะเหมือนกัน แต่อยู่กันคนละ

ระนาบการเปรียบเทียบก็จะเกิดปัญหาว่าค่าระยะความห่างระหว่างสัญญาณสองสัญญาณห่างกันมากดังแสดงในรูปที่ 3.6 งานวิจัยนี้จึงได้นำ SAX มาใช้ในการแปลงข้อมูลตัวเลขเป็นสายอักขระ ซึ่ง SAX ทำให้ค่าของสัญญาณคลื่นไฟฟ้าหัวใจอยู่ในรูปของค่าการกระจายมาตรฐานก่อน โดยใช้วิธีการแปลงค่าคะแนนซี (Z-Score) ค่าคะแนนซีที่ได้ถูกนำมาสร้างเป็นกราฟใหม่ที่มีค่าการกระจายตัวภายใต้กราฟการกระจายมาตรฐาน และถูกนำไปแบ่งระดับขั้นตามต้องการได้ต่อไป

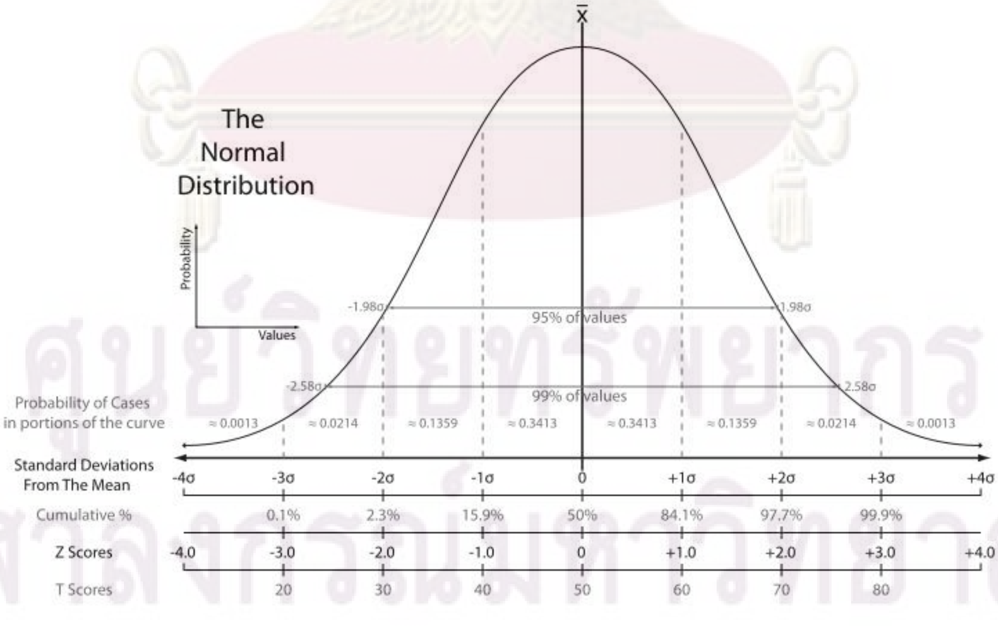
3.2.1 การคำนวณค่าคะแนนซี

การคำนวณค่าคะแนนซีเกิดจากสมการต่อไปนี้

$$Z_i \equiv \frac{x_i - \bar{x}}{\sigma}$$

โดยที่ $\bar{x}$ คือค่าเฉลี่ย และ σ คือค่าเบี่ยงเบนมาตรฐาน

การแปลงค่าคะแนนซีนั้นจะทำให้ช่วงที่มีข้อมูลขึ้นจำนวนมากไปตกอยู่ใกล้กับค่าศูนย์ ซึ่งงานวิจัยนี้ได้ให้ความสำคัญกับข้อมูลกลุ่มนี้เพื่อให้เส้นฐานของสัญญาณคลื่นไฟฟ้าหัวใจในทุก ๆ สัญญาณปรับเข้ามาใกล้กับค่าศูนย์ทำให้การเปรียบเทียบเพื่อหาความสัมพันธ์แม่นยำขึ้น



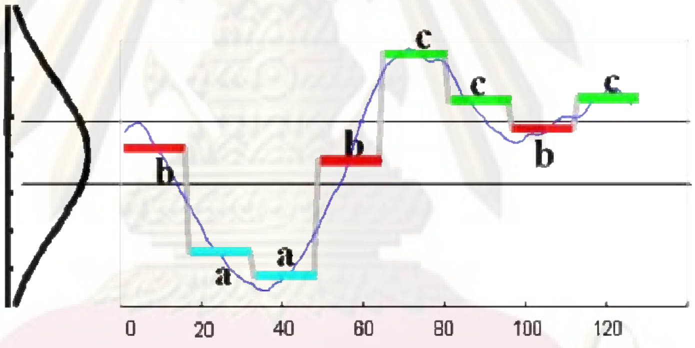
รูปที่ 3.3 ภาพความสัมพันธ์ของการกระจายแบบมาตรฐาน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากรูปที่ 3.3 จะเห็นความสัมพันธ์ของค่าคะแนนซีว่า ค่าส่วนใหญ่ที่เข้าใกล้ค่าเฉลี่ยที่เกิดขึ้นเมื่อผ่านกระบวนการคำนวณหาค่าคะแนนซีจะได้ค่าใหม่ซึ่งเข้าใกล้ศูนย์จึงทำให้เส้นฐานของคลื่นหัวใจถูกปรับให้อยู่ตรงกันหรือใกล้เคียงศูนย์ ประโยชน์อีกข้อของการปรับค่าคะแนนซีคือ จะสามารถเปรียบเทียบจำนวนข้อมูล กับพื้นที่ใต้กราฟได้โดยจุดตัดที่ร้อยละ 95 และร้อยละ 99 ดังรูปที่ 3.3 ซึ่งเป็นจุดที่ข้อมูลส่วนใหญ่ไปรวมตัวกันอยู่และเป็นระนาบแกน y ที่มีค่าสูง

3.2.2 การแบ่งระดับขั้นเพื่อแทนอักขระ

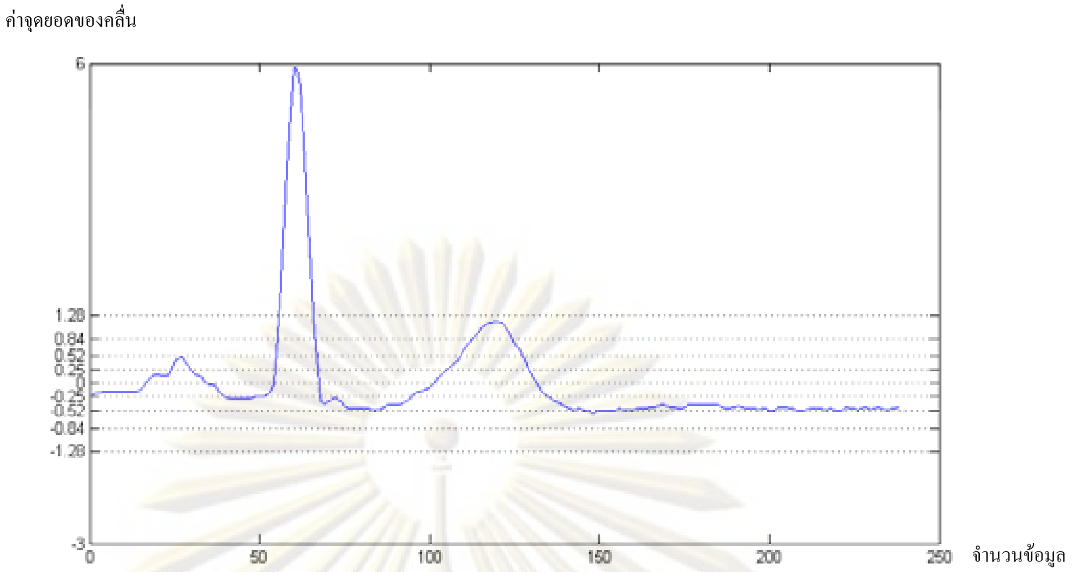
จากคุณสมบัติของการใช้ค่าคะแนนซี ถ้าต้องการแบ่งข้อมูลทั้งหมดออกเป็นสามส่วนเท่า ๆ กันในระนาบแกน x ตรงส่วนค่าใกล้เคียงศูนย์ซึ่งอยู่ตรงกลางของกราฟการกระจายมาตรฐานดังรูปที่ 3.3 จะถูกแบ่งให้แคบมากเพื่อบีบให้ได้ข้อมูลจำนวนหนึ่งในสาม เนื่องจากระนาบแกน y ในจุดที่ใกล้ค่า 0 ในระนาบแกน x จะมีค่าสูงการต้องการจำกัดจำนวนข้อมูลให้ได้ 1 ใน 3 ของพื้นที่ทั้งหมดทำให้จุดแบ่งในระนาบแกน x ต้องแคบมากเพื่อให้เหลือพื้นที่อีก 2 ใน 3 เพื่อจะนำมาแปลงเป็นตัวอักขระ a และ c ดังแสดงด้วยรูปที่ 3.4



รูปที่ 3.4 การปรับค่าอักขระด้วยความสัมพันธ์ของพื้นที่ใต้กราฟการกระจายมาตรฐาน

โดยค่าจุดตัดนั้นขึ้นอยู่กับลำดับชั้นที่ต้องการแปลงเป็นอักขระ และนำไปเปรียบเทียบกับตารางค่าความน่าจะเป็นของพื้นที่ใต้กราฟ ดังตารางที่ 2.1 เช่นถ้าต้องการแบ่งชั้นของอักขระออกเป็น 3 ชั้นคือมีอักขระ {a, b, c} เท่ากับในแต่ละชั้นจะต้องมีพื้นที่ใต้กราฟเท่ากันทั้ง 3 ส่วนคือ 33.33 เปอร์เซ็น ให้หาค่าที่ตำแหน่ง 0.3333 และ 0.6666 ซึ่งในตารางจะพบเพียงค่าบวกโดยที่ค่า 0.6666 เนื่องจากค่า 0.3333 จะได้ค่าเป็นลบจึงหาเพียงค่า 0.6666 จะได้ค่า 0.48 จากตารางที่ 2.1 ซึ่งจะเป็นจุดตัดที่บอกความเปลี่ยนแปลงของอักขระจากระดับชั้นหนึ่งไปอีกชั้นหนึ่ง ในทางกลับกัน -0.48 ก็จะเป็นจุดตัดเปลี่ยนแปลงของอักขระเช่นกันในระดับต่ำลงเนื่องจากเป็นค่าติดลบ

เราสามารถนำมาเขียนเป็นตารางค่าคงที่โดยดูจากตารางที่ 2.1 ที่ระดับขั้นที่ 3 จะมีค่าเท่ากับที่ได้คำนวณไว้ก่อนหน้านี้ คือ 0.48 และ -0.48



รูปที่ 3.5 ตัวอย่างการหาระดับชั้นทั้ง 10 และการแปลงอักษร

โดยได้ตัวอย่างสายอักขระดังแสดงต่อไปนี้

‘eeeeeeeeeeeeeeeeeeeeefggggffeeeeeeeddddddddddddddddegjjjjjjjjjjhfdccdddcccccccc ccccccccdddeeeeeeeeffffgggghhhhiiiiiiiiiiiihhhggfffeeeeeeeeeeeeeeeeeeeeeeeeeeeeee eeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeeee eee’

ต่อไปนี้เป็นการแสดงตัวอย่างการแปลงค่าตัวเลขเป็นอักขระ

ตัวอย่างที่ 3.1 แสดงการแปลงตัวเลขเป็นอักขระด้วยความสัมพันธ์ค่าคะแนนซีและกฎการกระจายมาตรฐาน

กำหนดให้ข้อมูลคลื่น x มีค่าต่อไปนี้

{-20, -20, -20, -20, 20, -20, -20, -20, -20, -20}

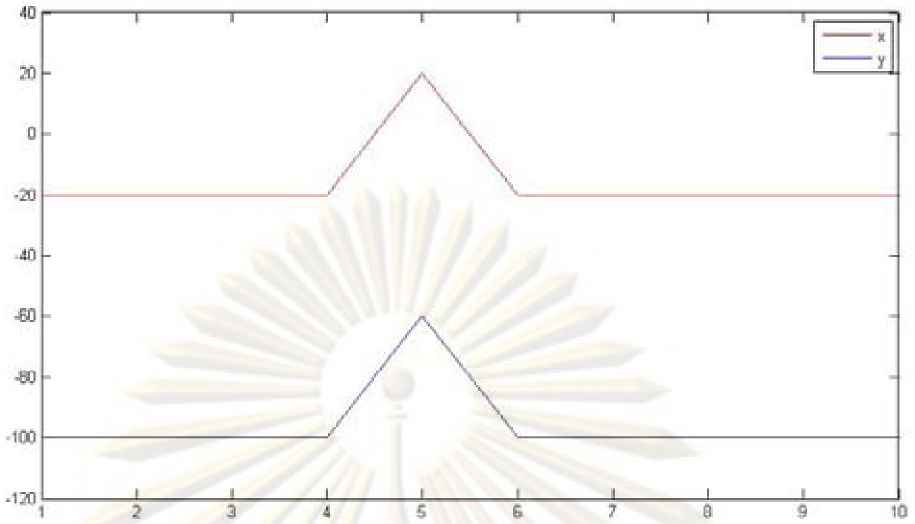
- และกำหนดให้ข้อมูลคลื่น y มีค่าต่อไปนี้

{-100, -100, -100, -100, -60, -100, -100, -100, -100, -100}

ข้อมูล x และ y ที่กำหนดแสดงดังรูปที่ 3.6

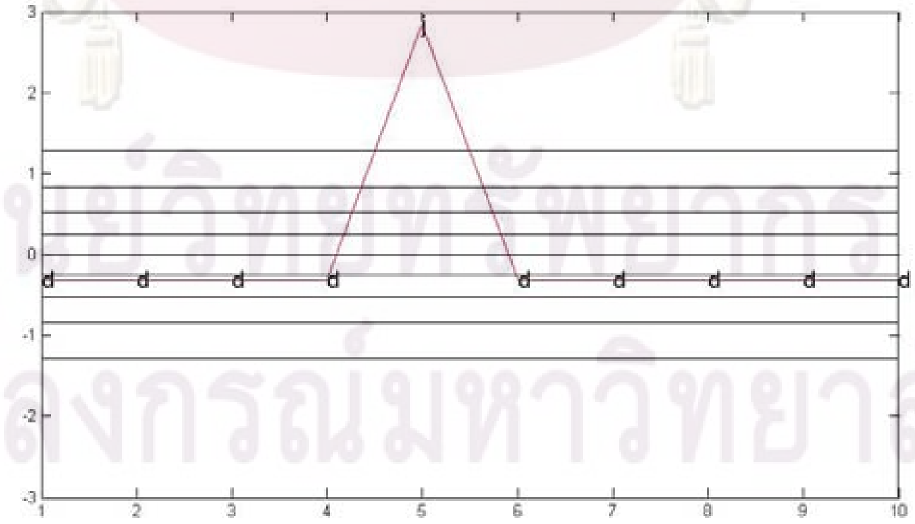
- ค่าคะแนนซีที่ได้ของทั้งสองค่า x และ y คือ

{-0.3162, -0.3162, -0.3162, -0.3162, 2.8460, -0.3162, -0.3162, -0.3162, -0.3162, -0.3162} ดังแสดงในรูปที่ 3.7

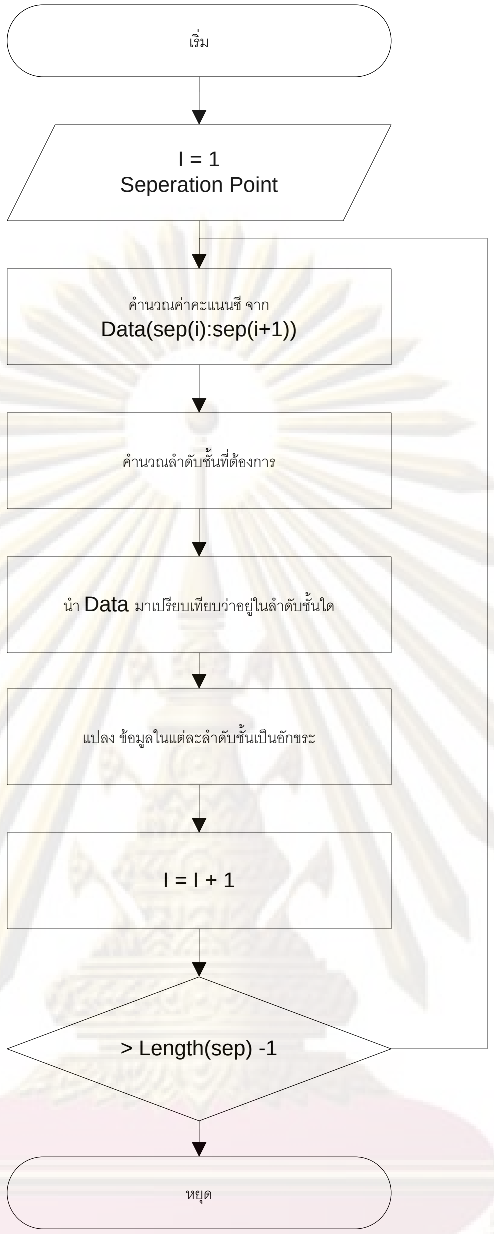


รูปที่ 3.6 ค่าคะแนนดิบ x และ y ที่ไม่ผ่านการหาค่าคะแนนซี

พบว่าถ้านำค่าที่ยังไม่มีการหาค่าคะแนนซีมาเปรียบเทียบกันโดยตรงแล้วภาพทั้งสองจะมีระยะในการเปรียบเทียบที่ห่างกันและถ้าแปลงเป็นอักขระแทนก็จะได้อักขระที่นำมาเปรียบเทียบคนละอักขระกัน โดยที่ถ้าพิจารณากราฟแล้วจะพบว่าคลื่นทั้งสองเหมือนกันดังรูปที่ 3.6 แต่เมื่อผ่านกระบวนการหาค่าคะแนนซี ซึ่งจะทำให้ค่าคะแนนทั้งสองนี้เท่ากัน ซึ่งจะได้อักขระแทนจำนวนที่เหมือนกันมาเปรียบเทียบ เช่นเมื่อนำค่าคะแนนซีมาสร้างเป็นกราฟ และวาดจุดตัดตาม ตารางที่ 2.2 ในที่นี้สมมุติให้ต้องการแบ่งข้อมูลออกเป็น 10 ชั้นในคอลัมภ์สุดท้ายของตารางที่ 2.2 คือจุดตัดแต่ละจุดที่จะใช้ในการแบ่งลำดับขั้น ดังแสดงในรูปที่ 3.7 และกราฟที่ได้ทั้งสองจะซ้อนทับกันพอดีเป็นผลมาจากการหาค่าคะแนนซีของกราฟทั้งสอง



รูปที่ 3.7 กราฟรูปที่ 3.6 เมื่อผ่านการหาคะแนนซี เส้นลำดับขั้นและอักขระแทน



รูปที่ 3.8 กระบวนการแบ่งนับ และแปลงค่าอักขระ

รูปที่ 3.8 แสดงถึงกระบวนการแบ่งนับ และแปลงค่าเป็นสายอักขระแทนตัวเลข โดยนำข้อมูลแต่ละชุดที่ผ่านการหาคาบสัญญาณคลื่นไฟฟ้าหัวใจเพียงหนึ่งรอบมาหาค่าคะแนนซีและ แปลงข้อมูลเป็นสายอักขระแทนด้วยการแบ่งระดับชั้นเป็น 10 ระดับ วนซ้ำจนหมดคาบสัญญาณคลื่นไฟฟ้าหัวใจที่แบ่งมาได้

3.3 การเปรียบเทียบและจัดกลุ่มสายอักขระโดยใช้เครื่องจักรตรวจสอบได้เค

เมื่อได้ข้อมูลของแต่ละคลื่นไฟฟ้าหัวใจที่ถูกตัดแบ่งและนำมาแปลงเป็นข้อมูลสายอักขระเรียบร้อยแล้ว ต่อมาคือการนำคลื่นแต่ละตัวมาเปรียบเทียบซึ่งกันและกันเพื่อหาความเหมือน หรือต่างกันของสายอักขระสองสายโดยใช้วิธีการสร้างเครื่องจักรตรวจสอบได้เค

เพื่อใช้เป็นต้นแบบในการหารูปแบบความเหมือนของสายอักขระ มีจุดเด่นที่สามารถที่จะหาความสัมพันธ์ของอักขระได้โดยใช้เพียงกลุ่มข้อมูลเชิงบวกเพียงอย่างเดียวซึ่งตรงกับข้อมูลที่นำมาใช้ วิธีการสร้างเครื่องจักรตรวจสอบได้แบบเค เมื่อได้เครื่องจักรตรวจสอบได้เค ซึ่งเป็นตัวแทนแต่ละกลุ่มเมื่อมีข้อมูลใหม่เข้ามาเปรียบเทียบก็จะใช้วิธีของการตรวจสอบที่มีการปรับปรุงให้มีความยืดหยุ่นมากยิ่งขึ้นในการตรวจสอบความเป็นสมาชิกในกลุ่ม

3.3.1 การสร้างเครื่องจักรตรวจสอบได้เค

1. กำหนดค่า k ใด ๆ ซึ่งมากกว่า 1
2. กำหนดให้ค่า x เป็นกลุ่มของสายอักขระของข้อมูลตัวอย่างบวก
3. I เป็นตัวเติมหน้า (prefix) ของสายอักขระย่อยจำนวน k – 1 ตัวทั้งหมดของแต่ละสายอักขระใน x
4. F เป็นตัวเติมท้าย (suffix) ของสายอักขระย่อยจำนวน k – 1 ตัวทั้งหมดของแต่ละสายอักขระใน x
5. T เป็นสายอักขระย่อยใด ๆ จำนวน k ตัวทั้งหมดของแต่ละสายอักขระใน x

ตัวอย่างที่ 3.3.1 ให้ X = {‘a’, ‘aa’, ‘abba’, ‘abbbba’} และกำหนดให้ k=3 จะได้เครื่องจักรตรวจสอบแบบเค เป็น

I = {‘Null‘, ‘a’, ‘aa’, ‘ab’}

F = {‘a’, ‘aa’, ‘ba’}

T = {‘abb’, ‘bbb’, ‘bba’}

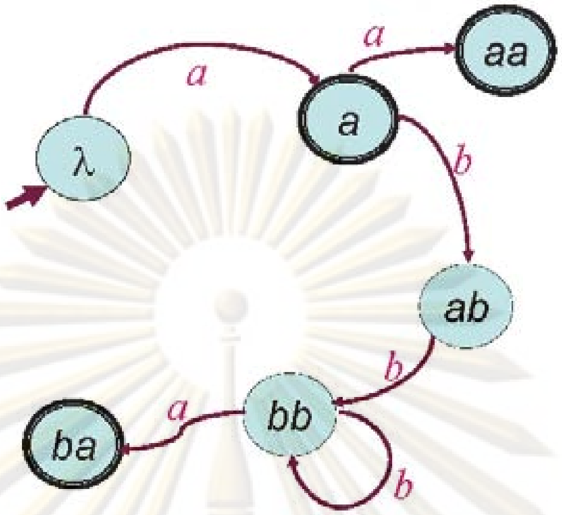
จากนั้นนำไปสร้างเป็นเครื่องจักรจำกัดสถานะเพื่อสามารถใช้เครื่องจักรจำกัดสถานะในการตรวจสอบความเป็นภาษาได้

ตัวอย่างที่ 3.3.2 I , F และ T ที่ได้จากตัวอย่างที่ 3.3.1 แปลงเป็นเครื่องจักรจำกัดสถานะ

- แต่ละสายอักขระใน I เป็นสถานะ
- แต่ละสายอักขระที่มีความยาวเท่ากับ k-1 ใน T เป็นสถานะ
- ค่าว่าง (Null) เป็นสถานะเริ่มต้น
- เพิ่มการเปลี่ยนสถานะโดยมีป้าย b จาก u ไปที่ ub ในแต่ละสายอักขระใน I
- เพิ่มการเปลี่ยนสถานะโดยมีป้าย b จาก au ไปที่ ub ในแต่ละสายอักขระใน T
- แต่ละสายอักขระใน F เป็นสถานะสุดท้าย หรือสถานะยอมรับ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ดังแสดงในรูปที่ 3.9 เป็นผลลัพธ์ของการสร้างเครื่องจักรจำกัดสถานะ โดยใช้ข้อมูล I , F และ T จากตัวอย่างที่ 3.3.1



รูปที่ 3.9 เครื่องจักรจำกัดสถานะ

จากวิธีการสร้างเครื่องจักรจำกัดสถานะนั้น พบว่ามีค่า k ใด ๆ เข้ามาเกี่ยวข้องซึ่งค่า k นี้เป็นตัวจำกัดขอบเขตในการมองสายอักขระไปทีละ k จำนวน ยิ่งค่า k มีน้อยจำนวนสายอักขระที่จะนำมาหาความสัมพันธ์ก็ยิ่งสั้นและสถานะในเครื่องจักรตรวจสอบจะมีจำนวนมากมีผลทำให้เครื่องจักรตรวจสอบได้เคที่สร้างขึ้นมีความยืดหยุ่นมาก กล่าวคือจะทำให้การเปลี่ยนสถานะไปสู่สถานะยอมรับของเครื่องจักรเป็นไปได้ง่าย ในทางกลับกันค่า k ที่มากสถานะของเครื่องจักรจะมีจำนวนน้อยลง เมื่อสถานะน้อยก็จะทำให้การเปลี่ยนสถานะไปในสถานะถัดไปมีความเป็นไปได้น้อยลง และทำให้เครื่องจักรไม่ยอมรับสายอักขระเข้าเป็นสมาชิกและเนื่องจากในการสร้างเครื่องจักรตรวจสอบได้เคนั้น ต้องนำกลุ่มข้อมูลที่เป็นสมาชิก หรือเป็นกลุ่มเดียวกันมาเพื่อสร้างเป็นเครื่องจักรตรวจสอบได้เคแต่ในงานวิจัยนี้ข้อมูลที่มีเป็นข้อมูลที่ผสมรวมกันอยู่ ซึ่งไม่สามารถที่จะนำข้อมูลทั้งหมดมาสร้างเครื่องจักรตรวจสอบได้เค เพราะจะทำให้เครื่องจักรที่ได้ยอมรับสัญญาณคลื่นไฟฟ้าหัวใจที่มีรูปแบบผิดปกติและไม่ผิดปกติซึ่งทำให้ไม่สามารถแบ่งกลุ่มสัญญาณได้อย่างถูกต้อง งานวิจัยนี้จึงต้องมีการเรียนรู้และสร้างเครื่องจักรตรวจสอบได้เคในขณะเดียวกันก็จะจัดกลุ่มร่วมไปด้วยกัน

ด้วยการปรับเปลี่ยนค่าเค จะทำให้เกิดผลลัพธ์ของการจัดกลุ่มเป็น

- การจัดกลุ่มได้ผลลัพธ์ที่มีสัญญาณคลื่นหัวใจที่แตกต่างกันบางส่วนถูกนำมารวม กันเป็นกลุ่ม ๆ เดียวกันแต่ได้ความรวดเร็วในการจัดกลุ่มเนื่องจากจำนวนกลุ่มให้เปรียบเทียบมีน้อย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

- การจัดกลุ่มได้ผลลัพธ์ที่เกิดกลุ่มมากซึ่งหลายกลุ่มอาจเป็นสัญญาณคลื่นหัวใจชนิดเดียวกัน แต่ไม่ถูกนำมารวมกับกลุ่มได้ อีกทั้งยังทำให้ความรวดเร็วในการจัดกลุ่มตกลง เนื่องจากมีกลุ่มที่ต้องนำมาเปรียบเทียบ มากขึ้น

จากผลลัพธ์ที่มีสัญญาณคลื่นไฟฟ้าหัวใจที่แตกต่างกันถูกนำมารวมกันเป็นกลุ่มเดียวกันนั้นงานวิจัยนี้จะไม่นำมาใช้ เนื่องจากผลลัพธ์ที่ได้ไม่สามารถที่จะนำมาใช้จริงได้ ซึ่งอาจทำให้เกิดการวินิจฉัยผิดพลาดซึ่งส่งผลให้การวินิจฉัยของแพทย์ผิดพลาด หลักการทำงานทั่วไปของการจัดกลุ่มมีดังนี้

3.3.2 การเปรียบสัญญาณคลื่นไฟฟ้าหัวใจและจัดกลุ่ม

1. สัญญาณคลื่นไฟฟ้าหัวใจแรกจะถูกจัดให้อยู่ในกลุ่มที่ 1

2. สัญญาณคลื่นไฟฟ้าหัวใจแรกที่ได้จากข้อ 1 จะถูกนำไปสร้างเครื่องจักรตรวจสอบได้เค ซึ่งจะใช้เป็นตัวแทนของกลุ่มที่ 1

3. สัญญาณคลื่นไฟฟ้าหัวใจลำดับถัดมาจะถูกเปรียบเทียบกับเครื่องจักรที่ได้ในข้อ 2 และถ้าเครื่องจักรยอมรับความเป็นสมาชิก ข้อมูลชุดนี้จะรวมเข้ากับกลุ่มที่ 1 ด้วยและนำไปสร้างเป็นเครื่องจักรตรวจสอบได้เค

4. ในกรณีสัญญาณคลื่นไฟฟ้าหัวใจไม่ถูกยอมรับความเป็นสมาชิกในกลุ่มที่ 1 จะถูกนำไปสร้างเป็นเครื่องจักรจำกัดสถานะตัวใหม่โดยเป็นตัวแทนของกลุ่มที่ 2 ต่อไป

5. สัญญาณคลื่นไฟฟ้าหัวใจในลำดับถัดมาจะถูกนำมาแจงส่วน (Parse) ด้วยเครื่องจักรตรวจสอบได้เคในแต่ละกลุ่มที่เกิดขึ้น ถ้าเครื่องจักรทุกกลุ่มไม่ยอมรับสัญญาณคลื่นไฟฟ้าหัวใจนั้นจะถูกจัดเป็นกลุ่มใหม่

6. ในกรณีที่ระหว่างการแจงส่วนไม่พบฟังก์ชันการผ่านไปที่สถานะใด ระบบจะให้ความหมายว่าสัญญาณคลื่นไฟฟ้าหัวใจที่นำมาตรวจสอบนั้นมีรูปแบบที่ไม่เหมือนกันโดยสิ้นเชิงจะให้สิ้นสุดการแจงส่วน และถือเป็นข้อมูลที่ไม่ยอมรับเข้าเป็นสมาชิก

7. ในกรณีที่การแจงส่วนสมบูรณ์และเกิดการยอมรับเข้ากลุ่มแล้วนั้น ระบบจะสร้างเครื่องจักรจำกัดสถานะขึ้นใหม่โดยรวมเอาสัญญาณคลื่นไฟฟ้าหัวใจตัวใหม่เข้ามาใช้ในการสร้างด้วย

8. เมื่อการแจงส่วนเสร็จสมบูรณ์ครบตามจำนวนสัญญาณคลื่นไฟฟ้าหัวใจที่ตัดแบ่งแล้ว ขั้นต่อไปจะตรวจสอบจำนวนสมาชิกสัญญาณคลื่นไฟฟ้าหัวใจแต่ละกลุ่ม โดยถ้ากลุ่มใดมีจำนวนสมาชิกน้อยมาก จะถูกนำมาแจงส่วนอีกครั้งกับกลุ่มที่มีจำนวนสมาชิกมาก ในกรณีนี้จะช่วยให้กลุ่มที่มีสมาชิกน้อยสามารถรวมเข้าไปอยู่ในกลุ่มจำนวนสมาชิกมากได้ เนื่องจากกลุ่มที่มีสมาชิกมากจะมีเครื่องจักรจำกัดสถานะที่สมบูรณ์ซึ่งเป็นไปตามจำนวนสมาชิกที่นำมาสร้างเครื่องจักรจำกัดสถานะ

3.3.2.1 การปรับปรุงความยืดหยุ่นในการตรวจสอบความเป็นสมาชิก

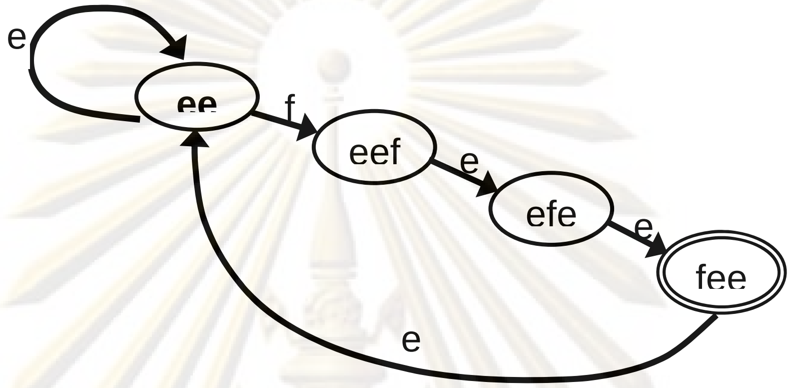
เมื่อสร้างเครื่องจักรตรวจสอบได้เคแล้ว การนำข้อมูลใหม่มาเปรียบเทียบในแบบวิธีปกติทำให้ข้อมูลที่มีการเปลี่ยนแปลงตลอดเวลาหรือมีความละเอียดสูงถูกจัดเป็นกลุ่มที่แตกต่างกันเกือบทั้งหมดเพราะความผิดพลาดเพียงแค่ลำดับของตัวอักขระที่ไม่ตรงกันในการตรวจสอบ การเปลี่ยนแปลงของอักขระที่ไม่เหมือนกัน หรือจำนวนอักขระทั้งหมดของสายอักขระหนึ่งมีความยาวที่ไม่เท่ากันเหล่านี้เป็นผลทำให้เครื่องจักรไม่ยอมรับเข้าเป็นสมาชิกของกลุ่มได้ทั้งหมด งานวิจัยนี้จึงมีการเปลี่ยนแปลงระบบการตรวจสอบความเป็นสมาชิกโดยกำหนดกฎที่ระดับของอักขระที่อยู่ใกล้เคียงกันจะยอมให้มีการผ่านไปทางฟังก์ชันการผ่านได้ ยกตัวอย่างเช่น สถานะลำดับที่ 1 จะผ่านไปสถานะลำดับที่ 2 ได้โดยใช้อักขระ c เป็นตัวส่งผ่าน และข้อมูลที่ได้จากข้อมูลนำเข้าเปรียบเทียบเป็นอักขระ b เครื่องจักรนี้ยังยอมให้ผ่านไปที่สถานะลำดับที่ 2 ได้ทางฟังก์ชันการผ่าน c การทำแบบนี้จะช่วยให้ระบบมีความยืดหยุ่นสูงขึ้น และทนทานต่อสัญญาณรบกวนมากขึ้นอีกด้วย

3.3.3 ปัญหาการสร้างเครื่องจักรตรวจสอบได้เคที่วนซ้ำได้

จากข้อมูลสัญญาณคลื่นไฟฟ้าหัวใจที่ถูกแปลงเป็นสายอักขระแทนจำนวนแล้ว จะพบว่าส่วนใหญ่ค่าที่ส่วนหัว และหางนั้นจะมีค่าซ้ำกันอยู่เป็นเส้นตรงและจะทำให้ได้อักขระแทนเป็นตัวอักขระเดียวกัน

ตัวอย่างสายอักขระที่ได้จากการแปลงโดยใช้อักขระแทนจำนวนของสัญญาณคลื่นไฟฟ้าหัวใจ ‘eeeeeffgedddjjjddddddddddeeeffgghhiiggffddddddddeeeeeeeeeeee’ พบว่าส่วนหัวมีอักขระแทน ‘e’ จำนวนห้าตัวติดกันและส่วนหางก็จะเป็น ‘e’ เช่นเดียวซึ่งในการสร้างเครื่องจักรตรวจสอบได้เคนั้น ถ้ากำหนดค่า k ให้มีค่าน้อยกว่าหรือเท่ากับ 5 ในกรณีของข้อมูลชุดนี้จะทำให้เกิดเครื่องจักร ที่มีการวนซ้ำไม่จำกัดจำนวน โดยในสถานะสุดท้ายหรือส่วนหางของข้อมูลจะวกกลับมาเริ่มต้นใหม่ที่สถานะเริ่มต้นที่เป็น ‘eeeee’ ได้ ซึ่งมีผลทำให้การเกิดสัญญาณคลื่นไฟฟ้าหัวใจหนึ่งลูกคลื่นมีค่าเท่ากับสัญญาณคลื่นไฟฟ้าหัวใจหลายลูกคลื่นและทำให้สัญญาณทั้งสองชนิดนี้รวมอยู่ในกลุ่มเดียวกันได้ดังตัวอย่างแสดงในรูปที่ 3.10 ที่ถ้ามีข้อมูล

สัญญาณคลื่นไฟฟ้าหัวใจที่มีอักขระแทนเป็น 'eeeeeeefeeeeee' โดยกำหนดให้ 'f' เป็นส่วนกลางข้อมูลสัญญาณคลื่นไฟฟ้าหัวใจทั้งหมดและสร้างเครื่องจักรตรวจสอบได้เคที่ค่า k เป็น 3 ดังรูปที่ 3.10 และถ้าใช้ข้อมูลตรวจสอบความเป็นสมาชิกในกลุ่มเป็น 'eeeeeeefeeeeeeeeeeeeefeeeeeee' เครื่องจักรตรวจสอบได้เคดังรูปที่ 3.10 ก็ยังคงตรวจสอบได้ว่าเป็นสมาชิกในกลุ่มเดียวกันกับข้อมูลที่นำมาสร้างเครื่องจักรตรวจสอบได้เคนี้ที่มีแบบรูป 'f' เพียงตัวเดียว



รูปที่ 3.10 ปัญหาการเกิดการวนซ้ำของเครื่องจักรตรวจสอบได้เค

จากปัญหาการวนซ้ำนี้ จึงแก้ปัญหาด้วยการแบ่งการสร้างเครื่องจักรตรวจสอบได้แบบเคออกเป็น 2 ส่วนหน้าหลังโดยใช้จุด R ที่หาได้จากขั้นตอนการตัดแบ่งแรกนั้นเป็นจุดตัด โดยส่วนหน้าจะนำไปสร้างเครื่องจักรตรวจสอบได้เคเพื่อใช้หาความสัมพันธ์กับชุดข้อมูลส่วนหน้าที่ถูกตัดที่จุด R ด้วยเช่นกันและในส่วนท้ายก็เช่นกันใช้เปรียบเทียบหาความสัมพันธ์กับชุดข้อมูลส่วนท้าย ถ้าทั้งสองส่วนเกิดการยอมรับด้วยเครื่องจักรตรวจสอบได้เค ก็จะถือว่าสัญญาณคลื่นไฟฟ้าหัวใจที่นำมาเปรียบเทียบเป็นข้อมูลในกลุ่มเดียวกันและในทางตรงกันข้ามถ้ามีส่วนใดส่วนหนึ่งไม่ถูกยอมรับก็จะให้ถือว่าเป็นสัญญาณคลื่นไฟฟ้าหัวใจที่อยู่ต่างกลุ่มกัน

บทที่ 4

การทดลอง

จากที่กล่าวมานั้น การสร้างไวยากรณ์ภาษานั้นสามารถที่จะนำมาเปรียบเทียบหาความเหมือนของสายอักขระใด ๆ ได้ ด้วยเครื่องจักรจำกัดสถานะที่สร้างขึ้นจากข้อมูลสายอักขระของกลุ่มข้อมูลที่ให้ได้ แต่ต้องแบ่งว่าสายอักขระนั้นเป็นข้อมูลของกลุ่มใดเพื่อสร้างไวยากรณ์ภาษาที่ถูกต้องของกลุ่มนั้น ๆ ซึ่งเป็นจุดด้อยที่ว่า ข้อมูลสายอักขระที่ได้มานั้นผสมรวมกันอยู่ระหว่างข้อมูลหลาย ๆ กลุ่มและยังไม่สามารถแบ่งได้ว่าข้อมูลสายอักขระนี้อยู่ในกลุ่มใดซึ่งนั่นคือเป้าหมายของงานวิจัยนี้

โดยในบทนี้จะกล่าวถึงวิธีการทดลอง เครื่องมือ และข้อมูลที่ใช้ในการทดลอง และการทดลองเพื่อหาผลลัพธ์ที่ได้จากวิธีการที่ได้กล่าวมา

4.1 วัตถุประสงค์การทดลอง

เพื่อให้ทราบถึงความถูกต้องของการแบ่งกลุ่มจากข้อมูลทั้งหมด ในแต่ละค่า k ใด ๆ และเพื่อให้ทราบถึงค่า k ที่ดีที่สุดในการสร้างเครื่องจักรจำกัดสถานะ รวมถึงเวลาในการประมวลผลจนกระทั่งเสร็จสิ้นกระบวนการ

4.2 เครื่องมือที่ใช้ในการทดลอง

เครื่องมือที่ใช้ในการทดลองคือโปรแกรมที่พัฒนาขึ้นเองด้วย Matlab

4.2.1 โปรแกรมที่พัฒนาขึ้นเองในการตัดแบ่งภาพคลื่นไฟฟ้าหัวใจออกเป็นส่วน ๆ

4.2.2 โปรแกรมที่พัฒนาขึ้นเองในการแปลงข้อมูลตัวเลขเป็นข้อมูลสายอักขระโดยใช้หลักการทางสถิติ

4.2.3 โปรแกรมที่พัฒนาขึ้นเองเพื่อการสร้างเครื่องจักรจำกัดสถานะ และใช้ในการเปรียบเทียบความแตกต่างระหว่างสายอักขระสองสาย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.3 ข้อมูลที่ใช้ในการทดลอง

เป็นข้อมูลจริงที่ได้มาโดยตรงจากแพทย์ โดยงานวิจัยนี้มีอยู่ด้วยกันสองประเภท

4.3.1 ข้อมูลที่ได้มาโดยตรงจากผู้ช่วยศาสตราจารย์ นายแพทย์ ทายาท ดีสุดจิต โรงพยาบาลจุฬาลงกรณ์ ซึ่งเป็นข้อมูลของคนไข้ซึ่งมีความเป็นไปได้ที่จะมีความผิดปกติ

4.3.2 ข้อมูลที่ได้จากเว็บไซต์สำหรับนักวิจัยที่เกี่ยวกับภาพคลื่นไฟฟ้า ที่สามารถตรวจจับได้จากมนุษย์ในปัจจุบัน http://www.physionet.org/ ซึ่งมีทั้งที่ไม่มีความผิดปกติ และมีความผิดปกติในด้านของรูปร่างอย่างเห็นได้ชัด

ซึ่งการทดลองจะใช้ภาพคลื่นไฟฟ้าหัวใจที่มีความยาวโดยประมาณ 5 – 8 ชม. ในการวัดผล

4.4 ขั้นตอนการทดลอง

4.4.1 นำข้อมูลเข้า และเริ่มการประมวลผลตัดแบ่งภาพคลื่นไฟฟ้าหัวใจ

4.4.2 นำแต่ละภาพของภาพคลื่นไฟฟ้าหัวใจเข้ากระบวนการหาค่าคะแนนซี

4.4.3 แปลงค่าตัวเลขหลังจากผ่านกระบวนการหาค่าคะแนนซีเป็นอักขระแทนค่า

4.4.4 นำภาพคลื่นไฟฟ้าหัวใจที่อยู่ในรูปสายอักขระ มาสร้างเครื่องจักรจำกัดสถานะแทนกลุ่มของตัวเอง เพื่อใช้ในการเปรียบเทียบกับภาพคลื่นไฟฟ้าหัวใจต่อไปไว้เปรียบเทียบ

4.4.5 พิจารณาจำนวนกลุ่มที่ได้ และจำนวนสมาชิกในกลุ่ม

4.4.6 พิจารณาสมาชิกในกลุ่มที่มีรูปแบบแตกต่างจากกลุ่ม เก็บค่าความผิดพลาด

4.4.7 เปลี่ยนค่า k และเริ่มทำกระบวนการที่ 4.4.4 – 4.4.6 อีกครั้ง

4.4.8 โดยในการทดลอง ใช้ค่าเคที่ 2 – 4 ในการทดลอง

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

ผลการทดลอง

ในบทนี้จะกล่าวถึงผลการทดลองจำแนกกลุ่มของภาพคลื่นไฟฟ้าหัวใจ ที่ถูกจำแนกโดยออโตมาตา โดยปกติแล้วหลักการในการจำแนกความแตกต่างระหว่างภาพคลื่นไฟฟ้าหัวใจแบ่งได้ออกเป็น 3 ส่วนหลักคือ

1. ที่ความยาวระหว่างคลื่นต่อคลื่นถัดไปโดยวัดจาก จุดอาร์จนถึงจุดอาร์ถัดไปซึ่งเรียกว่า ช่วงอาร์อาร์ (R-R Interval) ถ้าความยาวต่างกันที่ช่วงค่าเบี่ยงเบนมาตรฐานหนึ่งแพทย์จะถือว่ามีความผิดปกติของภาพคลื่นหัวใจนี้ที่การเต้นของหัวใจช้าลง

2. ลักษณะทั่วไปเช่น คลื่นที่มีลูกคลื่นผิดแบบ หรือช่วงอาร์ทิ่มลงในขณะที่ภาพคลื่นอื่น ๆ ตั้งขึ้นในแนวปกติ ซึ่งเป็นความผิดปกติที่จังหวะการเต้นของหัวใจมีความผิดปกติ

3. กรณีภาพคลื่นไฟฟ้าหัวใจขาดหายไปเป็นช่วง ๆ อาจเกิดขึ้นจากกล้ามเนื้อหัวใจบางส่วนตาย ซึ่งทำให้คลื่นไฟฟ้าไม่สามารถไหลเวียนไปได้ทำให้จังหวะการเต้นของหัวใจหยุดชะงัก

5.1 การทดลองด้วยภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติ

จากการทดลองกับภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติจำนวน 3 คน ดังตารางที่ 5.1 ตามลำดับ ซึ่งพิจารณาตามค่าเค ซึ่งเท่ากับ 2, 3 และ 4 ตามลำดับ ได้ผลการทดลองดังตารางที่ 5.2, 5.3 และ 5.4 ตามลำดับ

ตารางที่ 5.1 ข้อมูลของภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติ

คนที่	จำนวนภาพคลื่นไฟฟ้าที่แบ่งได้	ความยาวของชุดภาพคลื่นไฟฟ้าหัวใจ (เวลา)
1	24249	5 – 8 ชม.
2	24532	5 – 8 ชม.
3	26782	5 – 8 ชม.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.2 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติที่ค่า k เท่ากับ 2

จำนวนกลุ่มที่ได้	สมาชิกในกลุ่มมากที่สุด	% ของกลุ่มมากที่สุด
9	20113	82.94%
9	19103	71.32%
8	15664	63.85%

ตารางที่ 5.3 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติที่ค่า k เท่ากับ 3

จำนวนกลุ่มที่ได้	สมาชิกในกลุ่มมากที่สุด	% ของกลุ่มมากที่สุด
15	20838	85.93%
14	21301	86.83%
20	24547	91.65%

ตารางที่ 5.4 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่ไม่พบความผิดปกติที่ค่า k เท่ากับ 4

จำนวนกลุ่มที่ได้	สมาชิกในกลุ่มมากที่สุด	% ของกลุ่มมากที่สุด
38	11511	47.47%
23	20226	75.52%
28	15572	63.48%

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

จากทั้ง 3 ตารางจะเห็นว่าเมื่อค่าเคเพิ่มขึ้นมีผลทำให้จำนวนกลุ่มมากขึ้นตามไปด้วย ทั้งนี้เพราะค่าเคที่สูงได้แสดงให้เห็นถึงความละเอียดในการพิจารณา ผลจากการทดลองค่าเคที่เหมาะสมที่เลือกใช้ในงานวิจัยนี้คือ 3 ทั้งนี้เพราะกลุ่มที่มีจำนวนมากที่สุดเมื่อพิจารณาค่าเคเท่ากับ 3 จะครอบคลุมจำนวนภาพคลื่นไฟฟ้าหัวใจมากที่สุดจากการทดลอง

5.2 การทดลองด้วยภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติ

ทำนองเดียวกัน ภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติ 3 คน ดังตารางที่ 5.5 ทดลองจำแนกกลุ่มโดยพิจารณาค่าเค 2, 3 และ 4 ตามลำดับ ได้ผลตามตามตารางที่ 5.6, 5.7 และ 5.8 ตามลำดับ

ตารางที่ 5.5 ข้อมูลของชุดภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติ

ชุดที่	จำนวนภาพคลื่นไฟฟ้าที่แบ่งได้	ความยาวของชุดภาพคลื่นไฟฟ้าหัวใจ (เวลา)
1	22512	5 – 8 ชม.
2	20446	5 – 8 ชม.
3	29721	5 – 8 ชม.

ตารางที่ 5.6 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติที่ค่า k เท่ากับ 2

จำนวนกลุ่มที่ได้	จำนวนสมาชิกในกลุ่มที่มีสมาชิกมากที่สุด	% สมาชิก ในกลุ่มที่มีสมาชิกมากที่สุด	จำนวนสมาชิกในกลุ่มที่มีรูปแบบแตกต่าง	% สมาชิกในกลุ่มที่มีสมาชิกที่มีรูปแบบแตกต่าง
18	19826	88.06%	415	1.84%
8	19351	94.00%	0	0%
11	24521	82.25%	0	0%

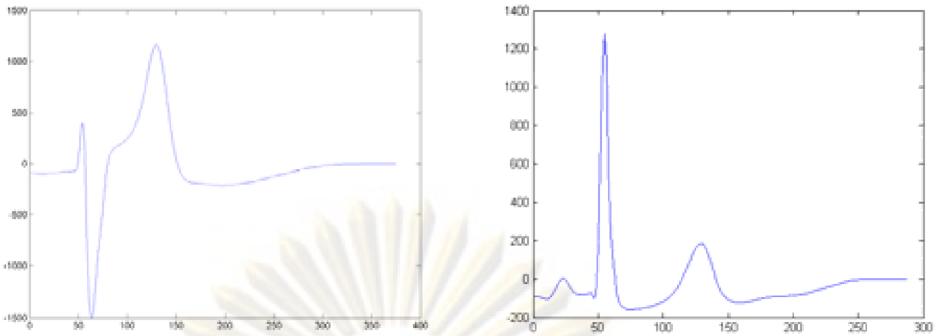
ตารางที่ 5.7 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติที่ค่า k เท่ากับ 3

จำนวนกลุ่มที่ได้	จำนวนสมาชิกในกลุ่มที่มีสมาชิกมากที่สุด	% สมาชิก ในกลุ่มที่มีสมาชิกมากที่สุด	จำนวนสมาชิกในกลุ่มที่มีรูปแบบแตกต่าง	% สมาชิกในกลุ่มที่มีสมาชิกที่มีรูปแบบแตกต่าง
25	16546	73.50%	742	3.30%
22	14331	70.09%	62	0.30%
25	19522	65.48%	35	0.12%

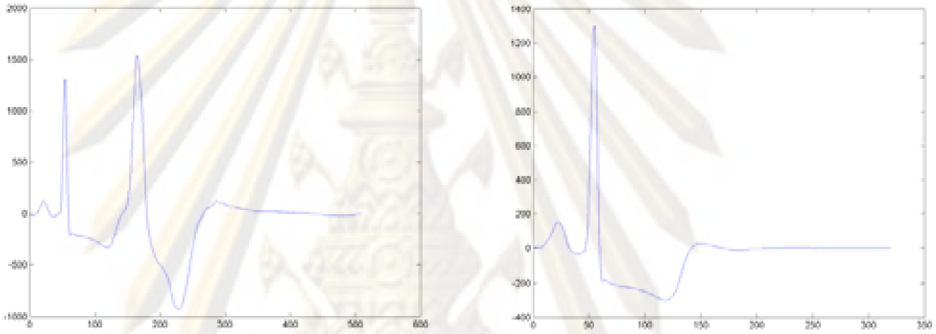
ตารางที่ 5.8 ผลการจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจของผู้ป่วยที่พบความผิดปกติที่ค่า k เท่ากับ 4

จำนวนกลุ่มที่ได้	จำนวนสมาชิกในกลุ่มที่มีสมาชิกมากที่สุด	% สมาชิก ในกลุ่มที่มีสมาชิกมากที่สุด	จำนวนสมาชิกในกลุ่มที่มีรูปแบบแตกต่าง	% สมาชิกในกลุ่มที่มีสมาชิกที่มีรูปแบบแตกต่าง
34	11098	49.29%	684	3.04%
39	13658	66.34%	70	0.34%
26	12493	41.90%	35	0.12%

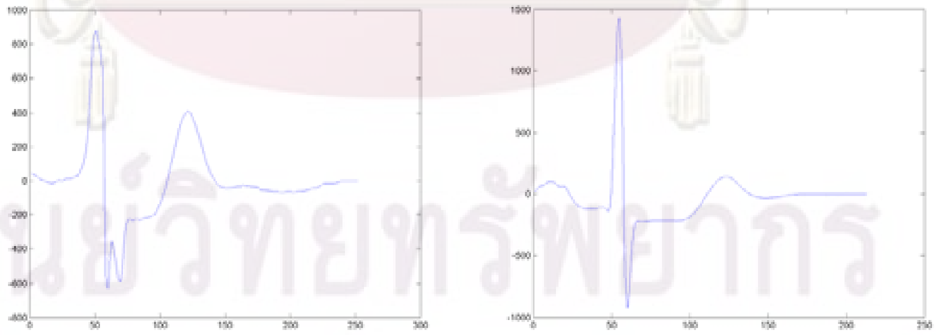
จากตารางทั้ง 3 ก็เช่นกันจะเห็นว่าเมื่อค่าเคเพิ่มขึ้นมีผลทำให้จำนวนกลุ่มมากขึ้นตามไปด้วย ถ้าพิจารณาจากผลการทดลองที่จำนวนสมาชิกในกลุ่มที่มีรูปแบบแตกต่างของค่าเคเท่ากับ 2 ในตารางที่ 5.6 จะพบว่าจำนวนสมาชิกในผู้ป่วยคนที่ 2 และ 3 ไม่พบกลุ่มของภาพที่มีรูปแบบแตกต่างเลย ในขณะที่ผลการทดลองที่ค่าเคเท่ากับ 3 และ 4 มีการตรวจพบ ซึ่งหมายความว่าที่ค่าเคเท่ากับ 2 นั้นเป็นค่าที่หลวมเกินไปจนทำให้ภาพที่มีรูปแบบแตกต่างรวมเข้าไปกับกลุ่มปกติจนไม่ตรวจพบกลุ่มรูปแบบแตกต่างนั้น



รูปที่ 5.1 (ซ้าย) ภาพคลื่นไฟฟ้าหัวใจจากชุดที่ 1 ที่ผิดปกติ (ขวา) ที่ปกติ



รูปที่ 5.2 (ซ้าย) ภาพคลื่นไฟฟ้าหัวใจจากชุดที่ 2 ที่ผิดปกติ (ขวา) ที่ปกติ



รูปที่ 5.3 (ซ้าย) ภาพคลื่นไฟฟ้าหัวใจจากชุดที่ 3 ที่ผิดปกติ (ขวา) ที่ปกติ

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 5.9 จำนวนภาพคลื่นไฟฟ้าหัวใจที่ผิดพลาดในแต่ละกลุ่ม ที่พบรูปแบบที่แตกต่างปะปนกับภาพคลื่นไฟฟ้าหัวใจรูปแบบปกติ

ชุด	ภาพคลื่นไฟฟ้าหัวใจกลุ่มรูปแบบปกติ									รูปแบบผิดปกติที่มากที่สุด
	มากที่สุด	2	3	4	5	6	7	8	9	
1	0	0	0	0	3	3	0	0	12	0
2	1	0	0	108	22	0	0	0	0	5
3	0	0	0	0	0	0	13	10	0	3

ตารางที่ 5.9 แสดงจำนวนภาพคลื่นไฟฟ้าหัวใจที่จำแนกผิด โดยพิจารณากลุ่มของภาพคลื่นไฟฟ้าหัวใจที่ปกติกลุ่มที่มากที่สุด 10 อันดับแรกซึ่งมีปริมาณภาพคลื่นไฟฟ้าหัวใจคิดเป็น 95% ของทั้งหมด และกลุ่มของภาพคลื่นไฟฟ้าหัวใจที่ผิดปกติที่มากที่สุด จะเห็นว่าในกลุ่มภาพคลื่นไฟฟ้าหัวใจที่มากที่สุดของชุดที่ 2 นั้นมีความผิดพลาดคิดเป็นเพียง 0.007% จากข้อมูลในกลุ่มที่มากที่สุด 14331 ซึ่งยังเป็นความผิดพลาดที่ยอมรับได้ถ้ายังไม่มากจนเกิดกรณีที่ภาพคลื่นไฟฟ้าหัวใจที่มีรูปแบบหนึ่ง รูปแบบใดหายไปทั้งกลุ่ม ในขณะที่ชุด 1 และ 3 ไม่มี ดังที่แสดงในตารางที่ 5.10 และจากตารางที่ 5.9 ข้อมูลชุดที่ 2 ในกลุ่มที่มีปริมาณมากสุดอันดับ 4 พบความผิดพลาดมากที่สุด 108 ภาพคลื่นไฟฟ้าหัวใจนั้น เนื่องจากภาพคลื่นไฟฟ้าหัวใจที่ปกติและผิดปกติมีความคล้ายคลึงกันมากดังแสดงในรูปที่ 5.2 ทำให้เกิดภาพคลื่นไฟฟ้าหัวใจที่ผิดปกติปะปนเข้าไป และเกิดการเรียนรู้ในการสร้างออโตมาตาที่ผิดทำให้เกิดออโตมาตาที่ยอมรับภาพคลื่นไฟฟ้าหัวใจที่ทั้งสองรูปแบบ ตารางที่ 5.10 นำค่าความผิดพลาด และเปอร์เซ็นความผิดพลาดแสดงร่วมกัน

ตารางที่ 5.10 เปอร์เซ็นต์ความผิดพลาดที่เกิดขึ้นจากตารางที่ 5.9

		กลุ่มที่รูปร่างปกติที่มากที่สุด	2	3	4	5	6	7	8	9	กลุ่มที่รูปร่างผิดแปลกที่มากที่สุด
ชุดที่ 1	สมาชิกในกลุ่ม	16,546	292	419	1,755	281	629	364	121	274	742
	ความผิดพลาดที่ระบบจัดกลุ่ม	-	-	-	-	3	3	-	-	12	-
	% ความผิดพลาด	-	-	-	-	1.0676	0.4769	-	-	4.3796	-
ชุดที่ 2	สมาชิกในกลุ่ม	14,331	1,822	453	1,763	447	129	153	147	173	62
	ความผิดพลาดที่ระบบจัดกลุ่ม	1			108	22					5
	% ความผิดพลาด	0.0070	-	-	6.1259	4.9217	-	-	-	-	8.0645
ชุดที่ 3	สมาชิกในกลุ่ม	19,503	4,471	442	1,136	1,368	330	815	323	252	34
	ความผิดพลาดที่ระบบจัดกลุ่ม	-	-	-	-	-	-	13	-	-	3
	% ความผิดพลาด	-	-	-	-	-	-	1.5951	-	-	8.8235

บทที่ 6

สรุปผลการวิจัยและข้อเสนอแนะ

6.1 สรุปผลการวิจัย

ในงานวิจัยนี้นำเสนอ การจัดกลุ่มภาพคลื่นไฟฟ้าหัวใจโดยใช้วิธีการหาความสัมพันธ์ของข้อมูลที่เรียงลำดับกันอยู่ในแต่ละภาพคลื่นไฟฟ้าหัวใจ ซึ่งสามารถทำได้ในภาษาสม่ำเสมอ และใช้ออโตมาตาในการตรวจสอบความผิดปกติของภาพคลื่นไฟฟ้าหัวใจซึ่งสามารถสรุปได้ดังตารางที่ 6.1

ตารางที่ 6.1 สรุปเปอร์เซ็นต์ความผิดพลาดในผู้ป่วยแต่ละคน และค่าเฉลี่ยความผิดพลาดทั้งหมด

ชุดที่	% ความผิดพลาดของกลุ่มปกติ	% ความผิดพลาดของกลุ่มผิดปกติ
1	5.91	0
2	11.06	8.06
3	7.99	8.80
เฉลี่ย	8.19	5.87

จากข้อมูลทั้ง 3 ชุดเราสามารถสรุปได้ว่าความผิดพลาดของการตรวจสอบภาพคลื่นไฟฟ้าหัวใจที่มีการยอมรับภาพคลื่นไฟฟ้าหัวใจที่ผิดปกติมี 8.19% ดังแสดงในตารางที่ 6.1 และความผิดพลาดของการปฏิเสธภาพคลื่นไฟฟ้าหัวใจที่ปกติมี 5.87% การผิดพลาดทั้งหมดที่เกิดขึ้นนั้นยังไม่ทำให้เกิดกรณีที่ภาพรูปแบบใด รูปแบบหนึ่งหายไปซึ่งจะทำให้เกิดผลเสียในด้านการวินิจฉัยโรค

อนึ่งประสิทธิภาพเชิงเวลาของขั้นตอนวิธีการสร้างออโตมาตาที่ใช้ในการตรวจสอบนี้มีอัตราการเติบโตเชิงเวลาเป็น $O(n^2)$ เมื่อ n หมายถึงจำนวนภาพคลื่นไฟฟ้าหัวใจของข้อมูลชุดทดสอบ เกิดขึ้นจากกรณีที่ถ้าในการจัดกลุ่มนั้นทุก ๆ ภาพคลื่นไฟฟ้าหัวใจเกิดเป็นกลุ่มใหม่ทั้งหมด

6.2 ปัญหาที่พบจากการวิจัย

ในงานวิจัยนี้เกิดปัญหาหลายส่วน ดังต่อไปนี้

6.2.1 ในการตัดแบ่งภาพคลื่นไฟฟ้าหัวใจออกมาเป็นเพียงหนึ่งคลื่นไฟฟ้าหัวใจ ถ้าในภาพคลื่นไฟฟ้าหัวใจมีคลื่นรบกวน และยาวต่อเนื่องกันเป็นจำนวนมาก จะทำให้ระบบมีความผิดพลาดในการหาจุดสูงสุดที่เรียกว่าจุด R ได้ อีกทั้งยังส่งผลกระทบไปถึงการจัดกลุ่มของภาพคลื่นไฟฟ้าหัวใจนั้นต่อไปอีกด้วย

6.2.2 เวลาในการจัดกลุ่มนั้น จะยิ่งใช้เวลามากถ้าเกิดกลุ่มมากขึ้น เนื่องจากข้อมูลภาพคลื่นไฟฟ้าหัวใจแต่ละอันจะต้องถูกเปรียบเทียบกับกลุ่มที่เกิดขึ้นมาทั้งหมด ว่าข้อมูลภาพคลื่นไฟฟ้าหัวใจนั้น จะไปตกอยู่ที่กลุ่มใด

6.2.3 ภาษาในระดับภาษาสม่ำเสมอนั้นมีข้อจำกัดของจำนวนนับในระยะทางของเวลา กล่าวคือไม่สามารถบอกระยะเวลา หรือจำนวนการเกิดได้ ยกตัวอย่างเช่น ถ้าภาพคลื่นไฟฟ้าหัวใจ เกิดคลื่นเหมือน ๆ กันซ้อนขึ้นมา ไวยากรณ์ภาษาระดับภาษาสม่ำเสมอก็จะยังคงตอบว่า ภาพคลื่นไฟฟ้าหัวที่เกิดซ้อนขึ้นสองครั้ง กับภาพคลื่นไฟฟ้าหัวใจที่ไม่เกิดซ้อนนั้นมีความเหมือนกัน

6.3 ข้อเสนอแนะ

6.3.1 การลดทอนคลื่นรบกวน ของภาพคลื่นไฟฟ้าหัวใจก่อนที่จะเริ่มกระบวนการตัดแบ่งออกมาเป็นแต่ละภาพคลื่นไฟฟ้าหัวใจในบทที่ 3.1

6.3.2 การแปลงข้อมูลตัวเลขเป็นตัวอักษรเพื่อการหาความสัมพันธ์นั้น อาจสามารถเปลี่ยนไปในรูปอื่น และช่วยให้ง่ายยิ่งขึ้นในการหารูปแบบที่แตกต่างกัน

6.3.3 ภาษาในระดับ ภาษาสม่ำเสมอนั้นสามารถจัดกลุ่ม ภาพคลื่นไฟฟ้าหัวใจได้ ซึ่งถ้าใช้ภาษาในระดับสูงขึ้น เช่น ไวยากรณ์ไม่พึ่งบริบทก็อาจทำให้ได้ผลที่ดียิ่งขึ้นอีก เนื่องจากภาษาไม่พึ่งบริบทสามารถบอกถึงปริมาณที่แน่นอนได้ ไม่เหมือนภาษาสม่ำเสมอที่ไม่สามารถบอกจำนวนที่แน่นอนได้

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รายการอ้างอิง

[1] Sakakibara, Y., “Grammatical Inference in Bioinformatics”. Digital Object Identifier, 27(2005): 1051 – 1062.

[2] Trahanias, P. and Skordalakis, E. “Syntactic Pattern Recognition of the ECG”. Pattern Analysis and Machine Intelligence, 12(1990): 648 – 657.

[3] Pullpothong, J., “On-line probabilistic automata analysis and design”. Master’s Thesis, Department of Computer Engineer, Chulalongkorn University, 2548.

[4] Koski, A., “Modelling ECG signals with hidden Markov models”. Artificial Intelligence in Medicine, 8(1996): 453 – 471.

[5] Vidal, E., Thollard, F., de la Higuera, C., Casacuberta, F. and Carrasco, R. C. “Probabilistic Finite State Machine – Part I”. Pattern Analysis and Machine Intelligence, 27(2005): 1013 – 1025.

[6] Vidal, E., Thollard, F., de la Higuera, C., Casacuberta, F. and Carrasco, R. C. “Probabilistic Finite State Machine – Part II”. Pattern Analysis and Machine Intelligence, 27(2005): 1026 – 1038.

[7] Ron, D., Singer, Y. and Tishby, N. “The power of amnesia:Learning probabilistic automata with variable memory length”. Machine Learning, 25(1996): 117 – 142.

[8] Horowitz, L. S. “A Syntactic Algorithm for Peak Detection in Wave forms with Applications to Cardiography”. Communications of the ACM, 18(1975): 281 - 285.

[9] Gold, E.M. “Language identification in the limit.” Information and Control, 10(1967): 447-474.

[10] P. García and E. Vidal, "Inference of k-testable languages in the strict sense and application to syntactic pattern recognition" IEEE Trans. Pattern Analysis and Machine Intelligence, 12(1990): 920-925.

[11] Lin, J., Keogh, E., Lonardi, S. & Chiu, B. “A Symbolic Representation of Time Series, with Implications for Streaming Algorithms.” In proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery. 13(2003): 2-11.

[12] ชมพูนุท อ่องจริต. คลื่นไฟฟ้าหัวใจทางคลินิก. พิมพ์ครั้งที่ 6. กรุงเทพมหานคร: สำนักพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย, 2543.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นายสาธิต ตั้งติดธรรม เกิดเมื่อวันที่ 3 เมษายน พ.ศ. 2523 ที่จังหวัดกรุงเทพมหานคร สำเร็จการศึกษาระดับปริญญาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยมหิดล ในปี พ.ศ. 2545 และศึกษาต่อในระดับปริญญามหาบัณฑิต ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย