สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ระบบแปลงเสียงเป็นตัวอักษรเพื่อผู้พิการทางหู

SPEECH TO TEXT FOR DEAF

ชาญวิศ ประสานเหลืองวิไล

ปัทมา พรสุขศิริ

รุ่งโรจน์ ศิริปิ่น

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เลขหมู่.....
เลขทะเบียน 110955
วัน,เดือน,ปี - 7 S.A. 2553

b.....
i.....

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2552

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2552

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง ระบบแปลงเสียงเป็นตัวอักษรเพื่อผู้พิการทางหู

SPEECH TO TEXT FOR DEAF

ผู้จัดทำ

1.นายชาญวิศ ประสานเหลืองวิไล รหัสนักศึกษา 49010194

2.นางสาวปัทมา พรสุขศิริ รหัสนักศึกษา 49010565

3.นายรุ่งโรจน์ ศิริปิ่น รหัสนักศึกษา 49010796

อาจารย์ที่ปรึกษา

(อาจารย์วัจนพงศ์ เกษมศิริ)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระบบแปลงเสียงเป็นตัวอักษรเพื่อผู้พิการทางหู

นายชาญวิศ ประสานเหลืองวิไล 49010194

นางสาวปัทมา พรสุขศิริ 49010565

นายรุ่งโรจน์ ศิริปิ่น 49010796

อาจารย์วัจนพงศ์ เกษมศิริ อาจารย์ที่ปรึกษา

ปีการศึกษา 2552

บทคัดย่อ

โครงงานนี้จัดทำขึ้นเพื่อศึกษา ออกแบบ และพัฒนาโครงข่ายประสาทเทียม (Neural Network) ในการรู้จำเสียงภาษาไทยโดยนำโครงงานนี้มาประยุกต์ใช้งานเพื่อช่วยเหลือผู้พิการทางหูให้สามารถสื่อสารกับคนทั่วไปได้ โดยการพัฒนาแอพพลิเคชัน (Application) ที่สามารถเปลี่ยนช่องทางการสื่อสารจากสื่อที่เป็นเสียงภาษาไทยไปเป็นตัวอักษรภาษาไทยซึ่งสามารถรับสารได้ด้วยการอ่าน แอพพลิเคชันนี้ยังถูกพัฒนาขึ้นให้ใช้งานบนพีดีเอ (PDA) ทำให้สามารถพกพาไปใช้งานได้สะดวกอีกด้วย

แอพพลิเคชันที่นำเสนอนี้จะแบ่งการทำงานออกเป็นสองส่วน ในการทำงานส่วนแรกจะทำการรับเสียงเข้ามาแล้วทำการตัดเสียงออกเป็นพยางค์ นำไปรีแซมปลิง (Resampling) และนอร์มอลไลเซชั่น (Normalization) เพื่อปรับขนาดแล้วจึงถอดลักษณะเด่นของเสียงออกด้วย MFCC จากนั้นจึงนำไปผ่านโครงข่ายประสาทเทียม โครงสร้างแบบ Multi-Layer Perceptron ใช้กฎการเรียนรู้แบบ Back Propagation โดยใช้พยางค์ในภาษาไทยในการทดสอบจำนวน 100 พยางค์ จนกระทั่งได้ค่า Mean Square Error เท่ากับ 0.00055 ซึ่งจะเป็นส่วนที่ใช้ในการรู้จำเสียง แล้วแสดงผลลัพธ์ตัวอักษรบนหน้าจอของเครื่องพีดีเอ ซึ่งจากผลการทดสอบประสิทธิภาพของโปรแกรมพบว่า ความถูกต้องในการตัดพยางค์เท่ากับ ร้อยละ 91.21 และความถูกต้องของการรู้จำเสียงโดยเฉลี่ยเท่ากับ ร้อยละ 71.83 ในการทำงานส่วนที่สองแอพพลิเคชันจะรับคำภาษาไทยมาจากผู้ใช้แล้วทำการแบ่งออกเป็นพยางค์ จากนั้นนำแต่ละพยางค์มาทำการจับคู่กับพยางค์ที่เหมือนกันใน XML ไฟล์(file) แล้วเล่นไฟล์เสียงที่สอดคล้องกับผลลัพธ์การจับคู่นั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Speech to Text for Deaf

Mr.Chanvit	Prasanluangvilai	49010194
Miss Pattama	Pornsuksiri	49010565
Mr. Roongroj	Siripin	49010796
Mr. Watjanapong	Kasemsiri	Advisor

Academic Year 2009

ABSTRACT

In this project, we develop Thai speech to text system by utilizing neural network technique as recognition engine. The goal of this project is to help deaf people to communicate with other people who do not understand sign language. This application is designed to work on PDA which deaf people can conveniently carry with them.

This application consists of two parts. The first part is to convert speech to text. This part starts from receiving a speech signal from microphone then divide the received speech signal to syllables after that apply resampling and normalization function to each syllables, next we extract features from the speech signal by using MFCC then pass these features to Backpropagation Neural Network with Multi-Layer Perceptron structure (which is trained by 100 Thai syllables until Mean Square Error is 0.00055) for recognizing syllables and display results on PDA's screen. Program's performance tests show that accuracy of the syllables segmentation part is 91.21% and the accuracy of recognition part is 71.83%. For the second part, the application uses Thai text as it's input and matches each syllables with the reference XML file then the application plays the matched speech syllable file.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้เกิดขึ้นและสำเร็จได้ด้วยดี เนื่องจากมีผู้ที่ให้ความรู้และคำปรึกษาเกี่ยวกับโครงงานนี้มาตลอดปีการศึกษา คืออาจารย์ที่ปรึกษาโครงงาน อาจารย์ วัจนพงศ์ เกษมศิริ นอกจากนี้ยังมีพี่ที่คอยให้ความรู้ คือ นายสุธรรม โรจนวิภาต และเพื่อนๆที่สละเวลามาช่วยอัดเสียง มีรายชื่อดังนี้ เลอศักดิ์ บุญเพ็ง, กรกมล ตั้งจิตเพียรพงศ์, ชุติมา พฤฒิชัย, ดลฤดี สุนารี, ตติยา พรหมทา, ธนภัทร เรืองสันติ, หทัยกาญจน์ ปินตา, นภาพร ทวีจิต, ปรเมตต์ ปราสาร, ปุณยวีร์ ภักดียิ่ง, พงศ์ศักดิ์ ชัยชีวชื่น, พจนา แก้วปัญญา,พรรักษา กมลเวชช์,พิทวัส ภมรชัยศิริกิจ,วริษธิฌีย์ ศรีพันธุ์,วิทยา คุ้มปีย๋ะผล, สิทธิธัช รักสัตย์, สุจิตรา ชาวสวน, สุรพัฒน์ พระศรีรัมย์, อภิชัย เก่งสุรการ, เสาวลักษณ์ ตะเคียนงาม

และบุคคลสำคัญที่สุดที่สมาชิกในกลุ่มทุกคนมีความรู้สึกขอบคุณอยู่ตลอดเวลา คือ พ่อแม่ ผู้อยู่เบื้องหลังความสำเร็จของสมาชิกในกลุ่มทุกคน ขอบคุณความห่วงใย กำลังใจและความเกื้อกูลที่มีเสมอมา ขอมอบความดีงามที่ปริญญานิพนธ์ฉบับนี้พึงมีให้ท่านเพื่อแทนคำขอบคุณจากสมาชิกในกลุ่มทุกคน

นายชาญวิศ ประสานเหลืองวิไล
นางสาวปัทมา พรสุขศิริ
นายรุ่งโรจน์ ศิริปิ่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....I

บทคัดย่อภาษาอังกฤษ.....II

กิตติกรรมประกาศ.....III

สารบัญ.....IV

สารบัญตาราง.....VI

สารบัญภาพ.....VII

บทที่ 1 บทนำ.....1

1.1 ความเป็นมา..... 1

1.2 ความมุ่งหมายและวัตถุประสงค์.....2

1.3 ประโยชน์ที่คาดว่าจะได้รับ.....2

1.4 ขอบเขตของโครงงาน.....2

1.5 ขั้นตอนการดำเนินงาน.....3

1.6 ข้อจำกัดโปรแกรม..... 4

บทที่ 2 การแปลงเสียงเป็นตัวอักษร

2.1 ทฤษฎีเบื้องต้นเกี่ยวกับเสียง.....5

2.2 การตรวจจับเสียงที่เป็นคำพูด.....6

2.3 การเตรียมข้อมูลก่อนการประมวลผล.....10

2.4 โครงข่ายประสาทเทียม..... 17

2.5 ระบบการแปลงตัวอักษรเป็นเสียง.....27

บทที่ 3 การออกแบบและพัฒนา

3.1 ความต้องการของระบบ..... 30

3.2 ภาพรวมของระบบ..... 31

3.3 การออกแบบอัลกอริทึม..... 32

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 4 การทดลอง

4.1 การทดลองส่วนการแปลงเสียงเป็นตัวอักษร..... 47

4.2 การทดลองส่วนการแปลงตัวอักษรเป็นเสียง.....76

บทที่ 5 สรุป

5.1 สรุป.....78

5.2 ปัญหาและอุปสรรค.....78

5.3 แนวทางการแก้ไข..... 79

5.4 แนวทางการพัฒนาต่อ..... 79

บรรณานุกรม..... 80

ภาคผนวก

ภาคผนวก ก..... 82

ภาคผนวก ข..... 83

ภาคผนวก ค..... 84

ภาคผนวก ง.....88

ภาคผนวก จ..... 89

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตาราง **หน้า**

1.1 ระดับการได้ยินของคนหูหนวก.....1

4.1 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ..... 48

4.2 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น.....49

4.3 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล.....50

4.4 สรุปผลการทดลอง.....51

4.5 สรุปผลการทดลองเป็นเปอร์เซ็นต์.....51

4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น.....53

4.7 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ..... 59

4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล.....65

4.9 สรุปผลการทดลอง.....70

4.10 จำนวนพยางค์ที่อยู่ในแต่ละช่วงเปอร์เซ็นต์ความถูกต้อง..... 74

4.11 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล.....75

4.12 ผลการทดลองส่วนการแปลงตัวอักษรเป็นเสียง.....76

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ

รูป **หน้า**

2.1 การบีบอัด-คลาย ของตัวกลางบริเวณรอบๆ แหล่งกำเนิดคลื่นเสียง..... 5

2.2 การแบ่งเฟรมของพรีโปรเซสซิง.....7

2.3 คลื่นก่อนและหลังคำนวณทีเจอร์ เอ็นเนอร์จี..... 8

2.4 ความหมายของตัวแปรในสมการมูฟวิง เอเวอร์เรจ.....9

2.5 คลื่นก่อนและหลังทำ มูฟวิง เอเวอร์เรจ.....9

2.6 สัญญาณเสียงก่อนและหลังทำรีแซมปลิง.....11

2.7 สัญญาณเสียงก่อนและหลังทำนอร์มอลไลเซชัน.....12

2.8 กราฟการกระจายตัวของความถี่บนสเกลเมล..... 13

2.9 แผนภาพลำดับการทำงานของสัมประสิทธิ์เซปสตรอลบนความถี่เมล.....14

2.10 ชุดตัวกรองแบบสามเหลี่ยม..... 15

2.11 สัญญาณเสียงถูกแปลงฟูริเยร์แบบเร็วและส่งผ่านชุดตัวกรองสามเหลี่ยม..... 16

2.12 แบบจำลองของเซลล์ประสาทในสมองมนุษย์..... 17

2.13 การคำนวณค่าเอาท์พุตของโครงข่ายประสาทเทียม.....18

2.14 ตัวอย่างโครงข่ายประสาทเทียมแบบชั้นเดียว.....20

2.15 ตัวอย่างโครงข่ายประสาทเทียมแบบหลายชั้น..... 21

2.16 การเรียนรู้แบบมีผู้สอน..... 22

2.17 การเรียนรู้แบบไม่มีผู้สอน.....22

2.18 โครงสร้างของโครงข่ายประสาทเทียมส่วนฟอร์เวิร์ดพาส.....24

2.19 โครงสร้างของระบบแปลงตัวอักษรเป็นเสียง.....28

3.1 แผนภาพโดยรวมของระบบ.....31

3.2 สัญญาณเสียงเมื่อผ่านสมการทีเจอร์ เอนเนอร์จี.....33

3.3 กระบวนการรู้จำคำพูด..... 34

3.4 โฟลว์ชาร์ตอัลกอริทึมในการตัดสัญญาณเสียงที่เป็น.....35

3.5 โฟล์วชาร์ตอัลกอริทึมการแปลค่าเอาท์พุท..... 36

3.6 โฟล์วชาร์ตอัลกอรึทึมการวิเคราะห์ข้อความ.....38

3.7 กระบวนการแยกพยางค์ออกจากประโยค “ผมชอบคุณ” 39

3.8 กระบวนการแยกพยางค์ออกจากประโยค “ราคา123บาท”.....40

3.9 โฟล์วชาร์ตอัลกอริทึมการอ่านตัวเลข.....43

3.10 หน้าหลักของโปรแกรม..... 44

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญภาพ(ต่อ)

รูป **หน้า**

3.11 หน้าต่างการตั้งค่า 45
3.12 หน้าต่างตั้งค่าตัวอักษร 45
3.13 หน้าต่างบันทึกการสนทนา.....46
4.1 กราฟความสัมพันธ์ระหว่างระยะเวลาที่พีดีเอประมวลผลต่อ1คำ กับจำนวนคำ.....75

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1
บทนำ

1.1 ความเป็นมา

ความพิการทางหูนั้นมีอยู่หลายประเภททั้งผู้ที่พิการมาตั้งแต่กำเนิด และผู้ที่พิการภายหลังด้วยอุบัติเหตุหรือสาเหตุอื่นๆ ความพิการทางหูนี้ สมาคมโสต ศอ นาสิก แพทย์ ได้แบ่งระดับความพิการของหูออกเป็น 6 ระดับ โดยใช้ค่าเฉลี่ยการได้ยินในช่วง 500-2,000 เฮิร์ท (Hz) เริ่มตั้งแต่หูปกติจนถึงหูหนวก ดังตาราง 1.1

ตาราง1.1 ระดับการได้ยินของคนหูหนวก

ระดับความพิการ	ค่าเฉลี่ยการได้ยินที่ 500-2,000 เฮิร์ท ในหูข้างที่ดีกว่า 150 เดซิเบล(dB)	ลักษณะการรับฟัง	ความสามารถในการพูดและฟัง
หูปกติ	น้อยกว่า 27 เดซิเบล	-	ปกติ
หูตึงระดับที่ 1	27-40 เดซิเบล	หูตึงน้อย	ไม่ได้ยินเสียงพูดเบาๆ
หูตึงระดับที่ 2	40-55 เดซิเบล	หูตึงปานกลาง	ไม่ได้ยินเสียงธรรมดา
หูตึงระดับที่ 3	55-70 เดซิเบล	หูตึงมาก	พูดดังเต็มที่แล้วไม่ได้ยิน
หูตึงระดับที่ 4	70-93 เดซิเบล	หูตึงอย่างรุนแรง	ต้องตะโกนหรือใช้เครื่องขยายจึงจะได้ยิน
หูหนวก	ตั้งแต่ 93 เดซิเบล ขึ้นไป	หูหนวก	ใช้เครื่องขยายแล้วยังไม่เข้าใจ

โดยส่วนใหญ่ผู้พิการจะมีระดับความพิการอยู่ที่ หูหนวก ฉะนั้นถึงแม้ว่าจะใช้เครื่องช่วยฟังแล้วก็ไม่สามารถที่จะเข้าใจสิ่งที่คนปกติพูดได้ จะได้ยินเพียงเสียงบางเสียง เช่น บางคนได้ยินเฉพาะเสียงสูงๆ บางคนได้ยินเสียงแตรรถ เป็นต้น จากตรงนี้จะเห็นว่าเครื่องช่วยฟังจะช่วยผู้พิการได้เพียงเล็กน้อยเท่านั้น และไม่สามารถช่วยด้านการสื่อสารได้อีกด้วย จึงจำเป็นอย่างยิ่งที่จะต้องมีเครื่องมือที่จะช่วยในการสื่อสารระหว่างผู้พิการทางหูกับคนปกติ เพราะส่วนมากคนปกติมักไม่สามารถเข้าใจภาษามือ ทำให้ผู้พิการถูกปิดกั้นทางสังคม และถูกตัดขาดจากข่าวสารที่อยู่ในรูปเสียงไปโดยปริยาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผู้พิการทางหูนั้นจะต้องรับรู้และจดจำทุกอย่างเป็นภาพ ดังนั้นถึงแม้ว่าจะเข้าโรงเรียนเพื่อเรียนหนังสือก็ไม่ได้หมายความว่าจะรับรู้ทุกอย่างได้เต็ม 100% จากงานวิจัยเรื่อง "การรับรู้ด้านสุขภาพของกลุ่มผู้พิการทางการได้ยิน" ทำให้ทราบว่าผู้พิการมักมีปัญหาด้านสุขภาพ อนามัย เช่นเวลาไปหาหมอ เพราะรู้ว่าไม่สบายแต่ไม่สามารถบอกถึงอาการได้ จากจุดนี้ถ้าหากมีเครื่องมือที่ช่วยสื่อสารระหว่างหมอ กับผู้พิการก็จะช่วยให้มีความเข้าใจตรงกันมากขึ้น และช่วยให้ผู้พิการสามารถไปหาหมอได้ด้วยตนเอง ไม่จำเป็นต้องพาล่ามไปด้วย

ปัจจุบันมีผู้พิการทางหูน้อยคนที่จะได้รับโอกาสในการทำงานร่วมกับคนปกติ เนื่องจากคนส่วนใหญ่ไม่สามารถสื่อสารด้วยภาษามือได้ จึงปิดกั้นโอกาสของผู้พิการ อีกทั้งข่าวสารต่างๆทางวิทยุและโทรทัศน์ที่ไม่มีภาษามือประกอบ ก็ทำให้ผู้พิการไม่สามารถรับรู้ข่าวสารนั้นๆได้เลย ดังนั้นหากเรามีเครื่องมือที่ช่วยให้บุคคลทั่วไปสามารถสื่อสารกับผู้พิการได้ ก็จะเป็นการเปิดโอกาสในการทำงาน ทำกิจกรรม และแลกเปลี่ยนทัศนคติกันได้ เป็นการช่วยให้ผู้พิการสามารถดำรงชีวิตในสังคมได้ด้วยตนเองอย่างสะดวกสบาย และเท่าเทียมกับคนทั่วไป

1.2 ความมุ่งหมายและวัตถุประสงค์

1) เพื่อศึกษา ออกแบบ และพัฒนาโครงข่ายประสาทเทียม (neural network) ในการรู้จำเสียงภาษาไทย
2) เพื่อเป็นแนวทางการสร้างเครื่องมือในการสื่อสารระหว่างผู้พิการทางหูกับผู้ที่ไม่เข้าใจภาษามือไทย

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1) ได้รับความรู้เรื่องโครงข่ายประสาทเทียม
2) ได้รับความรู้เรื่องภาษามือไทย
3) ได้รับความรู้เรื่องส่วนติดต่อผู้ใช้งาน (User Interface)
4) ได้รับความรู้เรื่องการสร้างแอพลิเคชัน (application) บนพีดีเอ (PDA)

1.4 ขอบเขตของโครงงาน

1) โปรแกรมที่พัฒนาขึ้นใช้คำในการทดสอบจำนวน 100 คำ ซึ่งคัดเลือกมาจากหนังสือภาษามือไทย ฉบับปรับปรุง
2) โปรแกรมนี้พัฒนาขึ้นให้สามารถทำงานบนเครื่องพีดีเอได้
3) การทำงานของโปรแกรมแบ่งเป็น 2 ส่วน คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1) ส่วนรับเสียง จะรับเสียงจากไมโครโฟน (Microphone) จากนั้นระบบจะแปลงเสียงเป็นตัวอักษรแสดงบนจอพีดีเอ

3.2) ส่วนตอบกลับ ผู้พิการทางหูจะต้องพิมพ์อักษรตอบกลับ เมื่อระบบรับประโยคมาแล้วจะตัดเป็นพยางค์แล้วนำแต่ละพยางค์มาจับคู่กับไฟล์เสียงที่เหมาะสม และแสดงเสียงออกทางลำโพง

1.5 ขั้นตอนการดำเนินงาน

1) สำรวจความต้องการของผู้ใช้งาน กลุ่มตัวอย่างที่สำรวจคือ สมาคมคนหูหนวกแห่งประเทศไทย ได้สอบถามเกี่ยวกับการสื่อสารกับคนทั่วไปในชีวิตประจำวัน การศึกษา และปัญหาที่เกิดขึ้นในการสื่อสาร รวมถึงการแก้ไขปัญหา

2) จากการสำรวจความต้องการของผู้ใช้งาน พบว่าผู้พิการทางหูในประเทศไทยส่วนใหญ่เรียนภาษามือจากหนังสือภาษามือไทย จึงคัดเลือกคำไทยจำนวน 100 พยางค์ ที่จะใช้ฝึกหัดจากหนังสือดังกล่าว โดยเลือกเฉพาะคำที่จำเป็นในการเดินทางและพบแพทย์

3) บันทึกเสียงคำไทย 100 พยางค์ โดยบันทึกพยางค์ละ 10 ครั้ง จากสมาชิกในกลุ่มจำนวน 3 คนเพื่อเป็นเสียงทดลองในขั้นตอนก่อนฝึกหัดโครงข่ายประสาทเทียม

4) ศึกษาอัลกอริทึม (algorithm) ต่างๆสำหรับตรวจจับสัญญาณส่วนที่เป็นคำพูด เพื่อหาอัลกอริทึมที่เหมาะกับโครงงานนี้

5) เขียนโปรแกรมจากทฤษฎี ซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอจี (Syllabel Segmentation Base on Energy) สำหรับตรวจจับสัญญาณส่วนที่เป็นคำพูด เพื่อใช้สำหรับการทดลองหาค่าพารามิเตอร์ (Parameter) ที่เหมาะสมที่สุดในการตรวจจับคำพูด

6) นำไฟล์เสียงทั้งหมดมาตัดคำพูดเพื่อใช้ในการฝึกสอนโครงข่ายประสาทเทียม

7) ศึกษาการทำรีแซมปลิ้ง (Resampling) เพื่อปรับขนาดคำพูดให้มีจำนวนแซมปลิงเท่ากัน

8) ศึกษาการทำนอร์มอล ไลเซชั่น (Normalization) เพื่อปรับค่าแอมปลิจูดสูงสุด (Max Amplitude) ของเสียงแต่ละคำให้มีขนาดเท่ากัน

9) ศึกษาการทำงานของโครงข่ายประสาทเทียม และหารูปแบบที่เหมาะสมในการรู้จำเสียงภาษาไทยโดยไม่ขึ้นกับผู้พูด

10) ศึกษาการหาค่าสำคัญของเสียงพูดโดยใช้ สัมประสิทธิเซปสตรอลบนความถี่เมล (Mel Frequency Cepstral Coefficient, MFCC) รวมถึงศึกษาขั้นตอนต่างๆโดยละเอียด ดังนี้

- แฮมมิ่งวินโดว์ (Hamming Window)
- การแปลงฟูริเยแบบเร็ว (Fast Fourior Transform)
- ชุดตัวกรองแบบสามเหลี่ยม (Filter Bank)
- การแปลงโคไซน์ไม่ต่อเนื่อง (Discrete Cosine Transformation)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.6 ข้อจำกัดโปรแกรม

1) ใช้ได้กับคำไทยจำนวน 100 คำเท่านั้น

2) ใช้ได้ดีเฉพาะในสถานที่ที่มีความดังระดับสำนักงานปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

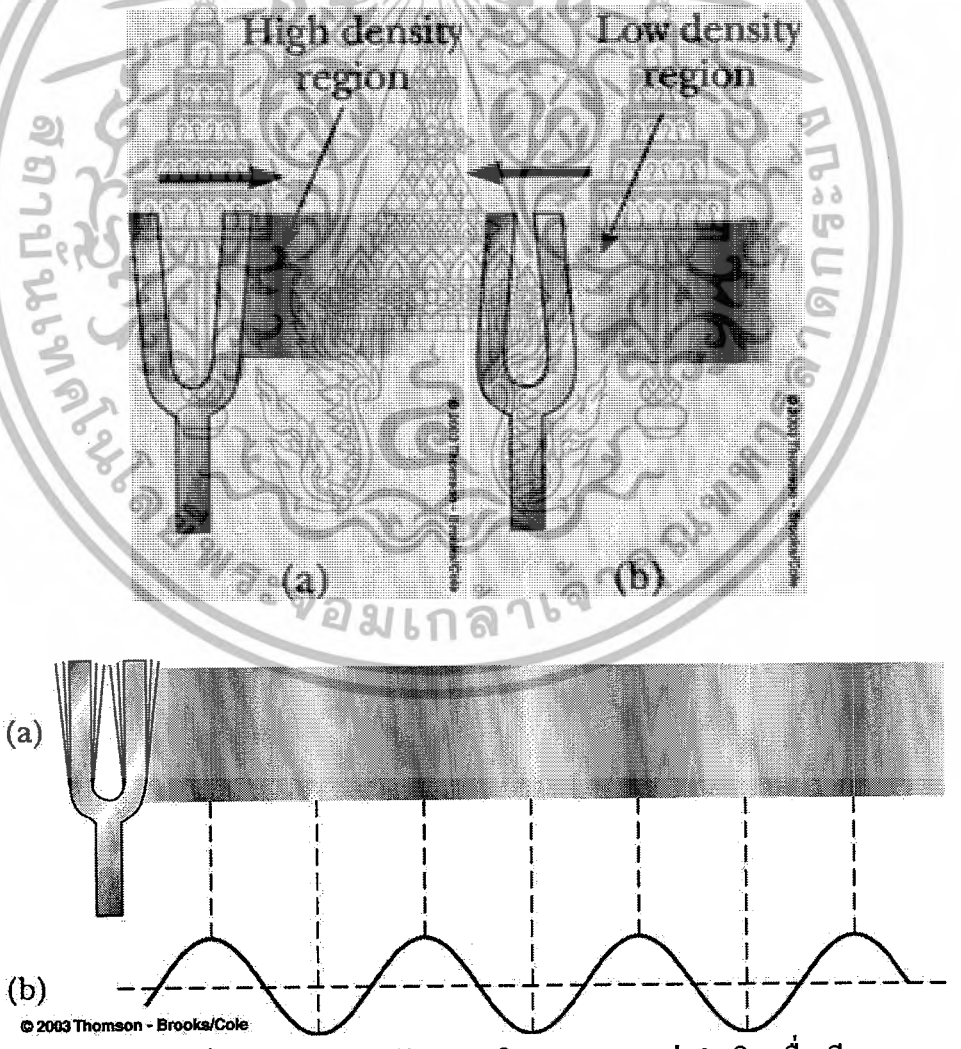
บทที่ 2
การแปลงเสียงเป็นตัวอักษร

2.1 ทฤษฎีเบื้องต้นเกี่ยวกับเสียง

2.1.1 คลื่นเสียง (Sound)

2.1.1.1 แหล่งกำเนิดคลื่นเสียง

เสียงเป็นคลื่นตามยาว (Longitudinal waves) เคลื่อนที่ผ่านตัวกลาง ตัวอย่าง คลื่นเสียงจากซ่อมเสียง ขณะที่ตีซ่อมเสียงจะเกิดการบีบอัด-คาย ของตัวกลางบริเวณรอบๆ แหล่งกำเนิดคลื่นเสียงและ เคลื่อนที่ออกไปโดยรอบและสามารถแทนระดับความเข้มของการบีบอัด-คาย ได้ด้วยรูปคลื่น ดังรูป 2.1



รูป 2.1 การบีบอัด-คลาย ของตัวกลางบริเวณรอบๆ แหล่งกำเนิดคลื่นเสียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.2 รายละเอียดทั่วไปของคลื่นเสียง

ความถี่คลื่นเสียงคือระดับความถี่ที่มนุษย์สามารถรับฟังได้ปกติ ระหว่าง 20 - 20,000 เฮิร์ท คลื่นต่ำ (Infrasonic waves) ความถี่ที่ต่ำกว่า 20 เฮิร์ท คลื่นเหนือเสียง (Ultrasonic waves) ความถี่ที่สูงกว่า 20,000 เฮิร์ท ความสัมพันธ์ระหว่างความถี่ ความยาวคลื่น และความเร็วคลื่นเป็นดังสมการ 2.1

$$v = \lambda f \tag{2.1}$$

2.1.1.3 ความดังเสียง (Threshold of hearing)

ความเข้มเสียงที่เบาสุดที่มนุษย์สามารถได้ยิน ประมาณ 1.0×10^{-12} W/m^2 ดังสุดที่หูมนุษย์สามารถทนฟังได้ 1 W/m^2 หูมนุษย์มีความสามารถตอบสนองต่อการได้ยินเสียงมาก

2.1.1.4 ระดับความเข้มเสียง

ระดับความดังที่กำหนดเป็นสมการล็อกการิทึมของสัดส่วนความเข้มเสียง β มีหน่วยเป็นเดซิเบล ดังสมการ 2.2

$$\beta = 10\log\frac{I}{I_0} \tag{2.2}$$

เมื่อ I_0 เป็นความเข้มเสียงเบาสุดที่ได้ยินได้เท่ากับ 1.0×10^{-12} W/m^2 เบาสุดจนรับฟังไม่ได้เท่ากับ 0 เดซิเบล ดังสุดจนหูทนฟังไม่ได้คือมากกว่า 120 เดซิเบล (เสียงเครื่องบินประมาณ150 เดซิเบล)

2.2 การตรวจจับเสียงที่เป็นคำพูด

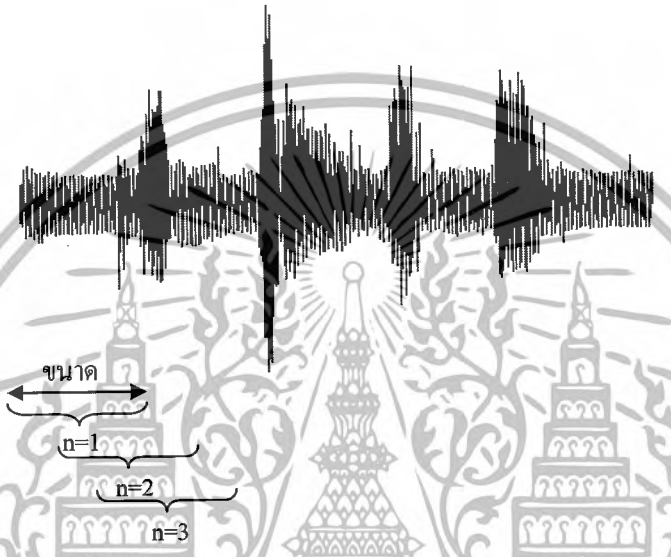
ในปัจจุบันมีหลายอัลกอริทึมที่ใช้ สำหรับการตรวจจับคำพูดจากสัญญาณเสียงส่วนใหญ่ อัลกอริทึมที่นิยมใช้มักจะคำนวณบนพื้นฐานของพลังงาน เนื่องจากง่ายและรวดเร็วในการคำนวณ โดยในแต่ละอัลกอริทึมจะมีรูปแบบในการนำพลังงานมาคำนวณโดยรูปแบบที่แตกต่างกันไป

ซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอร์จี เป็นอัลกอริทึมหนึ่งที่มีความเหมาะสมในการนำมาคำนวณเพื่อการตรวจจับคำพูดจากสัญญาณเสียง ประกอบด้วย 3 ขั้นตอนคือ พรีโปรเซสซิง (Pre-processing), พารามิเตอร์โพวิเดนซ์ (Parameter Providence), และดีเทคชัน (Detection) มีรายละเอียดดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.1 พรีโปรเซสซิง

โครงงานนี้ศึกษาสัญญาณเสียงบนอัตราการสุ่ม (Sampling rate) 44,100 เฮิร์ท ควอนทิเซชั่น (Quantization) 16 บิท (Bit) ในขั้นตอนพรีโปรเซสซิงเราจะทำการแบ่งพลังงานออกเป็นเฟรม (frame) ดังรูป 2.2 ใน 1 เฟรมมี 1024 ตัวอย่าง (Sample) โดยแต่ละเฟรมจะมีการซ้อนทับ (overlap) กัน 3/4 เฟรม (738 ตัวอย่าง) ซึ่งเราจะนำข้อมูลในแต่ละเฟรมมาทำการคำนวณในส่วนของ พารามิเตอร์โพวิเดนซ์ ต่อไป



รูป 2.2 การแบ่งเฟรมของพรีโปรเซสซิง

2.2.2 พารามิเตอร์โพวิเดนซ์

เป็นขั้นตอนการนำข้อมูลเสียงในเฟรมมาทำการคำนวณหาค่าจำเพาะของแต่ละเฟรม และทำการปรับค่าที่ได้จากแต่ละเฟรมให้ดียิ่งขึ้น โดยในโครงงานนี้จะใช้สองขั้นตอนคือ

2.2.2.1 การคำนวณหาค่าจำเพาะ

แอ็บโซลูท เอ็นเนอร์จี (Absolute Energy)

$$E_n = \frac{1}{w}\sum_{i=1}^{w}\left|S_n[i]\right| \tag{2.3}$$

เดอะ รูท มีน สแควร์ เอ็นเนอร์จี (The Root Mean Square Energy, RMS)

$$E_n = \left[\frac{1}{w}\sum_{i=1}^{w}S_n^2[i]\right]^2 \tag{2.4}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เดอะ สแควร์ เอ็นเนอร์จี (The Square Energy)

$$E_n = \frac{1}{w}\sum_{i=1}^{w} S_n^2[i] \tag{2.5}$$

ทีเจอร์ เอ็นเนอร์จี (Teager Energy)

$$E[i] = S^2[i] - S[i-1]S[i+1] \tag{2.6}$$

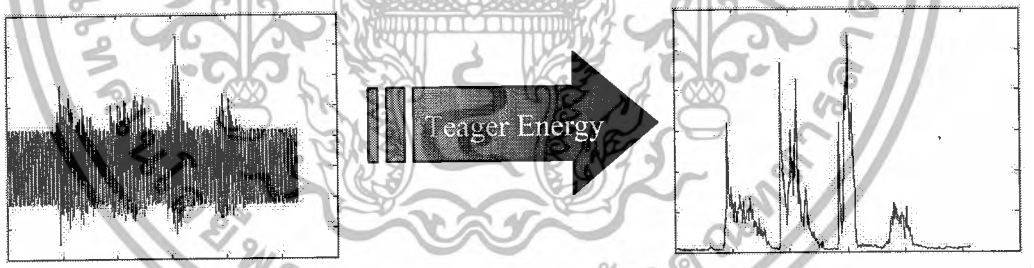
$$E_n = \sum_{i=1}^{w} E_n[i] \tag{2.7}$$

โดย E_n คือ ค่าจำเพาะของข้อมูลเสียงเฟรมที่ n

$S_n[i]$ คือ ค่าข้อมูลเสียงของเฟรมที่ n ตัวที่ i

W คือ ขนาดของเฟรม

ในโครงงานนี้เลือกใช้วิธี ทีเจอร์ เอ็นเนอร์จี ในการคำนวณขั้นตอนพารามิเตอร์โพวิเดนซ์ เนื่องจากเป็นวิธีที่ให้ความแม่นยำสูงสุด ซึ่งได้ลักษณะผลลัพธ์ดังรูป 2.3



คลื่นเสียงที่รับเข้ามาทางอินพุท (input) คลื่นหลังทำ ทีเจอร์ เอ็นเนอร์จี

รูป 2.3 คลื่นก่อนและหลังคำนวณ ทีเจอร์ เอ็นเนอร์จี

2.2.2.2 การนำค่าจำเพาะมาปรับความราบเรียบ

โดยขั้นตอนนี้จะใช้วิธี มูฟวิง เอเวอร์เรจ (Moving Average) ในการปรับค่าจำเพาะที่ได้จากหัวข้อ 2.2.2.1 ให้มีความราบเรียบมากยิ่งขึ้นราบเรียบโดยใช้สมการ 2.8 แบ่งจำนวนเฟรมด้านซ้ายและด้านขวาดังรูป 2.4

$$E_i^{'} = \frac{1}{m_1 + m_2 + 1}\sum_{i=n-m_1}^{n+m_2} E_i \tag{2.8}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

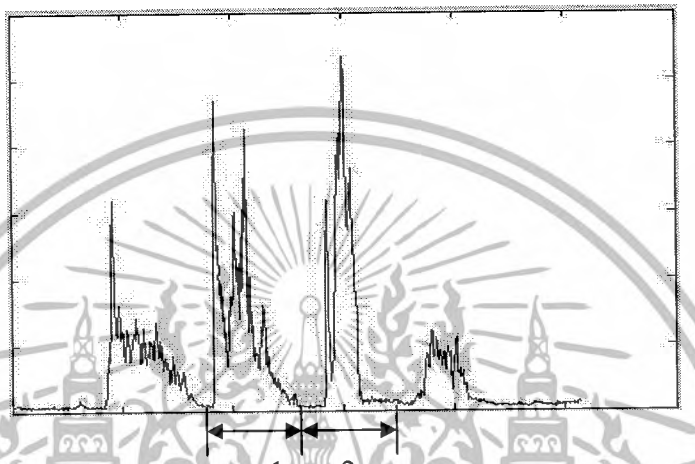
โดย E_n' คือค่าจำเพาะของเฟรมที่ n ใหม่หลังจาการทำ มูฟวิง เอเวอร์เรจ

m1 คือค่าจำนวนเฟรมที่อยู่ทางด้านซ้ายของเฟรมที่เราสนใจ

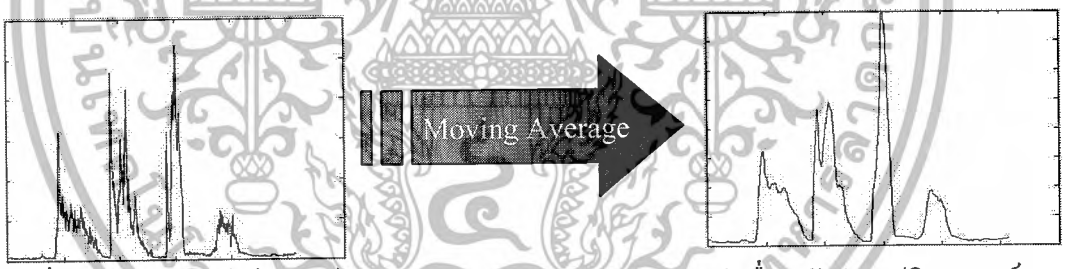
m2 คือ ค่าจำนวนเฟรมที่อยู่ทางด้านขวาของเฟรมที่เราสนใจ

E_n คือ ค่าจำเพาะของเฟรมที่ n

ผลลัพธ์หลังจากคำนวณด้วยมูฟวิง เอเวอร์เรจเป็นดังรูป 2.5



รูป 2.4 ความหมายของตัวแปรในสมการ มูฟวิง เอเวอร์เรจ



รูปคลื่นหลังทำ ทีเจอร์ เอ็นเนอร์ รูปคลื่นหลังทำ มูฟวิง เอเวอร์เรจ

รูป 2.5 คลื่นก่อนและหลังทำ มูฟวิง เอเวอร์เรจ

2.2.3 การแยกสัญญาเสียงที่เป็นคำพูดออกจากพื้นหลัง

เนื่องจากโครงงานนี้ทำการรู้จำเสียงบนอุปกรณ์ที่สามารถพกพาได้โดยที่แต่ละสถานที่จะมีเสียงพื้นหลัง (Background sound) ที่แตกต่างกัน ดังนั้นในโครงงานนี้จะทำการนำค่าเสียงพื้นหลัง มาทำการเข้าสมการ ทีเจอร์ เอ็นเนอร์จี เพื่อนำมาคำนวณหา ค่าเทรสโฮลด์ (Thresholds) โดยโครงงานนี้จะใช้ค่าค่าเทรสโฮลด์ ในการตรวจจับสัญญาณเสียงที่เป็นคำพูด 2 ค่า ด้วยกันคือ ค่าเทรสโฮลด์ โลว์ (Thresholds low) และค่าเทรสโฮลด์ ไฮ (Thresholds high) ซึ่งสามารถคำนวณหาค่าดังกล่าวได้ตามสมการข้างล่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$T_L = \alpha \times B_n^{'} \tag{2.9}$$

$$T_H = \beta \times B_n^{'} \tag{2.10}$$

โดย T_L คือ ค่าเทรสโฮลด์ โลว์

T_h คือ ค่าเทรสโฮลด์ ไฮ

$B_n^{'}$ คือ ค่าทีเจอร์ เอ็นเนอร์จี ของสัญญาณเสียงที่เป็นพื้นหลัง

สำหรับค่า β และ α เป็นตัวแปรที่มีไว้สำหรับปรับค่าเทรสโฮลด์ โลว์และค่าเทรสโฮลด์ ไฮ สามารถเลือกใช้ได้ตามความเหมาะสม หากเลือกค่า α ที่ต่ำเกินไปจะทำให้ไม่สามารถแยกคำออกจากคลื่นเสียงได้ แต่ถ้าสูงไปจะทำให้คำพูดที่ตัดมาได้ไม่ครบตัวอย่าง หรือถ้าสูงมากอาจทำให้ไม่สามารถแยกคำพูดออกมาได้เลย ส่วนค่า β หากต่ำไปจะทำให้แยกสัญญาณรบกวนออกมาด้วยหากสูงไปจะทำให้ไม่สามารถแยกคำพูดออกมาได้ สำหรับในโครงงานนี้ปรับค่า $\alpha = 1.5$ และ $\beta = 7$

การตรวจจับสัญญาณเสียงที่เป็นคำพูดโดยขั้นตอนในการตรวจจับอยู่ด้วยกัน 4 ขั้นตอนดังนี้

2.2.3.1 หาจุดเริ่มต้นของคลื่นเสียงที่เป็นคำพูด

โดยทำการเปรียบเทียบค่าพลังงานแต่ละเฟรมของสัญญาณเสียงหลังทำ มูฟวิง เอเวอร์เรจ กับค่าเทรสโฮลด์แล้ว หากเจอค่าพลังงานมีค่ามากกว่าค่าค่าเทรสโฮลด์ โลว์ ให้ถือว่าเฟรมเสียงเฟรมนั้นเป็นจุดเริ่มต้นของสัญญาณเสียงที่น่าจะเป็นคำพูดและเข้าสู่ขั้นตอนในข้อต่อไป

2.2.3.2 ยืนยันว่าเป็นคำพูด

โดยทำการเทียบค่าพลังงานของเฟรมถัดไปกับค่าค่าเทรสโฮลด์ ไฮและค่าค่าเทรสโฮลด์ โลว์ หากค่าพลังงานของเฟรมถัดไปมีค่ามากกว่าค่าเทรสโฮลด์ ไฮ ให้ถือว่าก้อนเสียงที่กำลังพิจารณาอยู่เป็นคำพูดและเข้าสู่ขั้นตอนต่อไป หากมีค่าต่ำกว่าค่าเทรสโฮลด์ โลว์จะถือว่าก้อนเสียงที่กำลังพิจารณาอยู่เป็นเพียงเสียงพื้นหลังจะทำการกลับไปตรวจสอบต่อในขั้นตอนแรก

2.2.3.3 หาจุดสิ้นสุดของคำพูด

โดยการเปรียบเทียบค่าพลังงานในเฟรมถัดไปจนกว่าจะเจอเฟรมที่มีพลังงานน้อยกว่าค่าค่าเทรสโฮลด์ โลว์ จากขั้นตอนนี้จะทำให้เรารู้ว่าคำพูดที่กำลังตรวจจับอยู่มีเฟรมไหนบ้าง

2.2.3.4 เปลี่ยนเฟรมแต่ละเฟรมกลับไปเป็นสัญญาณเสียง

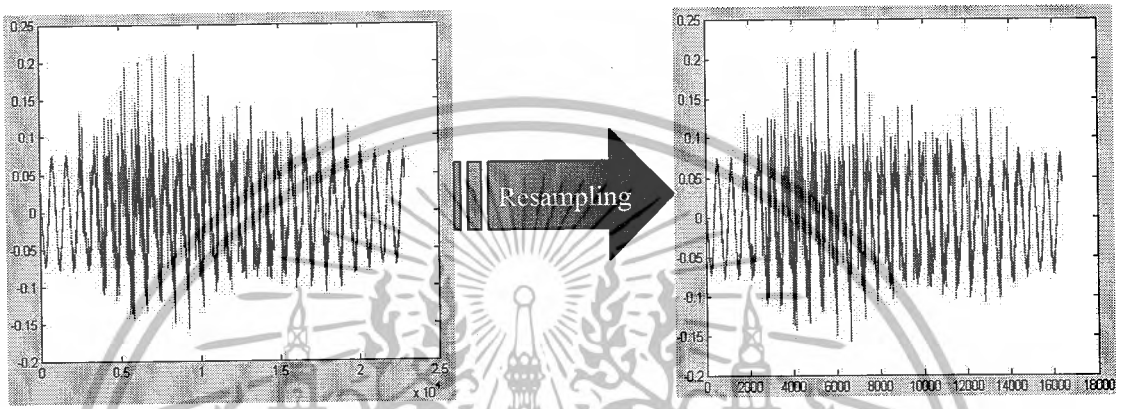
โดยการนำเฟรมที่เป็นจุดเริ่มต้นและจุดสิ้นสุดไปจับคู่กับสัญญาณเสียงก่อนการทำซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอร์จี เพื่อหาตัวอย่างเริ่มต้นและสิ้นสุดของสัญญาณเสียงที่เป็นคำพูด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3 การเตรียมข้อมูลก่อนประมวลผล

2.3.1 การเปลี่ยนแปลงความถี่ในการสุ่ม (Resampling)

เป็นการนำสัญญาณเสียงมาทำการสุ่ม (Sampling) ใหม่เพื่อให้มีจำนวนตัวอย่างที่ต่างไปจากเดิมดังรูป 2.6 โดยในโครงนี้จะนำสัญญาณเสียงที่เป็นคำพูด มาทำการรีแซมปลิงให้แต่ละคำพูดมีจำนวนตัวอย่างที่เท่าๆกัน



สัญญาณเสียงก่อนทำรีแซมปลิง สัญญาณเสียงหลังทำรีแซมปลิง

รูป 2.6 สัญญาณเสียงก่อนและหลังทำรีแซมปลิง

โดยการรีแซมปลิงจะใช้สมการดังข้างล่าง

$$y[k] = x[round((k-1) \times q / p) + 1)] \tag{2.11}$$

โดยที่ y[k] คือ ตัวอย่างที่ k ของสัญญาณเสียงหลังทำรีแซมปลิง

x[k] คือ ตัวอย่างที่ k ของสัญญาณเสียงก่อนทำรีแซมปลิง

round (i) คือ ค่าจำนวนเต็มที่ใกล้เคียงค่า i ที่สุดหากมีสองค่าจะเลือกค่าที่มากกว่า i

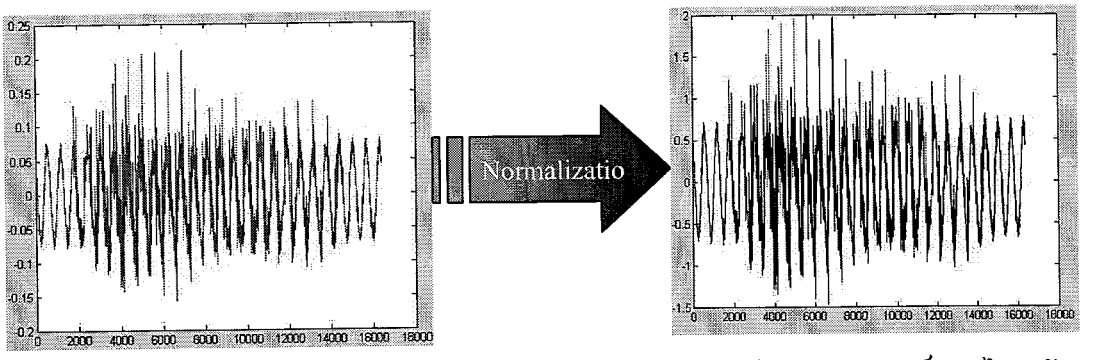
q คือ จำนวนตัวอย่างใหม่

p คือ จำนวนตัวอย่างเก่า

2.3.2 นอร์มอลไลเซชัน

เป็นการปรับแอมปลิจูด (Amplitude) ของสัญญาณเสียงโดยที่เสียง ไม่มีการบิดเบือนไป เนื่องจากเสียงคำๆเดียวกันแม้ว่าจะมาจากคนๆเดียวกันในแต่ละครั้งมีความดังไม่เท่ากันดังนั้นเราจึงต้องทำการปรับค่าแอมปลิจูดของคำพูดแต่ละคำให้เท่ากันก่อนที่จะนำไปทำการรู้จำเสียงดังรูป 2.7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



สัญญาณเสียงก่อนทำนอร์มอลไลเซชัน สัญญาณเสียงหลังทำ นอร์มอลไลเซชัน

รูป 2.7 สัญญาณเสียงก่อนและหลังทำนอร์มอลไลเซชัน

โดยการนอร์มอลไลเซชันจะใช้สมการดังข้างล่าง

$$y[k] = \frac{x]k] \times \max(x)}{Max} \tag{2.12}$$

โดยที่ y[k] คือ แอมปลิจูดใหม่ที่ k หลังทำนอร์มอลไลเซชัน

x[k] คือ แอมปลิจูดเก่าที่ k ก่อนทำนอร์มอลไลเซชัน

Max(x) คือ แอมปลิจูดสูงสุดก่อนทำนอร์มอลไลเซชัน

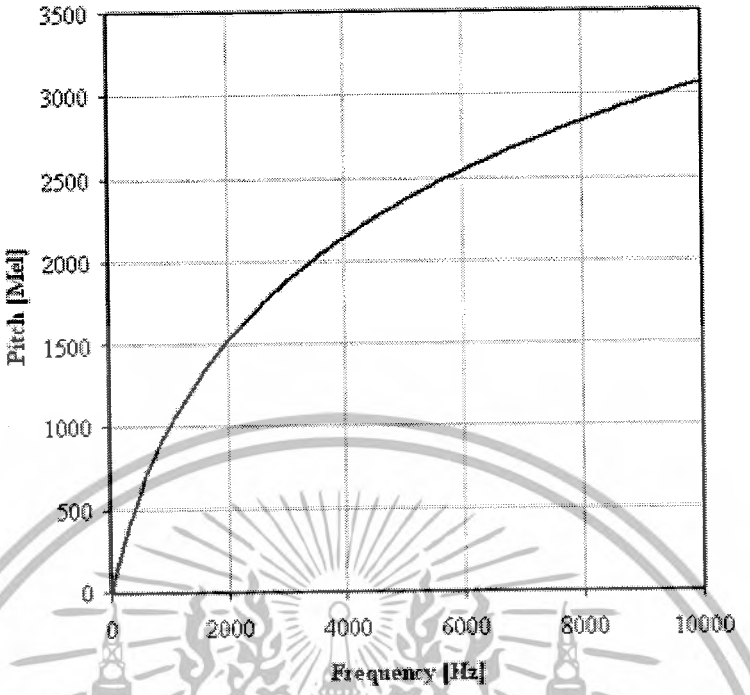
Max คือ แอมปลิจูดสูงสุดหลังทำนอร์มอลไลเซชัน

2.3.3 การหาลักษณะเด่นของเสียงพูด (Feature Extraction)

เนื่องด้วยในการรู้จำเสียงพูดนั้นไม่ได้นำสัญญาณเสียงทั้งหมดไปเปรียบเทียบ แต่จะนำเอาเฉพาะค่าสำคัญของเสียงนั้นๆ ไปใช้ในการฝึกฝนรู้จำในโครงข่ายประสาทเทียมซึ่ง กระบวนการหาค่าลักษณะเด่นที่นำมาใช้ในโครงงานนี้ คือ สัมประสิทธิ์เซปสตรอลบนความถี่เมล ซึ่งมีรายละเอียดดังนี้

สัมประสิทธิ์เซปสตรอลบนความถี่เมล เป็นกระบวนการหาค่าลักษณะเด่น ที่ปรับปรุงมาจากสัมประสิทธิ์เซปสตรัมปกติ ด้วยการปรับสเกลของสเปกตรัมให้อยู่บนสเกลที่เหมาะสมสำหรับการรับฟังของมนุษย์ เนื่องจากสัญญาณเสียงพูดในช่วงความถี่ต่ำ จะมีนัยสำคัญมากกว่าสัญญาณเสียงพูดในช่วงความถี่สูง จึงมีการออกแบบสเกลของสเปกตรัมให้สามารถเก็บรายละเอียดของสัญญาณเสียงช่วงความถี่ต่ำได้มากกว่า เรียกว่า สเกลเมล (Mel scale) ซึ่งกราฟความถี่ของสเกลเมลมีลักษณะเป็นกราฟเส้นตรงในบริเวณพื้นที่ต่ำกว่า 1,000 เฮิร์ท และเป็นกราฟลอการิทึมในบริเวณพื้นที่ที่สูงกว่า 1000 เฮิร์ทดังรูป 2.8

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 2.8 กราฟการกระจายตัวของความถี่บนสเกลเมล

ดังนั้นความถี่ของเสียงพูดจึงถูกปรับให้อยู่ในรูปสเกลเมลดังสมการ

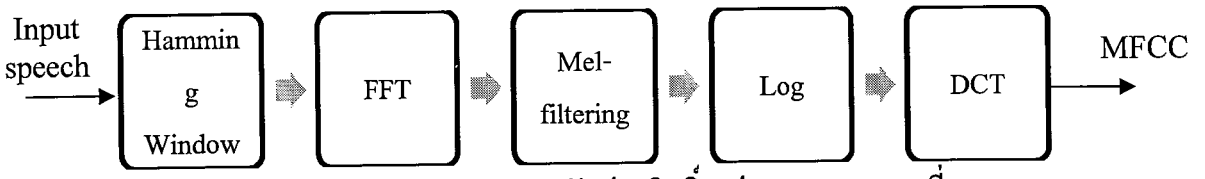
$$f_{mel} = 2595 \times \log_{10}\left(\frac{f}{700} + 1\right) \tag{2.13}$$

กระบวนการคำนวณลักษณะเด่น สัมประสิทธิ์เซปสตรอลบนความถี่เมล เริ่มต้นจากการวางแฮมมิงวินโดว์ โดยแต่ละหน้าต่างจะเลื่อนไปแบบทับซ้อนกัน สัญญาณที่ได้แต่ละหน้าต่างนำไปผ่านกระบวนการแปลงฟูริเยร์ เพื่อเปลี่ยนสัญญาณเสียงให้อยู่ในโดเมนความถี่ ปรับความถี่ที่ได้ให้อยู่ในสเกลเมลตามสมการที่ 2.13 หลังจากนั้นส่งผ่านเสียงไปยังชุดตัวกรองสามเหลี่ยม แล้วคำนวณหาค่าพลังงาน (*Ej*) ใส่ลอการิทึมแล้วนำมาแปลงโคซายน์แบบไม่ต่อเนื่องจะได้ค่าสัมประสิทธิ์เซปสตรอลบนความถี่เมล ลำดับที่ n (C_{mel}) ดังสมการ 2.14

$$C_{mel}(n) = \sum_{j=1}^{J} \log\left(E_j\right)\cos(n(j-0.5)\frac{\pi}{J}) \tag{2.14}$$

ผลลัพธ์ที่ได้จากสมการ 2.14 จะได้ค่าลักษณะเด่น สัมประสิทธิ์เซปสตรอลบนความถี่เมล ของแต่ละส่วนย่อยของเสียงซึ่งสามารถนำมาแทนลักษณะของหน่วยเสียงของภาษาได้ สำหรับลำดับการทำงานของสัมประสิทธิ์เซปสตรอลบนความถี่เมลเป็นดังรูป 2.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 2.9 แผนภาพลำดับการทำงานของสัมประสิทธิ์เซปสตรอลบนความถี่เมล

2.3.3.1 แฮมมิ่งวินโดว์

เป็นกระบวนการแบ่งเสียงออกเป็นเฟรมย่อยๆ โดยแต่ละเฟรมย่อยๆจะทับซ้อนกันดังสมการ 2.15, 2.16

$$W(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \tag{2.15}$$

โดยที่ N คือ จำนวนของข้อมูลในแต่ละเฟรม และ n มีค่าตั้งแต่ 0, 1, ...,N-1 เพื่อให้แน่ใจว่าจะไม่มีข้อมูลบริเวณขอบของเฟรมสูญหายไประหว่างการแบ่งเฟรม

2.3.3.2 การแปลงฟูริเยร์แบบเร็ว

การแปลงฟูริเยร์ ตั้งชื่อตาม โจเซฟ ฟูริเยร์ หมายถึงการแปลงเชิงปริพันธ์ โดยเป็นการเขียนแทนฟังก์ชันใดๆ ในรูปผลบวก หรือปริพันธ์ ของฐาน ที่เป็นฟังก์ชันรูปคลื่น ไซน์ หรือ โคไซน์ โดยการแปลงฟูริเยร์นั้นมีอยู่หลายประเภท แต่ประเภทที่เราเลือกนำมาศึกษา เป็นวิธีการฟูริเยร์แบบเร็วซึ่งเป็นวิธีการหนึ่งของการแปลงฟูริเยร์แบบไม่ต่อเนื่อง

การแปลงฟูริเยร์แบบไม่ต่อเนื่อง (Discrete Fourier Transform) ใช้สำหรับการคำนวณด้วยเครื่องคอมพิวเตอร์ ค่าสัญญาณในทั้งสองโดเมน (Domain) นั้นจำเป็นต้องมีค่าเป็นดิจิทัล ซึ่งคือ ฟังก์ชันค่าไม่ต่อเนื่อง x[n] บนโดเมนไม่ต่อเนื่องแทนที่จะเป็นโดเมนต่อเนื่องในช่วงจำกัด หรือ เป็นคาบ ในกรณีนี้เราจะใช้ การแปลงฟูริเยร์ไม่ต่อเนื่อง ซึ่งเขียนแทน x[n] ด้วยผลบวกของฟังก์ชันคาบ ดังสมการข้างล่าง

$$X[n] = \frac{1}{N}\sum_{K=0}^{N-1} X[k]e^{\frac{2\pi ink}{N}} \qquad n = 0,\ldots, N-1 \tag{2.16}$$

โดยที่ X[k] คือ ค่าขนาดบนโดเมนการแปลงการคำนวณจากสมการข้างต้น โดยจะใช้ความซับซ้อนในการคำนวณ O (N^2) ซึ่งสามารถลดลงเหลือเพียง O (N log N) โดยการใช้อัลกอริทึม การแปลงฟูริเยร์แบบเร็ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การแปลงฟูริเยร์แบบเร็ว เป็นขั้นตอนการนำเฟรมแต่ละเฟรมย่อยซึ่งผ่านการทำแฮมมิ่งวินโดว์แล้วมาเข้ากระบวนการแปลงให้อยู่ในรูปของโดเมนความถี่ (Frequency Domain) แล้วจึงปรับความถี่ที่ได้ให้อยู่ในสเกลเมลตามสมการที่ 2.13

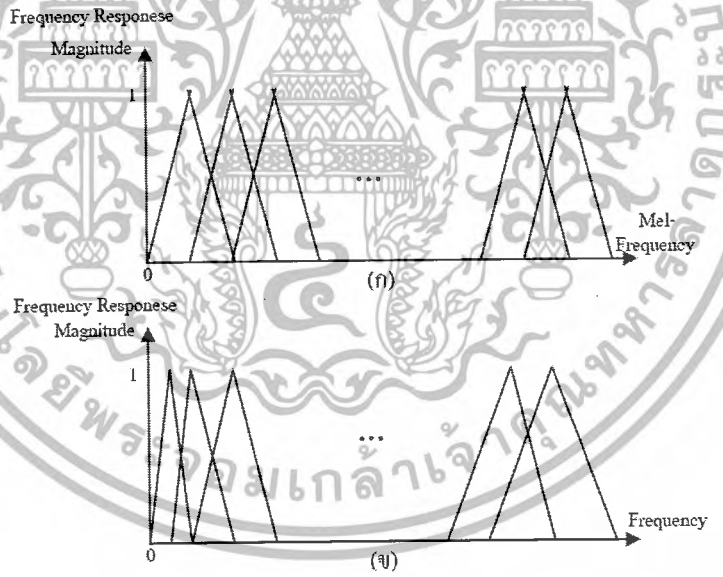
2.3.3.3 ชุดตัวกรองแบบสามเหลี่ยม

ขั้นตอนนี้เป็นการนำแต่ละความถี่ที่ได้จากการแปลงฟูริเยร์แบบเร็ว มาผ่านชุดตัวกรองแบบสามเหลี่ยมในสเกลเมล เพื่อเน้นความสำคัญของความถี่ ที่อยู่ในช่วงกลางของชุดตัวกรองแต่ละตัวกรองตามสมการที่ 2.17 ซึ่งชุดตัวกรองสามเหลี่ยมมีลักษณะดังรูปที่ 2.10

$$E_j = \sum_{k=0}^{\frac{n}{2}-1} \Phi_j \bullet (k)\left|\tilde{x}(k)\right|^2; \quad 0 \le j \le J \tag{2.17}$$

เมื่อ Φ_j คือ ค่าประจำตัวกรองที่ j

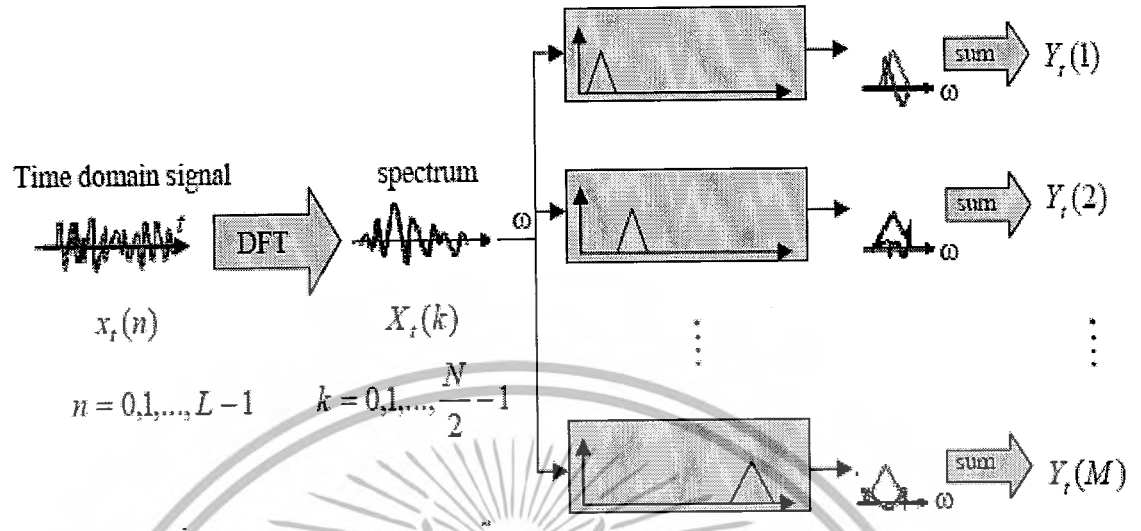
$\tilde{x}(k)$ คือ สเปคตรัม



รูป 2.10 ชุดตัวกรองแบบสามเหลี่ยม (ก) ตัวกรองบนสเกลเมล (ข) ตัวกรองบนสเกลปกติ

จากรูปที่ 2.10 (ก) ในการคำนวณสร้างชุดตัวกรองความถี่แบบสามเหลี่ยม นำความถี่ (fs/2) มาคำนวณหาความถี่สูงสุดบนสเกล แล้วนำความถี่ที่คำนวณได้มาแบ่งระยะห่างเท่าๆกันบนสเกลเมลตามขนาดตัวกรองดังรูป 2.11 โดยทั่วไปจะกำหนดตัวกรองสามเหลี่ยมแต่ละตัวให้เหลื่อมกัน 50 เปอร์เซ็นต์ ความถี่บนสเกลเมล (*mel(f)*) หาได้จากการแปลงความถี่ปกติ (f) ดังสมการที่ 2.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 2.11 สัญญาณเสียงถูกแปลงฟูริเยร์แบบเร็วและส่งผ่านชุดตัวกรองสามเหลี่ยม

2.3.3.4 การแปลงโคไซน์ไม่ต่อเนื่อง

การแปลงโคไซน์ไม่ต่อเนื่อง เป็นการแปลง ออร์ทอโกนัล (Orthogonal) ที่เป็นจำนวนจริง และมีฟังก์ชัน โคไซน์ เป็นฐานมีทั้งหมด 8 ชนิด คือ ดีซีที-1 (DCT-1) ถึง ดีซีที-4 (DCT-4) ความยาวคู่ (หรือดีซีที-ไออี (DCT-IE) ถึงดีซีที-ไอวีอี (DCT-IVE)) และ ดีซีที-5 (DCT-5) ถึง ดีซีที-8 (DCT-8) ความยาวคี่ (หรือดีซีที-ไอโอ(DCT-IO) ถึงดีซีที-ไอวีโอ (DCT-IVO))

การแปลงโคไซน์ ที่รู้จักกันมากที่สุด คือ ดีซีทีชนิดที่สองความยาวคู่ ซึ่งมักจะเรียกสั้นๆว่า "การแปลงดีซีที " และ เรียกการแปลงกลับ ซึ่งเท่ากับการแปลง ดีซีที-3 (ดีซีที-III) ว่า "การแปลงกลับดีซีที " หรือ " ไอ ดีซีที (Inverse DCT ,IDCT)"

ดีซีทีนั้นมีการประยุกต์ใช้งานที่รู้จักกันดีในการประมวลผลสัญญาณและการประมวลผลภาพ โดยเฉพาะอย่างยิ่งในการเข้ารหัสแบบแปลง (Transform Coding) เพื่อการบีบอัดข้อมูลแบบที่ไม่มีการสูญเสีย ไม่ว่าจะเป็นตามมาตรฐานการบีบอัดภาพนิ่ง เจเพก (JPEG) และมาตรฐานการบีบอัด ภาพเคลื่อนไหว เอ็มเพก (MPEG) ทั้งนี้เนื่องมาจากคุณสมบัติของ ดีซีที ที่เรียกว่า เอ็นเนอร์จี คอมแพ็กชัน (Energy Compaction) ที่ดี คือ สามารถอัดพลังงานส่วนใหญ่ของสัญญาณ โดยเฉพาะภาพไปไว้ในสัมประสิทธิ์ย่านความถี่ต่ำในโดเมนของการแปลงและการคำนวณการแปลงในทางปฏิบัติสามารถกระทำได้อย่างมีประสิทธิภาพ

โดยในโครงงานนี้ได้นำ ดีซีที มาใช้ในการแปลงแต่ละเฟรมย่อยๆ ของเสียงในรูปเชิงความถี่ให้อยู่ในรูปเชิงเวลา โดยแต่ละเฟรมจะนำมาผ่านค่าล็อก แล้วนำมาเข้าสมการ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

ดีซีที ในสมการ 2.14 ซึ่งผลลัพธ์ที่ได้ออกมาจะเป็นค่าสำคัญของเสียง ซึ่งเราจะนำไปเข้ากระบวนการรู้จำต่อไป

2.4 โครงข่ายประสาทเทียม

2.4.1 ทฤษฎีเกี่ยวกับโครงข่ายประสาทเทียมเบื้องต้น

2.4.1.1 ทฤษฎีเซลล์ประสาทเบื้องต้น

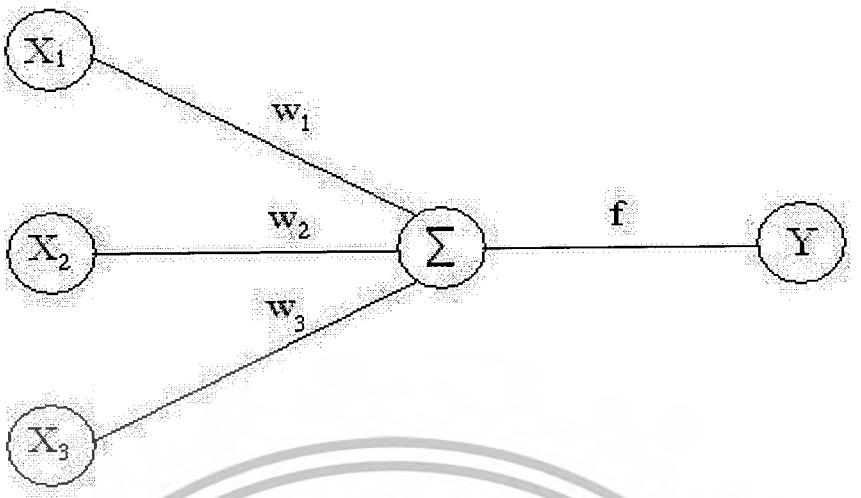
ระบบประสาทมนุษย์ประกอบด้วยเซลล์ประสาท (neuron) ประมาณ 10^{11} เซลล์ และมีการเชื่อมต่อกันระหว่างเซลล์ประมาณ 10^{15} จุด ลักษณะเซลล์ประสาททางชีวภาพประกอบด้วย 3 ส่วนหลักคือ เดนไดร์ (Dendrites) โซมา (Soma) หรือตัวเซลล์ (Cell body) และแอคซอน (Axon) ดังรูป 2.12 ส่วนที่เชื่อมระหว่างตัวเซลล์กับแอคซอนของเซลล์อื่น ๆ ที่อยู่รอบ ๆ ข้างเรียกว่า ไซแนปส์ (Synapse) โดยตัวเดนไดร์จะรับสัญญาณจากเซลล์ประสาทที่อยู่รอบ ๆ ผ่านทางไซแนปส์ ด้วยปฏิกิริยาทางเคมี ซึ่งสัญญาณที่เข้ามานี้คือประจุอิเล็กตรอน (Electron) โดยปฏิกิริยาเคมีที่เกิดขึ้นจะทำการปรับเปลี่ยนความถี่ของสัญญาณที่เข้ามา ตัวเซลล์ทำหน้าที่รวมสัญญาณที่เข้ามาแล้วทำการส่งสัญญาณออกให้กับแอคซอนเพื่อผ่านต่อให้กับเซลล์อื่น ๆ ซึ่งสัญญาณที่ออกจากตัวเซลล์มีลักษณะเป็นสัญญาณกระตุ้นให้กับเซลล์อื่น

รูป 2.12 แบบจำลองของเซลล์ประสาทในสมองมนุษย์

2.4.1.2 ทฤษฎีโครงข่ายประสาทเทียม (Artificial Neural Network)

ตัวเซลล์ที่ใช้มีคุณลักษณะพื้นฐานเหมือนกับเซลล์ประสาททางชีววิทยา โดยทั่วไปใช้พื้นฐานจากตัวแบบที่เสนอโดย แมคคาลอค (McCulloch) และ พิทส์ (Pitts) ในปี 1943 จากรูป 2.13 เซลล์ประสาทเทียมรับสัญญาณ x_1-x_n จากเซลล์รอบข้าง โดยไซแนปส์หรือตัวเชื่อมต่อจะทำการปรับเปลี่ยนหรือถ่วงน้ำหนักค่าสัญญาณที่เข้ามาด้วย w_1-w_n

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้
110955



รูป 2.13 การคำนวณค่าเอาท์พุทของโครงข่ายประสาทเทียม

$$x_i = \sum x_j w_{ij} \qquad (2.18)$$

$$y_i = f_i(x_i) \quad I = 1,2...m \qquad (2.19)$$

โดยที่ x_j คือ ค่าสัญญาณลำดับที่ j ที่เข้าสู่เซลล์ประสาทเทียมโหนดที่ i

w_{ij} คือ ค่าน้ำหนักสัญญาณลำดับที่ j ที่เข้าสู่เซลล์ประสาทเทียมโหนดที่ i

x_i คือ ค่า x อินพุทของเซลล์ประสาทเทียมโหนดที่ i

$f_i(x_i)$ คือ ฟังก์ชันกระตุ้นของเซลล์ประสาทเทียมโหนดที่ i

y_i คือ ค่าสัญญาณที่ออกจากเซลล์ประสาทเทียมโหนดที่ i

m คือ จำนวนเซลล์ประสาทเทียมของโครงข่าย

จากสมการที่ 2.18 แสดงการหาค่าผลรวมสัญญาณที่เข้าสู่เซลล์ประสาทเทียมโหนดที่ i ซึ่งเรียกว่า x_i input และค่า y_i คือค่าสัญญาณที่ออกจากเซลล์ประสาทเทียม โดยค่าสัญญาณนี้จะส่ง ผ่านทางฟังก์ชันกระตุ้น (Activated Function) $f_i(x_i)$ ซึ่งฟังก์ชันนี้ทำหน้าที่เลือกผ่านค่าสัญญาณให้กับแอคซอน โดยจะทำการย่อขนาดของผลรวมสัญญาณที่ออกจากตัวเซลล์ให้อยู่ในช่วง 0 – 1 สมการที่ 2.19 แสดงการหาค่าฟังก์ชันการกระตุ้นของสมการที่ 2.18

โหนด Y จะถูกเชื่อมไปยัง Z_1 และ Z_2 กับค่าน้ำหนัก v_1 และ v_2 ตามลำดับ โหนด Y จะส่งสัญญาณ y ไปยังแต่ละหน่วย (Unit) ที่เอาท์พุท ซึ่งค่าที่ได้รับจะแตกต่างกัน เพราะว่าในแต่ละสัญญาณจะถูกปรับแต่งให้มีค่าที่เหมาะสมตามค่าน้ำหนัก โดยขึ้นอยู่กับค่าอินพุทที่ได้รับ ซึ่งเป็นโครงข่ายประสาทเทียมแบบมีฮิดเดน ยูนิก (Hidden Unit) พร้อมกับฟังก์ชันที่ใช้ในการกระตุ้นแบบไม่เป็นเชิงเส้น ซึ่งสามารถแก้ปัญหาได้มากกว่า และมีความยากเพิ่มขึ้นสำหรับการฝึกสอนโครงข่ายในส่วนของฮิดเด้นยูนิก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าของส่วนเชื่อมต่อของเซลล์ประสาทเทียมอาจมีค่าเป็นบวก ลบ หรือศูนย์ก็ได้ ดังนั้นในการปรับแต่งค่าสัญญาณที่เข้าสู่เซลล์ประสาทด้วยค่าน้ำหนักของส่วนเชื่อมต่อ ถ้าข้อมูลและค่าส่วนเชื่อมต่อมีค่าเป็นบวก ก็จะก่อให้เกิดสภาพเสริมกัน แต่ถ้าค่าของส่วนเชื่อมต่อมีค่าเป็นลบ ก็จะลดค่าความเข้มของข้อมูลที่เข้ามา ส่วนกรณีที่ค่าของส่วนเชื่อมต่อเท่ากับศูนย์หมายถึง ส่วนเชื่อมต่อตรงส่วนนั้นไม่ทำงาน

ค่าผลลัพธ์ที่ได้จากฟังก์ชันกระตุ้นจะเป็นค่าที่อยู่ในช่วง 0 – 1 ซึ่งมีลักษณะเหมือนกับเซลล์ประสาททางชีววิทยา ถ้ามีลักษณะเป็นสถานะของการกระตุ้นเราจะสามารถแทนความหมายสถานะการกระตุ้นด้วยค่า 1 และการไม่กระตุ้น 0 รูปแบบฟังก์ชันกระตุ้นที่นิยมใช้มีดังนี้

2.4.1.2.1 สเต็ป ฟังก์ชัน (Step Function)

นิยมใช้ในโครงข่ายแบบชั้นเดียว (Single Layer) ซึ่งใช้ในการแปลงค่า X อินพุท ที่อยู่ในรูปของค่าต่อเนื่องให้อยู่ในรูปไบนารี (Binary) คือ 1 และ 0 หรือไบโพลา (Bipolar) คือ 1 และ - 1 ฟังก์ชันนี้จะใช้ค่าเทรสโฮลด์ θ ในการกำหนดการแปลงค่า เรียกว่า เทรสโฮลด์ ฟังก์ชัน (Threshold Function หรือ Heaviside Function) ดังสมการ 2.20

$$f(x) = 1 \text{ if } x \geq \theta$$
$$= 0 \text{ if } x < \theta \qquad (2.20)$$

โดยที่ x คือ ค่า x อินพุทที่ของเซลล์ประสาทเทียม

θ คือ ค่าเทรสโฮลด์ที่ใช้ในการแปลงค่าของเซลล์ประสาทเทียม

2.4.1.2.2 ซิกมอยฟังก์ชัน (Sigmoid Function, S – shaped curves)

นิยมนำมาใช้ในการแปลค่า X อินพุทให้อยู่ในรูปของช่อง 0 – 1 ซึ่งเรียกว่า ไบนารี ซิกมอย (Binary Sigmoid) หรืออยู่ในรูป -1 – 1 ซึ่งเรียกว่าไบโพลาซิกมอย (Bipolar Sigmoid) ฟังก์ชันที่ใช้ในการแปลงค่าเป็นซิกมอยฟังก์ชัน (Sigmoid Function) คือ ลอจิสติกฟังก์ชัน (Logistic Function) และ ไฮเปอร์โบลิก แทนเจน ฟังก์ชัน (Hyperbolic Tangent Function) ซึ่งฟังก์ชันเหล่านี้มีประโยชน์อย่างมากในการสอนโครงข่ายแบบแพร่ย้อนกลับ (Backpropagation Network) สามารถแสดงดังสมการ 2.21 และ 2.22

Binary sigmoid

$$f(x) = \frac{1}{1 + e^{(-\alpha x)}} \qquad (2.21)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

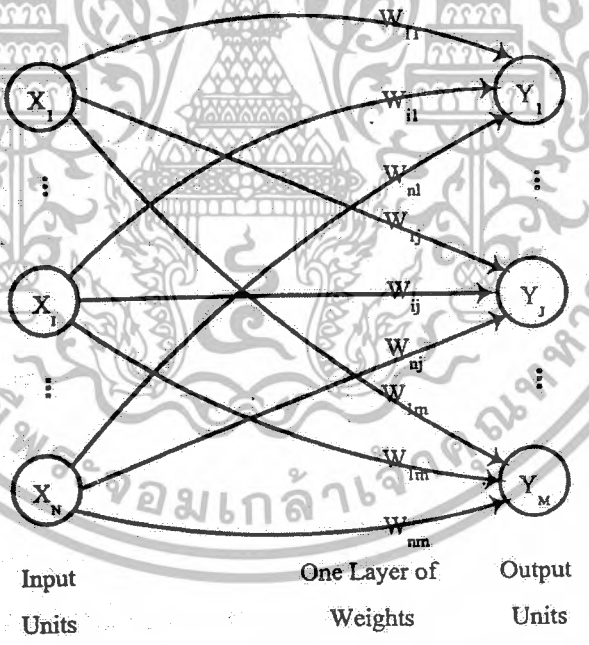
Bipolar sigmoid

$$f(x) = \frac{1 - e^{(-2x)}}{1 + e^{(-2x)}} \tag{2.22}$$

2.4.2 สถาปัตยกรรมของโครงข่ายประสาทเทียม (Neural Network Architecture)

การเชื่อมต่อโครงข่ายประสาทเทียมแบ่งออกเป็น 2 แบบคือ โครงข่ายแบบชั้นเดียว (Single Layer) และโครงข่ายแบบหลายชั้น (Multilayer) จำนวนชั้นของโครงข่ายคือ จำนวนชั้นของส่วนเชื่อมต่อที่ถูกถ่วงน้ำหนัก โครงข่ายแบบหลายชั้นสามารถแก้ปัญหาที่มีความซับซ้อนได้ดีกว่าโครงข่ายแบบชั้นเดียว

2.4.2.1 โครงข่ายประสาทเทียมแบบชั้นเดียว (Single Layer Artificial Neural Network)



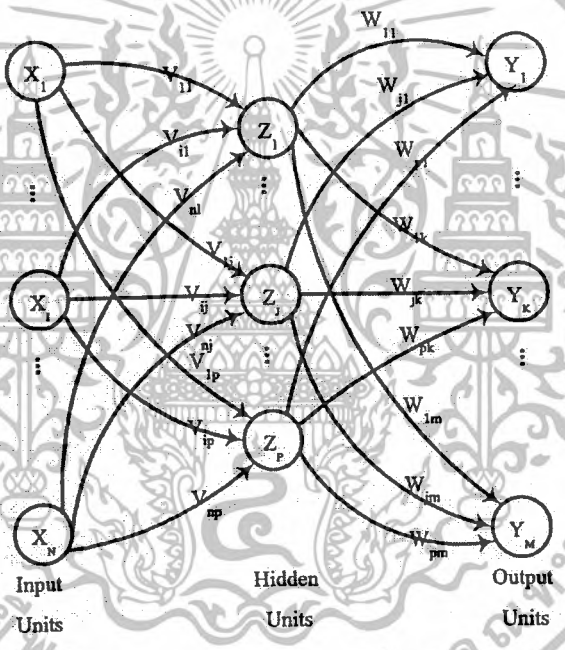
รูป 2.14 ตัวอย่างโครงข่ายประสาทเทียมแบบชั้นเดียว

จากรูป 2.14 เป็นโครงข่ายประสาทเทียมแบบชั้นเดียวที่ประกอบด้วยเซลล์ประสาทเทียมง่ายๆหลายชุด ความสามารถในการคำนวณของโครงข่ายประสาทเทียมได้จากลักษณะการเชื่อต่อเป็นโครงข่ายประสาทเทียมง่ายๆเป็นกลุ่มโมดุล (Module) ประสาทเทียมเชื่อมต่อกันเป็นชั้นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อย่างไรก็ดีลักษณะการเชื่อมโยงระหว่างโครงข่ายไม่ได้มีแค่แบบเดียวการเชื่อมโยงระหว่างชั้นอาจมีการเชื่อมโยงกลับมาในชั้นอินพุทอีก ซึ่งโครงข่ายประสาทชีวภาพก็มีลักษณะดังกล่าวเช่นกัน สำหรับค่าน้ำหนักมีวิธีพิจารณาในรูปเวจท์เมตริก (Weight matrix) ซึ่งหากโครงข่ายมีหลายชั้นจะทำให้ช่วยระบุค่าถ่วงได้ง่ายขึ้น เพื่อหลีกเหลี่ยงความสับสนจะกำหนดเป็นมิติ (Dimensions) ของเมตริก โดยให้ m แทนจำนวนแถวหรือจำนวนอินพุทและ n แทนจำนวนของเซลล์ประสาทที่สร้างขึ้นตัวอย่าง เช่น ค่าถ่วงน้ำหนักที่เชื่อมระหว่างอินพุทตัวที่ 3 กับนิวรอนตัวที่ 2 คือ $W_{3,2}$

2.4.2.2 โครงข่ายประสาทเทียมแบบหลายชั้น (Multilayer Artificial Neural Network)



รูป 2.15 ตัวอย่างโครงข่ายประสาทเทียมแบบหลายชั้น

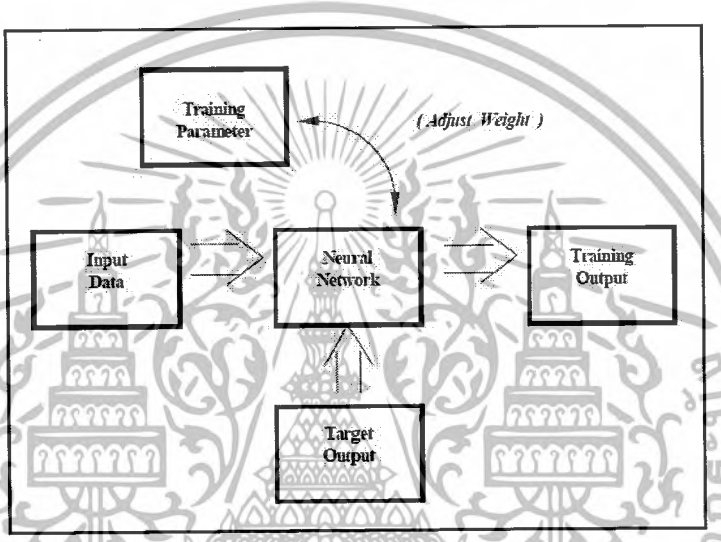
จากรูป 2.15 เป็นโครงข่ายประสาทเทียมแบบหลายชั้น โดยชั้นแรกเป็นอินพุทเลเยอร์ (Input Layer) และโหนดในชั้นนี้ได้รับสัญญาณจากภายนอก ชั้นถัดไปคือ ฮิดเดนเลเยอร์ ในชั้นนี้ประกอบด้วยประสาทที่จะรับสัญญาณมาจากอินพุทเลเยอร์และส่งสัญญาณไปยังชั้นเอาท์พุทเลเยอร์ในชั้นฮิดเดนเลเยอร์สามารถที่จะมีได้มากกว่าหนึ่งชั้น ชั้นถัดมาคือ เอาท์พุทเลเยอร์ (Output Layer) เป็นชั้นที่จะนำผลลัพธ์ส่งออกไปข้างนอก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.3 การฝึกสอนโครงข่ายประสาทเทียม

2.4.3.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

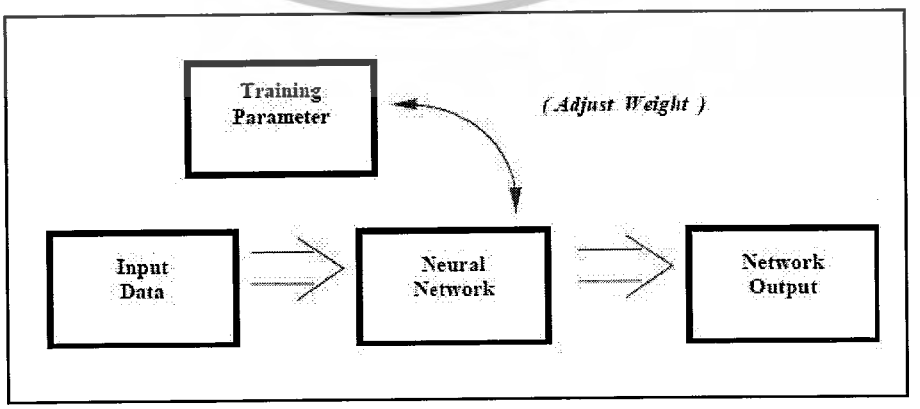
โครงข่ายชนิดนี้จะกำหนดค่าผลลัพธ์เป้าหมายให้กับข้อมูลแต่ละตัว โครงข่ายจะนำค่าที่ผิดพลาดที่ได้จากการคำนวณโครงข่ายกับผลลัพธ์ ที่ต้องการให้โครงข่ายสร้างมาเป็นค่าน้ำหนักเพื่อปรับข้อมูลให้ได้ตามที่ต้องการดังรูป 2.16 สำหรับโครงข่ายประสาทเทียมที่เป็นการเรียนรูปแบบมีผู้สอนได้แก่ แบคพรอพพาเกชั่น (Backpropagation), อะดาไลน์ (Adaline), เพอร์เซปตรอน (Perceptron)



รูป 2.16 การเรียนรู้แบบมีผู้สอน

2.4.3.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

โครงข่ายชนิดนี้จะไม่มีการส่งค่าผลลัพธ์เป้าหมายให้กับข้อมูลแต่ละตัว การปรับค่าน้ำหนักใช้ข้อมูลที่นำมาสอนเป็นตัวปรับค่าดังรูป 2.17 โดยจะปรับตามกลุ่มที่ข้อมูลเข้ามีรูปแบบคล้าย ๆ กัน อาทิเช่น ซีพีเอ็น (CPN), เออาร์ที (ART)



รูป 2.17 การเรียนรู้แบบไม่มีผู้สอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.4 กฎการเรียนรู้โครงข่ายประสาทเทียม (Neural Network Learning Rules)

ค่าน้ำหนักจะถูกเปลี่ยนแปลงได้ขึ้นอยู่กับสัญญาณอินพุทที่มันได้รับ ซึ่งทำให้มีค่าเอาท์พุทเกิดขึ้น และการร่วมกันตอบสนองของผู้สอน

2.4.4.1 กฎการเรียนรู้แบบเฮบเบรน (Hebbian Learning Rule)

ต้องการค่าน้ำหนักในการเริ่มต้นโดยการสุ่มค่าน้อย ๆ ประมาณ $w_i = 0$ ก่อนที่จะทำการเรียนรู้ เป็นการแสดงการเรียนรู้แบบไม่มีผู้สอน ถ้าการคูณกันระหว่างเอาท์พุทและอินพุท หรือ $o_i x_j$ มีค่าเป็นบวก ผลลัพธ์นี้จะเพิ่มเข้าไปในค่าน้ำหนัก w_{ij}

2.4.4.2 กฎการเรียนรู้แบบเพอร์เซปตรอน (Perceptron Learning Rule)

เป็นการเรียนรู้ถึงความแตกต่างของสัญญาณ ระหว่างความต้องการและการตอบสนองที่ถูกต้อง การเรียนรู้นี้เป็นการเรียนรู้แบบมีผู้สอนสามารถนำไปใช้สำหรับโครงข่ายประสาทเทียมที่เป็นไบนารี ค่าน้ำหนักจะถูกปรับเปลี่ยนถ้าเอาท์พุทมีค่าไม่ถูกต้อง เนื่องจากค่าที่ผิดพลาดเป็นสิ่งจำเป็นในการเรียนรู้ การปรับเปลี่ยนค่าน้ำหนัก จะมีค่าเป็นศูนย์เมื่อความต้องการและการตอบสนองถูกต้องตรงกัน

2.4.4.3 กฎการเรียนรู้แบบเดลต้า (Delta Learning Rule)

เป็นการฝึกสอนแบบมีผู้สอน โดยทำไปพร้อม ๆ กับกฎการฝึกสอนของเพอร์เซปตรอน ซึ่งจะถูกเรียกว่าเป็นกฎการฝึกสอนเพอร์เซปตรอนแบบต่อเนื่อง

2.4.4.4 กฎการเรียนรู้แบบมีการแข่งขัน (Winner – Take – All Learning Rule)

นำไปใช้สำหรับการฝึกสอนโดยไม่มีผู้ฝึกสอน ใช้เรียนรู้โดยเก็บสถิติค่าอินพุท ค่าเอาท์พุทที่ได้ ที่มีค่าสูงสุดในการตอบสนองค่าอินพุท x ซึ่งค่าเอาท์พุทที่ได้จะถูกประกาศให้เป็นผู้ชนะ

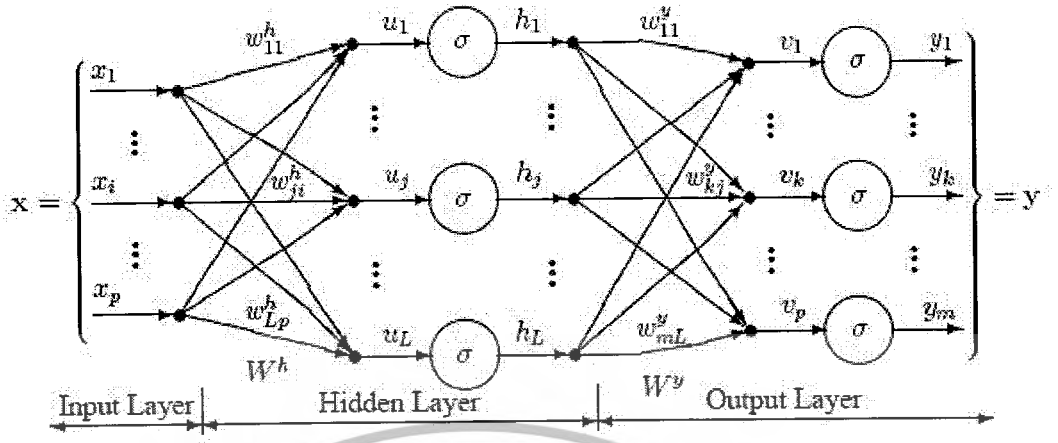
2.4.5 การนำโครงข่ายประสาทเทียมไปใช้กับการรู้จำคำพูด

ในโครงนี้เลือกใช้โครงข่ายประสาทเทียมแบบมีผู้สอน โดยมีสถาปัตยกรรมแบบมัลติเลเยอร์ ใช้กฎการเรียนรู้แบบเพอร์เซปตรอน โดยแบ่งการทำงานของโครงข่ายประเทียมออกเป็น 2 ส่วนดังนี้

2.4.5.1 ฟอร์เวิร์ดพาธ (Forward Pass)

เป็นส่วนในการคำนวณค่าผลลัพธ์จากค่าอินพุท โดยหลังจากนำสัญญาณเสียงที่เป็นคำพูดผ่านกระบวนการพรีโปรเซสซิงเรียบร้อยแล้วเอาท์พุทจากส่วนพรีโปรเซสซิงจะเป็นอินพุทของโครงข่ายประสาทเทียมที่จะนำมาใช้คำนวณหาเอาท์พุทดังรูป 2.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป2.18 โครงสร้างของโครงข่ายประสาทเทียมส่วนฟอร์เวิร์ดพาธ

สมการที่ใช้คำนวณผลลัพธ์จากอินพุทเลเยอร์ไปยังฮิดเดนเลเยอร์เป็นดังนี้

$$u_j(p) = \sum x_i(p) \times w_{ij}(p) - \theta_j \tag{2.23}$$

สมการซิกมอยฟังก์ชัน

$$h_j(p) = \frac{1}{1 + e^{-u_j(p)}} \tag{2.24}$$

โดย u_j คือ ค่าอินพุทไปยังโหนดที่ j

x_i คือ ค่าอินพุทตัวที่ i

w_{ij} คือ ค่าเวจท์ (Weight) ของอินพุทตัวที่ i ไปยังโหนดที่ j

h_j คือ ค่าเอาท์พุทของโหนดที่ j ในฮิดเด้นเลเยอร์

p คือ รอบที่ของการฝึกสอน

สมการที่ใช้การคำนวณผลลัพธ์จากฮิดเดนเลเยอร์ไปยังเอาท์พุทเลเยอร์

$$v_k(p) = \sum h_j(p) \times w_{jk}(p) - \theta_k \tag{2.25}$$

สมการซิกมอยฟังก์ชัน

$$y_k(p) = \frac{1}{1 + e^{-v_k(p)}} \tag{2.26}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดย v_k คือ ค่าอินพุทไปยังโหนดที่ k ในเอาท์พุทเลเยอร์

h_j คือ ค่าเอาท์พุทของโหนดที่ j ในฮิดเด้นเลเยอร์

w_{jk} คือ ค่าเวจท์จากโหนดที่ j ไปยังโหนดที่ i

y_k คือ เอาท์พุทโหนดที่ k ของเอาท์พุทเลเยอร์

2.4.5.2 รีเวิร์ดพาธ (Reverse Pass)

เป็นส่วนในการปรับค่าเวจท์เพื่อให้ได้ค่าเอาท์พุทที่ถูกต้องยิ่งขึ้น โดยในส่วนนี้จะมีทิศทางในการคำนวณตรงกันข้ามกับส่วนฟอร์เวิร์ดพาธ โดยจะเริ่มคำนวณจากชั้นเอาท์พุทเลเยอร์ไปสิ้นสุดที่ชั้นอินพุทเลเยอร์

2.4.5.2.1 การปรับค่าเวจท์ในชั้นฮิดเด้นเลเยอร์ไปยังเอาท์พุทเลเยอร์

ก่อนที่จะทำการคำนวณหาค่าเวจท์ใหม่ได้นั้นจำเป็นต้องทำการหาค่าความผิดพลาด (error) และสัมประสิทธิ์ความผิดพลาดซะก่อนด้วยสมการข้างล่าง

$$e_k(p) = y_{d,k} - y_k(p) \tag{2.27}$$

$$\delta_k(p) = \frac{\partial y_k(p)}{\partial v_k(p)} \times e_k(p) \tag{2.28}$$

โดย e_k คือ ค่าความผิดพลาดของโหนดที่ k ในเอาท์พุทเลเยอร์

y_{dj} คือ ค่าที่เราคาดหวังของเอาท์พุทของโหนดที่ k ในเอาท์พุทเลเยอร์

y_k คือ ค่าเอาท์พุทที่ได้จริงของโหนดที่ k ในเอาท์พุทเลเยอร์

p คือ รอบที่ของการฝึกสอน

δ_k คือ ค่าสัมประสิทธิ์ความผิดพลาดของโหนดที่ k ในเอาท์พุทเลเยอร์

เมื่อเรานำสมการที่ (2.24) ไปแทนในสมการที่ (2.28) จะได้สมการดังนี้

$$\delta_k(p) = y_k(p) \times [1 - y_k(p)] \times e_k(p) \tag{2.29}$$

จากนั้นนำค่าสมประสิทธิ์ความคาดเคลื่อนเพื่อหาค่าเวจท์ที่จะใช้สำหรับการฝึกสอบในรอบถัดไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\Delta w_{jk}(p) = \alpha \times \delta_k(p) \times h_j(p) \tag{2.30}$$

$$w_{jk}(p+1) = w_{jk}(p) + \Delta w_{jk}(p) \tag{2.31}$$

โดย Δw_{jk} คือ ค่าผลต่าเวจท์ของรอบนี้กับรอบถัดไป

α คือ อัตราการเรียนรู้ของโครงข่ายประสาทเทียม

2.4.5.2.2 การปรับค่าเวจท์ในชั้นอินพุทเลเยอร์ไปยังฮิ้ดเดนเลเยอร์

สำหรับการคำนวณหาค่าเวจท์ใหม่ในชั้นอินพุทเลเยอร์ไปฮิดเดนเลเยอร์นั้นแตกต่างจากการคำนวณค่าเวจท์ใหม่จากชั้นฮิดเดนเลเยอร์ไปยังเอาท์พุทเลเยอร์ตรงที่แต่ละโหนดของชั้นฮิดเดนเลเยอร์จะไม่มีค่าเอาท์พุทที่คาดหวังไว้ ดังนั้นสมการที่ใช้สำหรับการหาสัมประสิทธิ์ความผิดพลาดจะแตกต่าออกไป โดยสมการที่ใช้หาสัมประสิทธิ์ความผิดพลาดเป็นดังนี้

$$\delta_j(p) = h_j(p) \times [1 - h_j(p)] \times \sum(\delta_k(p) \times w_{jk}(p)) \tag{2.32}$$

โดยสมการที่ใช้คำนวณหาสัมประสิทธิ์ความผิดพลาดจะใช้ค่า $\sum(\delta_k(p) \times w_{jk}(p))$ แทนค่าความผิดพลาดสำหรับสมการในการหาค่าเวจท์ใหม่ก็ยังคงเหมือนเดิมคือ

$$\Delta w_{ij}(p) = \alpha \times \delta_j(p) \times x_i(p) \tag{2.33}$$

$$w_{ij}(p+1) = w_{ij}(p) + \Delta w_{ij}(p) \tag{2.34}$$

โดย δ_j คือ ค่าสัมประสิทธิ์ความผิดพลาดของโหนดที่ j ในชั้นฮิดเดนเลเยอร์

Δw_{ij} คือ ค่าผลต่างเวจท์ของรอบนี้กับรอบถัดไป

2.4.5.3 การออกแบบโครงข่ายประสาทเทียม

เนื่องจากโครงงานนี้ ต้องการให้โครงข่ายประสาทเทียมรู้จำคำ100พยางค์ โดยจำนวนเอาท์พุทโหนดของโครงข่ายประสาทเทียมที่เหมาะสม ควรมีจำนวนเท่ากับจำนวนของสิ่งที่ต้องการให้โครงข่ายประสาทเทียมเรียนรู้ ดังนั้นจำนวนเอาท์พุทโหนดของโครงข่ายประสาทเทียมที่ใช้ในโครงงานนี้จะมีจำนวนเอาท์พุทโหนดเท่ากับ 100 โหนด สำหรับจำนวนอินพุทในชั้นอินพุทเลเยอร์นั้นขึ้นอยู่กับเอาท์พุทของขบวนการพรีโปรเซสซิง ว่าจะได้กี่เอาท์พุท สำหรับโครงงานนี้เลือกใช้อินพุทโหนด 280 โหนด ฮิดเดนโหนด 560 โหนด และเอาท์พุทโหนด 100 โหนด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.6 การแปลค่าเอาท์พุทของโครงข่ายประสาทเทียม

สำหรับค่าเอาท์พุตจากโหนดหนึ่งโหนด จะหมายถึงพยางค์ภาษาไทยหนึ่งพยางค์ ในการแปลค่าเอาท์พุท จึงใช้วิธีการหาค่าเอาท์พุทที่มีค่ามากที่สุดในโหนดทุกโหนด และแปลผลเป็นพยางค์ประจำโหนดนั้น ตัวอย่างเช่น โครงข่ายประสาทเทียมสำหรับการรู้จำคำไทย 5 พยางค์ดังนี้

"กลับ" "คุณ" "ฉัน" "เขา" "เธอ"

โดยให้ โหนด(0) คือ พยางค์ "กลับ"

โหนด(1) คือ พยางค์ "คุณ"

โหนด(2) คือ พยางค์ "ฉัน"

โหนด(3) คือ พยางค์ "เขา"

โหนด(4) คือ พยางค์ "เธอ"

ว่ามีค่าเอาท์พุทดังข้างล่าง

โหนด(0) = 0.26

โหนด(1) = 0.84

โหนด(2) = 0.33

โหนด(3) = 0.12

โหนด(4) = 0.08

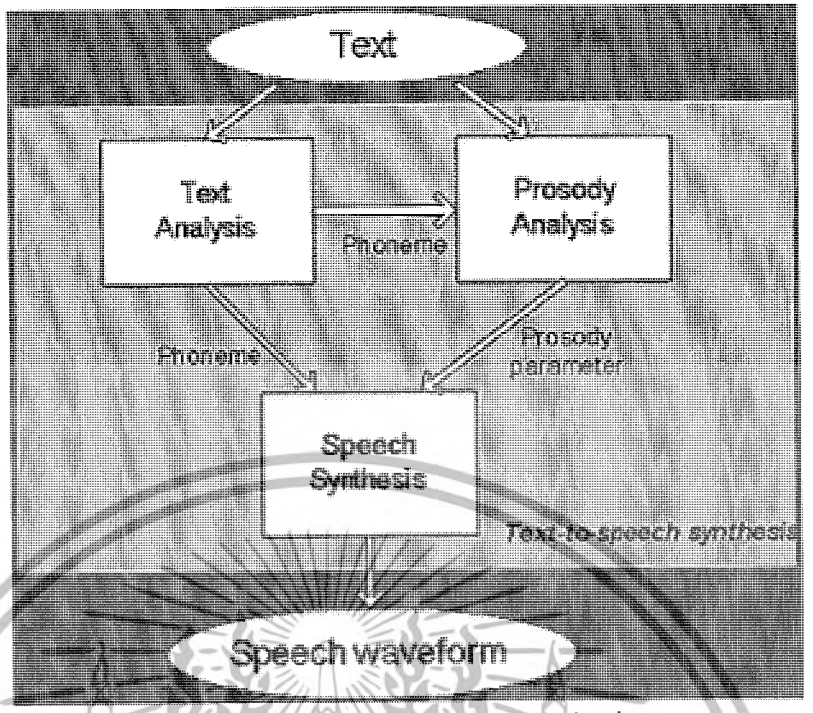
พบว่าโหนด(1) มีค่ามากที่สุดดังนั้นจึงแปลผลออกมาได้เป็นพยางค์ "คุณ"

แต่หากทำการแปลผลตามวิธีข้างต้นแล้ว จะพบว่าเมื่อนำโครงข่ายประสาทเทียมไปใช้งานจริงหากมีการป้อนค่าสัญญาณเสียงที่อยู่นอกเหนือจากพยางค์ 100 พยางค์ที่ทำการฝึกสอนโครงข่ายประสาทเทียมจะพบว่าเมื่อนำค่าดังกล่าวมาแปลผลแล้ว ยังได้ผลลัพธ์เป็นพยางค์ใดพยางค์หนึ่งใน 100 พยางค์ดังนั้นเราอาจจะต้องมีการกำหนดค่าเทรสโฮลด์ ขึ้นเมื่อค่าเอาท์พุทสูงสุดของแต่ละโหนดมีค่าน้อยกว่าค่าเทรสโฮลด์ จึงให้ทำการแปลผลว่าพยางค์ที่รับเข้ามาไม่ตรงกับพยางค์ใดใน 100 พยางค์

2.5 ระบบการแปลงตัวอักษรเป็นเสียง

ลักษณะของการเขียนประโยคในภาษาไทยจะใช้วิธีการนำคำมาเรียงต่อกันโดยจะมีการเว้นวรรคเมื่อจบประโยคเท่านั้น ซึ่งต่างจากรูปประโยคในหลายๆภาษาเช่น ภาษาอังกฤษ ภาษาญี่ปุ่น ที่ไม่เขียนคำติดกัน ทำให้การนำตัวอักษรมาแปลงเป็นเสียงค่อนข้างง่ายกว่าภาษาไทย สำหรับภาษาที่มีการเขียนคำติดกันการแปลงเสียงเป็นตัวอักษรจะประกอบด้วย 3 ส่วนหลักๆคือ ส่วนการวิเคราะห์ข้อความ (Text Analysis) ส่วนการวิเคราะห์สัทสัมพันธ์ (Prosody Analysis) ส่วนการสังเคราะห์เสียง (Speech Synthesis) ซึ่งสามารถเขียนเป็นโครงสร้างได้ดังรูป 2.19

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 2.19 โครงสร้างของระบบแปลงตัวอักษรเป็นเสียง

2.5.1 ส่วนการวิเคราะห์ข้อความ (Text Analysis)

ส่วนนี้จะมีหน้าที่วิเคราะห์ข้อความอินพุทเพื่อแปลงเป็นข้อมูล เสียงอ่าน (Phoneme) ของคำนั้น และส่งต่อให้ส่วนของการ สังเคราะห์เสียง (Speech synthesis) ต่อไป นอกจากนี้ส่วนนี้ยังทำหน้าที่อย่างอื่น ได้แก่

1) การแบ่งประโยคจากข้อความที่ยาว (Sentence breaking)
2) การทำข้อความให้อยู่ในรูปปกติ (Text Normalization) ได้แก่ การแปลงตัวเลข, คำย่อและเครื่องหมายอื่นๆ ที่ไม่ใช้ข้อ ความให้ กลายเป็นข้อความ
3) การหาขอบเขตของวลีของการอ่านในประโยคอื่นๆ

2.5.2 ส่วนการวิเคราะห์สัทสัมพันธ์

ส่วนนี้ทำหน้าที่ในการวิเคราะห์และสังเคราะห์ข้อมูลสัทสัมพันธ์ของประโยคใดๆ จากข้อมูลเสียงอ่าน และข้อความ ข้อมูลสัทสัมพันธ์ ที่วิเคราะห์ ออกมาได้ในระบบทั่วไปได้แก่

1) เซ็กเมนท์ ดูเรชัน (Segment Duration) คือความยาวของเสียงย่อยที่ต้องการสังเคราะห์ ค่านี้จะ มีผลต่อ จังหวะของเสียงที่ทำการสังเคราะห์ เช่น ถ้ากำหนด ให้ค่า ความ ยาว ของเสียงย่อยที่ต้องการสังเคราะห์มี ขนาดสั้น เสียงที่ทำ การสังเคราะห์ก็จะเหมือนกับการพูดเร็ว
2) พิทช์ คอนทัวร์ (Pitch Contour) คือ ค่าความสัมพันธ์ของความถี่มูลฐานกับเวลา ค่านี้จะมีผลต่อเสียงสูง ต่ำ (intonation) ของประโยคนั้นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.3 ส่วนการสังเคราะห์เสียง

ส่วนนี้ทำหน้าที่ในการสร้างสัญญาณคลื่นเสียง จากข้อมูลเสียง อ่าน (Phonetic transcription) และข้อมูลสัทสัมพันธ์ (Prosody transcription) จากข้อ 1 และ 2 และส่งออกสู่ลำโพง เพื่อให้เราได้ยินเสียงพูดประโยคนั้นๆ โดยทั่วไป ส่วนนี้สามารถแบ่งตามเทคนิควิธีการสังเคราะห์เสียง ได้ 3 ประเภท คือ

2.5.3.1 ฟอร์แมน ซินเทอะซิส (Formant Synthesis)

เทคนิคการสังเคราะห์วิธีแบบนี้ ข้อมูลเสียงอ่านใดๆ จะถูกกำหนด ไว้อยู่ในรูป ของความถี่ฟอร์แมนท์ต่างๆ (F1, F2, F3) ของเสียงนั้นๆ เมื่อต้องการ สังเคราะห์ เสียงใดๆ ก็นำข้อมูลเหล่านี้มาทำการสังเคราะห์ให้ เป็นสัญญาณ เสียง ซึ่งวิธีการนี้จะมีข้อดี ที่สามารถควบคุมค่าความเปลี่ยนแปลงของ ความถี่ฟอร์แมนท์ (Formant transition) ที่บริเวณรอย ต่อระหว่าง เสียงได้ง่าย แต่มีข้อเสีย คือ การจะแทนเสียงใดๆด้วยค่าฟอร์แมนท์ทำได้ยาก จะต้องมีกฎในการสังเคราะห์เสียงใดๆจำนวนมาก และเสียงที่ สังเคราะห์ออก มาได้จะไม่ค่อยเป็นธรรมชาติ ตัวอย่างของระบบแบบนี้ได้แก่ มิททอล์ค (MITALK) และ เดซทอล์ค (DECTALK)

2.5.3.2 อาร์ทิคิวเลชัน ซินเทอะซิส (Articulation Synthesis)

สำหรับวิธีนี้ข้อมูลเสียงที่ต้องการสังเคราะห์โมเดลจะอยู่ในรูปของค่าพารามิเตอร์ของโครงสร้างทางกายภาพของการเคลื่อนไหวของอวัยวะในช่องปากที่ทำให้เกิดเสียงต่างๆวิธีการนี้ค่อนข้างยากในแง่การ โมเดล เสียงต่างๆซึ่งจะต้องศึกษาจากอวัยวะในการออกเสียงจริงๆ

2.5.3.3 คอนแคททิเนชัน ซินเทอะซิส (Concatenation synthesis)

เทคนิควิธีการนี้เป็นวิธีที่ได้รับความนิยมมากที่สุดในปัจจุบัน โดยเสียงที่ทำการสังเคราะห์ขึ้น นี้เกิดจากการนำหน่วยเสียงย่อยที่ทำการเก็บไว้ก่อนแล้วมาต่อกันเป็นเสียงพูดที่ต้องการ โดยทั่วไปหน่วยเสียงย่อยที่ทำเก็บ ไว้จะอยู่ระดับต่ำกว่าคำ เช่น หน่วยของเสียงพยางค์ หน่วยของเสียง ครึ่งพยางค์ (demisyllable) หน่วยของเสียงคู่เสียง (diphone) เป็นต้น ซึ่งในโครงงานนี้ได้เลือกใช้เทคนิควิธีการนี้ โดยใช้หน่วยของเสียงครึ่งพยางค์เป็นพื้นฐาน

สำหรับโครงงานนี้เน้นหนักไปที่การแปลงเสียงไปเป็นตัวอักษร ส่วนการแปลงตัวอักษรไปเป็นเสียงเป็นเพียงแค่ส่วนเสริมเท่านั้น ดังนั้นในโครงงานนี้จึงเน้นไปที่ส่วนการสังเคราะห์เสียงซึ่งเลือกใช้ วิธีคอนแคททิเนชัน ซินเทอะซิส และส่วนการวิเคราะห์ข้อความ (Text Analysis) จะใช้วิธีนำคำมาเทียบกับกับพจนานุกรมเพื่อแยกคำออกมา สำหรับส่วนวิเคราะห์สัทสัมพันธ์ขอละเอาไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การออกแบบและพัฒนา

3.1 ความต้องการของระบบ

3.1.1 ฟังก์ชัน (Function)

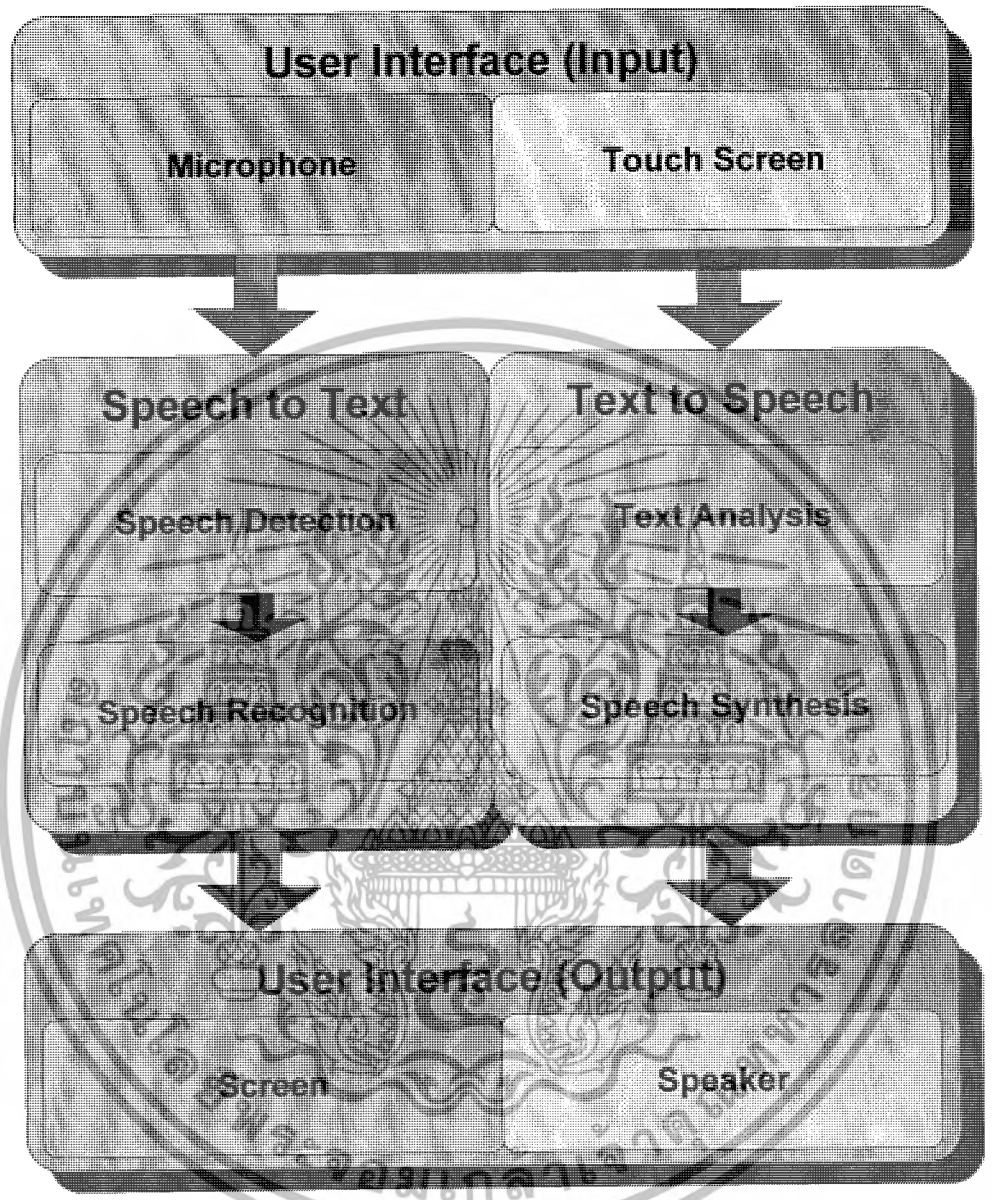
1) สามารถแปลงเสียงไปเป็นข้อความได้
2) สามารถแปลงข้อความไปเป็นเสียงได้
3) สามารถแยกพยางค์ในภาษาไทยได้
4) สามารถรับเสียงผ่านทางไมโครโฟนได้
5) สามารถบันทึกบทสนทนาเป็นข้อความได้
6) สามารถเลือกขนาด ของตัวอักษรที่จะแสดงบนหน้าจอได้

3.1.2 นอนฟังก์ชัน (Non-function)

1) สามารถจำเสียงพยางค์ภาษาไทยได้ 100 พยางค์
2) สามารถเล่นเสียงพยางค์ภาษาไทยได้ 100 พยางค์
3) สามารถปรับขนาดตัวอักษรได้ 12 ขนาด
4) บันทึกบทสนทนาเป็นตัวอักษรได้ในรูปของไฟล์นามสกุล .txt เท่านั้น
5) เสียงที่ใช้ตอบกลับมีให้เลือกได้ 2 แบบเสียงคือ เสียงผู้หญิงและเสียงผู้ชาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ภาพรวมของระบบ



รูป 3.1 แผนภาพโดยรวมของระบบ

จากรูป 3.1 สามารถแบ่งได้เป็น 3 ส่วนหลักๆ คือ

3.2.1 ส่วนติดต่อผู้ใช้งาน (User Interface)

เป็นตัวกลางในการติดต่อระหว่างผู้ใช้กับเครื่องพีดีเอ ซึ่งแบ่งออกเป็นอินพุทของส่วนติดต่อผู้ใช้งาน (User Interface Input) และเอาท์พุทของส่วนติดต่อผู้ใช้งาน (User Interface Output)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2 ส่วนการแปลงเสียงเป็นตัวอักษร

เป็นกระบวนการนำสัญญาณเสียงที่รับผ่านทางส่วนติดต่อผู้ใช้งานไปแปลผลให้เป็นข้อมูลที่อยู่ในรูปตัวอักษรที่คนพิการทางหูสามารถเข้าใจได้ ซึ่งในส่วนนี้จะประกอบด้วยส่วนย่อย 2 ส่วนคือ ส่วนการตรวจจับสัญญาณเสียงที่เป็นคำพูดกับส่วนรู้จำสัญญาณเสียงที่เป็นคำพูด

3.2.3 ส่วนการแปลงตัวอักษรเป็นเสียง

เป็นกระบวนการนำตัวอักษรที่รับผ่านส่วนติดต่อผู้ใช้งานไปสังเคราะห์สัญญาณเสียงออกมา ซึ่งส่วนนี้ประกอบด้วยส่วนย่อย 2 ส่วนคือ ส่วนการวิเคราะห์ข้อความและส่วนการสังเคราะห์สัญญาณเสียง

3.3 การออกแบบอัลกอริทึม

3.3.1 การออกแบบส่วนการแปลงเสียงเป็นตัวอักษร

แบ่งการออกแบบเป็น 2 ส่วนดังนี้

3.3.1.1 การออกแบบกระบวนการตรวจจับสัญญาณเสียงที่เป็นคำพูด

เนื่องจากเสียงพื้นหลังในสถานที่ที่แตกต่างกัน หรือแม้แต่สถานที่เดียวกันแต่คนละช่วงเวลามีคุณลักษณะที่แตกต่างกันไปดังนั้นหากต้องการให้กระบวนการตรวจจับสัญญาณเสียงที่เป็นคำพูดมีความยืดหยุดจึงควรใช้ค่าจากสัญญาณเสียงที่เป็นพื้นหลัง ณ สถานที่นั้นๆ เป็นตัวเปรียบเทียบว่าส่วนไหนควรจะเป็นสัญญาณเสียงที่เป็นคำพูด ซึ่งจากบทที่แล้วได้พูดถึงการแปลค่าสัญญาณเสียงด้วยวิธีซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอจี โดยใช้สมการทีเจอร์ เอ็นเนอจี ในทำให้ข้อมูลอยู่ในรูปที่ง่ายต่อการคำนวณ ซึ่งในบทนี้จะกล่าวถึงการออกแบบวิธีการนำค่าทีเจอร์ เอ็นเนอจีของเสียงพื้นหลังมาตรวจจับสัญญาณเสียงที่เป็นคำพูด

3.3.1.1.1 กระบวนการตรวจจับเสียงที่เป็นคำพูด

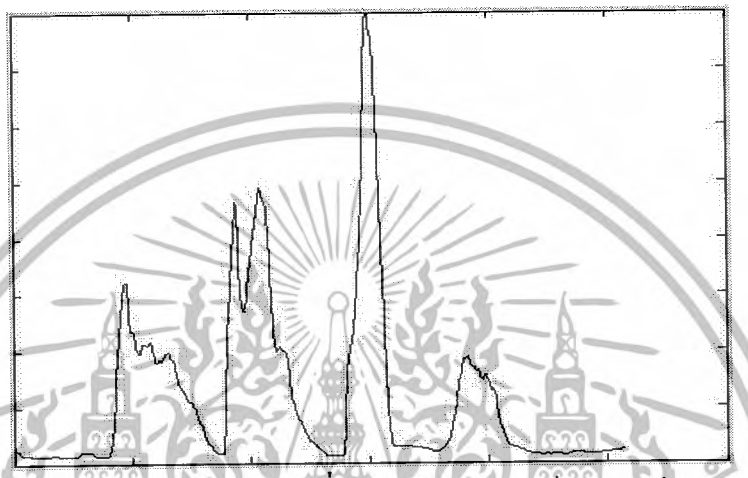
หลังจากรับสัญญาณเสียงเข้ามาจะนำอินพุทมาตรวจจับเสียงที่เป็นคำพูดด้วยวิธีซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอจีซึ่งประกอบด้วย

1) พรีโปรเซสซิ่งในโครงงานนี้จะใช้เฟรมขนาด 512 ตัวอย่าง โดยแต่ละ เฟรมจะซ้อนทับกัน 3/4 ของเฟรม ที่อัตราสุ่ม 11.25 กิโลเฮิร์ท ควอนติเซชั่น 16 บิท
2) ทีเจอร์ เอ็นเนอจี จะนำค่าในแต่ละเฟรมมาคำนวณหาค่าเฉพาะ
3) มูฟวิง เอเวอร์เรจ ทำการปรับค่าเฉพาะแต่ละเฟรม เพื่อให้ราบเรียบยิ่งขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.1.1.2 ธรรมชาติของเสียงเมื่อผ่านสมการทีเจอร์ เอ็นเนอจี

เสียงพื้นหลังเมื่อผ่านสมการทีเจอร์ เอ็นเนอจีจะมีค่าเป็นบวกเสมอโดยแต่ค่าที่ได้ในแต่ละเฟรมมิได้มีค่าเท่ากันหมด เพียงแต่จะมีค่าใกล้เคียงกันซึ่งมีความแตกต่างไปจากช่วงเสียงที่เป็นคำพูดเมื่อผ่านสมการทีเจอร์ เอ็นเนอจีจะมีความแตกต่างที่เห็นได้อย่างชัดเจนดังรูป 3.2



รูป 3.2 สัญญาณเสียงเมื่อผ่านสมการ ทีเจอร์ เอ็นเนอจี

จากรูป 3.2 จะเห็นว่าค่าทีเจอร์ เอ็นเนอจีในแต่ละเฟรมของสัญญาณเสียงที่เป็นคำพูดจะมีค่ามากกว่าสัญญาณเสียงที่เป็นพื้นหลัง ดังนั้นในกระบวนการนี้จึงนำค่าทีเจอร์ เอ็นเนอจีของเสียงพื้นหลังมาคำนวณค่าเทรดโฮล์ดโดยวิธีการคูณกับค่าถ่วงน้ำหนักเพื่อเปรียบเทียบค่า เหตุผลที่ใช้การคูณค่าถ่วงน้ำหนักแทนการบวกค่าเพิ่มเติมเข้าไปเพราะจากการศึกษาเสียงพื้นหลังในสภาพแวดล้อมที่แตกต่างกันพบว่ายิ่งทีเจอร์ เอ็นเนอจีของสัญญาณเสียงพื้นหลังมีค่าสูง ความแตกต่างของทีเจอร์ เอ็นเนอจีของแต่ละสัญญาณเสียงก็มีค่ามากขึ้นตามไปด้วย ดังนั้นการนำค่าของสัญญาณเสียงพื้นหลังมาหาค่าเทรสโฮลด์ โดยการคูณค่าถ่วงน้ำหนักจึงเหมาะสมกว่าการนำไปบวกกับค่าคงที่

นอกจากนี้ในระหว่างที่รับสัญญาณเสียงเข้ามาอาจมีเสียงที่ไม่ต้องการเข้ามาด้วย เช่น เสียงพูดของคนรอบข้าง เสียงของหล่น เป็นต้น ซึ่งเสียงเหล่านี้เป็นเสียงที่เกิดเป็นครั้งคราวทำให้เมื่อตัดเสียงออกมา เสียงเหล่านี้อาจถูกตัดออกมาด้วยทั้งๆที่ไม่ต้องการ ดังนั้นหากใช้ค่าเทรสโฮลด์ เปรียบเทียบเพียงค่าเดียวอาจทำให้เสียงเหล่านี้ถูกตัดออกมาเยอะ เพราะเหตุนี้ในโครงงานนี้จึงใช้ค่าเทรสโฮลด์ 2 ค่าด้วยกันในการตรวจจับ โดยอาศัยหลักธรรมชาติที่ได้จากการสังเกตว่าเสียงที่เป็นคำพูดเมื่อผ่านทีเจอร์ เอ็นเนอจี จะได้เป็นรูปภูเขา กล่าวคือ ณ จุดเริ่มต้นเป็นต้นไป ค่าจะสูงขึ้นเรื่อยๆ จนถึงจุดๆหนึ่งค่าจะต่ำลงจนถึงจุดสิ้นสุดคำพูด

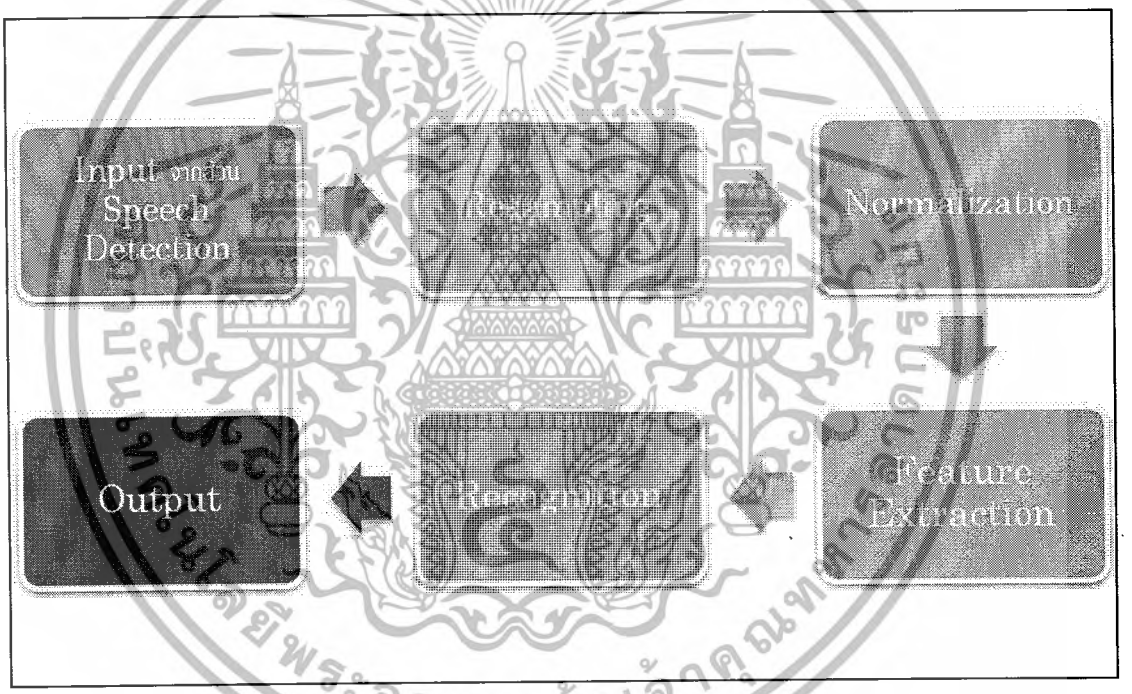
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.1.1.3 อัลกอริทึมในการตัดสัญญาณเสียงที่เป็นคำพูด

อย่างที่ได้กล่าวไว้ในหัวข้อที่แล้วว่าการตัดเสียงควรใช้ค่าเทรสโฮลด์ 2 ค่า โดยที่ค่าสองค่าเอามาจากค่าทีเจอร์ เอ็นเนอจีของสัญญาณเสียงพื้นหลังเอามาคูณกับค่าถ่วงน้ำหนักที่แตกต่างกันซึ่งเรียกค่าเทรสโฮลด์ที่มีค่าน้อยว่าค่าเทรสโฮลด์ โลว์และค่าเทรสโฮลด์ ที่มีค่ามากว่าเทรสโฮลด์ ไฮสำหรับอัลกอริทึมที่ใช้ในการตรวจจับสัญญาณเสียงที่เป็นคำพูดสามารถเขียนเป็นโฟลว์ ชาร์ต (Flow Chart) ดังรูป 3.4 หน้า 35

3.3.1.2 การออกแบบกระบวนการรู้จำ (Speech Recognition) ส่วนการแปลผลฟอร์เวิร์ดพาธ

ขบวนการรู้จำเสียงส่วนการแปลผลฟอร์เวิร์ดพาธเป็นดังรูป 3.3



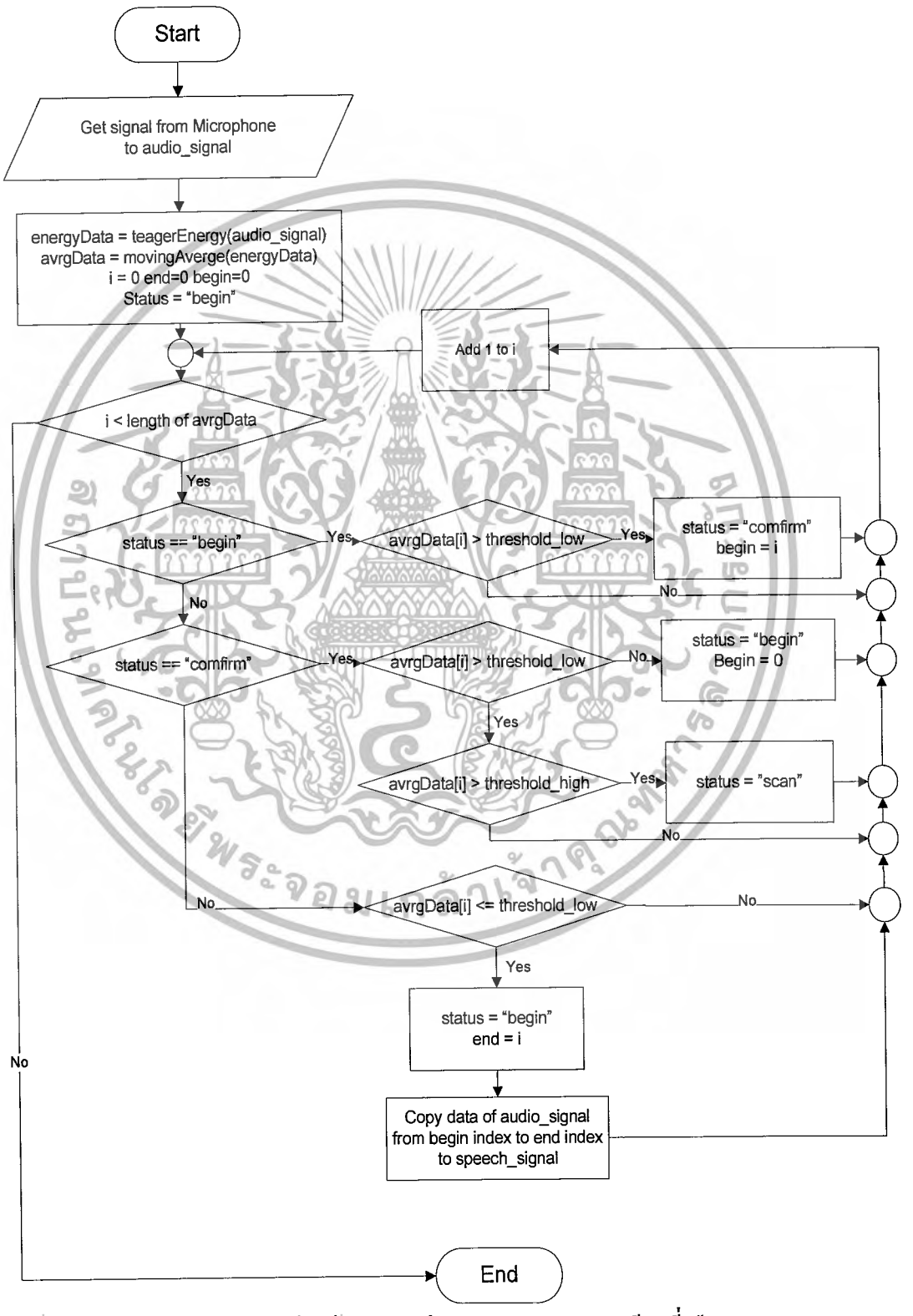
รูป 3.3 กระบวนการรู้จำ

เนื่องจากคำในเสียงภาษาไทยมีหลายคำแต่ด้วยธรรมชาติของโครงข่ายปราสาทเทียมเอาท์พุทที่ได้มีเท่ากับจำนวนคำดังนั้นในโครงงานนี้หากมีการพูดคำที่ไม่ได้อยู่ใน 100 คำ โครงข่ายประสาทเทียมไม่สามารถที่จะปฏิเสธได้ แต่ระบบที่ดีควรจะสามารถบอกได้ว่าคำๆนั้นมิใช่คำใดเลยใน 100 คำ จึงต้องมีการออกแบบการแปลผลสำหรับกระบวนการรู้จำเพื่อให้สามารถบอกได้ว่าคำใดเป็นใน 100 คำ คำใดไม่ใช่

การแปลผลจะยึดเอาโหนดที่มีค่ามากที่สุดเป็นคำตอบ เพื่อให้สามารถปฏิเสธคำที่ไม่ใช่คำใน 100 คำได้จึงมีค่าเทรสโฮลด์ เอาไว้สำหรับเปรียบเทียบกับโหนดที่มีค่ามาก

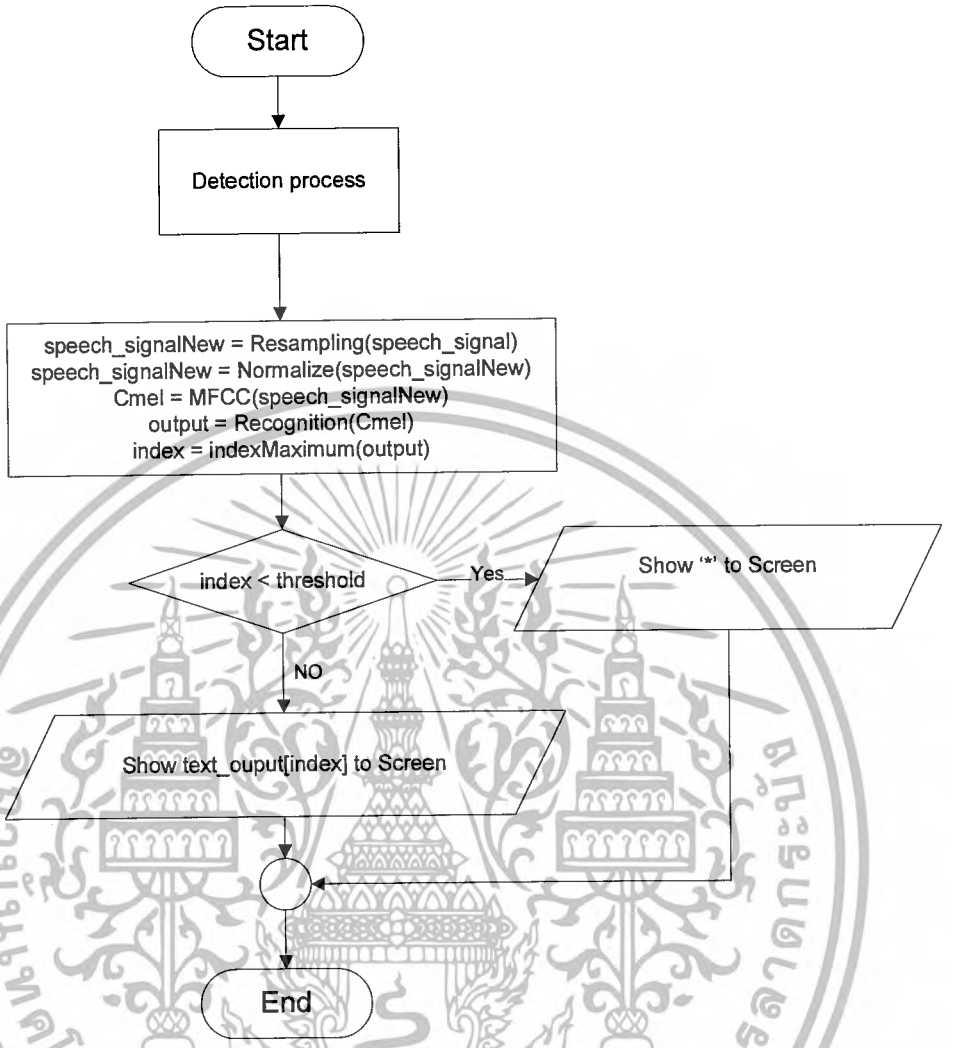
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ที่สุด หากโหนดที่มีค่ามากที่สุดมีค่าน้อยกว่าเทรสโฮลด์ จะถือว่าคำนั้นไม่ตรงกับคำใดเลยในร้อยคำ โดยกำหนดแทนคำๆนั้นด้วยสัญลักษณ์ '*' สำหรับอัลกอริทึมจากที่อธิบายมาข้างต้นเป็นดังรูป 3.5 หน้า 36



รูป 3.4 โฟลว์ชาร์ตอัลกอริทึมในการตัดสัญญาณเสียงที่เป็นคำพูด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.5 โฟล์วชาร์ตอัลกอริทึมการแปลค่าเอาท์พุท

สำหรับส่วนรีเวิร์ดพาธ โครงงานนี้ใช้โปรแกรม นิวโรโซลูชัน (Neuro solution) ช่วยเพื่อหาค่าเวจน์ที่ดีจึงๆ ไม่ขอกล่าวเรื่องการออกแบบ

3.3.2 การออกแบบส่วนการแปลงตัวอักษรไปเป็นเสียง

สำหรับโครงงานนี้เลือกทำเฉพาะสองส่วนก็คือ ส่วนการวิเคราะห์ข้อความและส่วนการสังเคราะห์เสียง

3.3.2.1 ธรรมชาติของภาษาไทย

พยางค์ภาษาไทยประกอบด้วยพยัญชนะต้น พยัญชนะควบ สระ ตัวสะกด วรรณยุกต์ โดยความยาวอยู่ระหว่าง 2-8 ตัว ประกอบด้วยพยัญชนะต้น 1ตัว พยัญชนะควบ 0-1 ตัว สระ 0-3 ตัว ตัวสะกด 0-2 ตัว วรรณยุกต์ 0-1 ตัว นอกจากพยางค์ที่เป็นภาษาไทยแล้วยังมีตัวเลข อักษร ‘ ’ ด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2.2 การออกแบบส่วนการวิเคราะห์ข้อความ

สำหรับส่วนนี้ใช้วิธีการตรวจสอบพยางค์จากที่มีอยู่ในไฟล์เอ็กซ์เอ็มแอล ซึ่งใช้พอยเตอร์(pointer) 2 ตัว ตัวที่ 1 บอกจุดเริ่มของพยางค์ ตัวที่ 2 บอกจุดสิ้นสุดของพยางค์ โดยมีขั้นตอนดังนี้

1) ตอนแรกพอยเตอร์ตัวที่ 1 และตัวที่ 2 อยู่ท้ายประโยค จากนั้นเลื่อนพอยเตอร์ตัวที่ 1 จากท้ายของประโยค นำตัวอักษรตั้งแต่พอยเตอร์ตัวที่ 1 จนถึงพอยเตอร์ตัวที่ 2 หากพยางค์ๆ นั้นไม่ตรงกับพยางค์ในไฟล์ให้ทำการเลื่อนพอยเตอร์ตัวที่ 1 ไปข้างหน้าจนกว่าจะตรงกับพยางค์ในไฟล์เอ็กซ์เอ็มแอล จากนั้น แทรกสัญลักษณ์ '%' ที่พอยเตอร์ตัวที่ 2 และเลื่อนพอยเตอร์ตัวที่ 2 มาที่พอยเตอร์ตัวที่ 1 เลื่อพอยเตอร์ตัวที่หนึ่งมาข้างหน้า 2 ช่อง

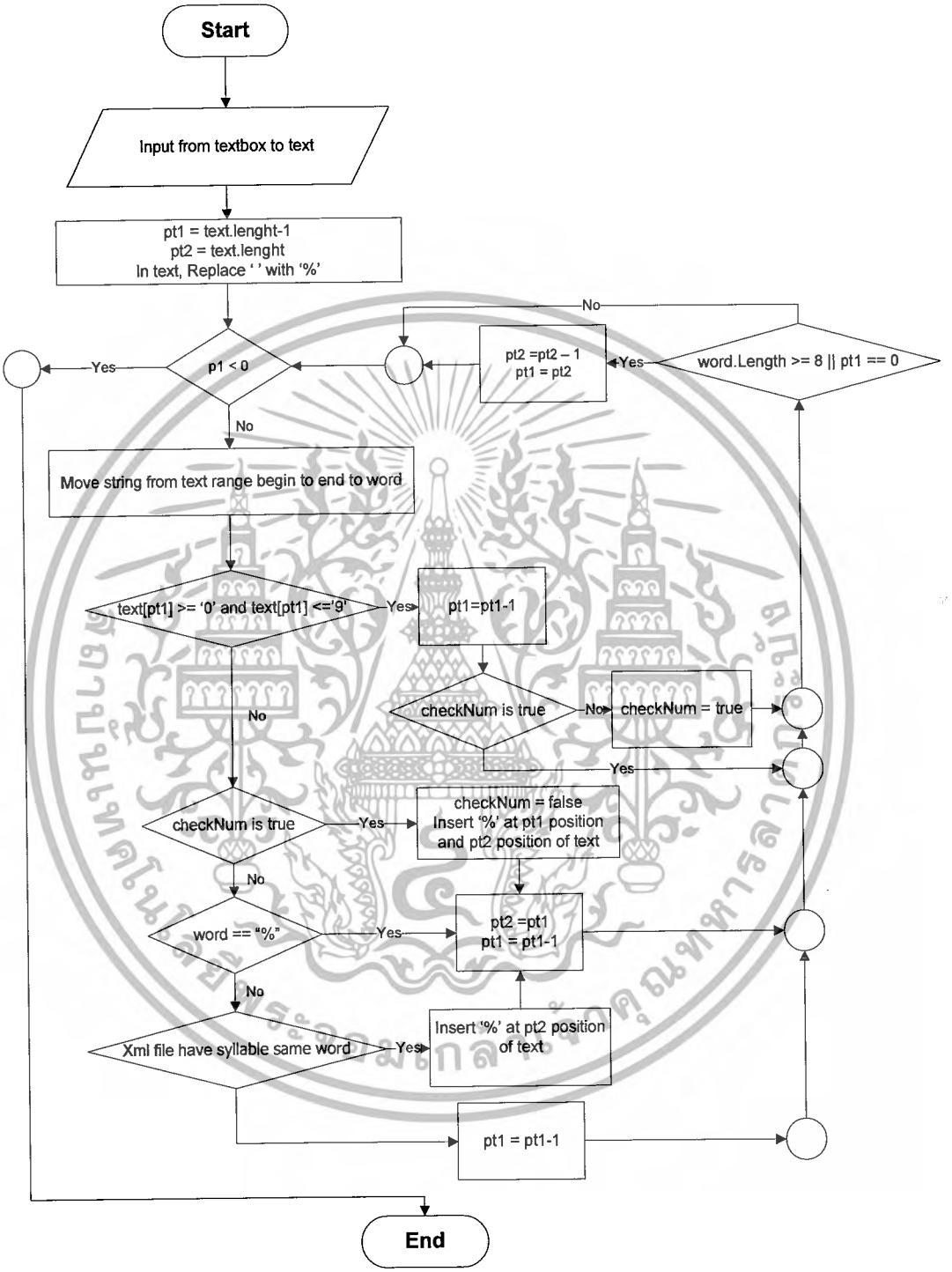
2) ถ้ายังไม่ตรงกับพยางค์ในไฟล์เอ็กซ์เอ็มแอล จนกระทั่งความยาวของอักษรระหว่างพอยเตอร์ตัวที่ 1 กับพอยเตอร์ตัวที่ 2 มีจำนวนเกิน 8 ตัวให้ทำการเลื่อนพอยเตอร์ตัวที่ 2 ไปข้างหน้า และเลื่อนพอยเตอร์ตัวที่ 1 กลับมาอยู่ข้างหน้าพอยเตอร์ตัวที่สองหนึ่งช่อง

3) หาก ณ ตำแหน่งที่พอยเตอร์ตัวที่ 1 ชี้เป็นอักขระ ' ' ให้ทำการแทนด้วย '%' และเลื่อน พอยเตอร์ตัวที่ 2 มาที่พอยเตอร์ตัวที่หนึ่ง เลื่อนพอยเตอร์ตัวที่ 1 ไปข้างหน้า 1 ช่อง

4) หากพอยเตอร์ตัวที่ 1 ชี้ไปที่ตัวเลข ให้ทำการแทรก % ที่ตำแหน่งก่อนหน้าพอยเตอร์ตัวที่ 1 และ เลื่อนพอยเตอร์ตัวที่ 2 มายังตำแหน่งก่อนหน้าพอยเตอร์ตัวที่ 1 เลื่อนพอยเตอร์ตัวที่หนึ่งไปข้างหน้าจนกระทั่งไม่เจอตัวเลข ทำการแทรก '%' ไปที่ตำแหน่งพอยเตอร์ตัวที่ 1 ชี้ ทำตามสี่ขั้นตอนแรกจนครบทั้งข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

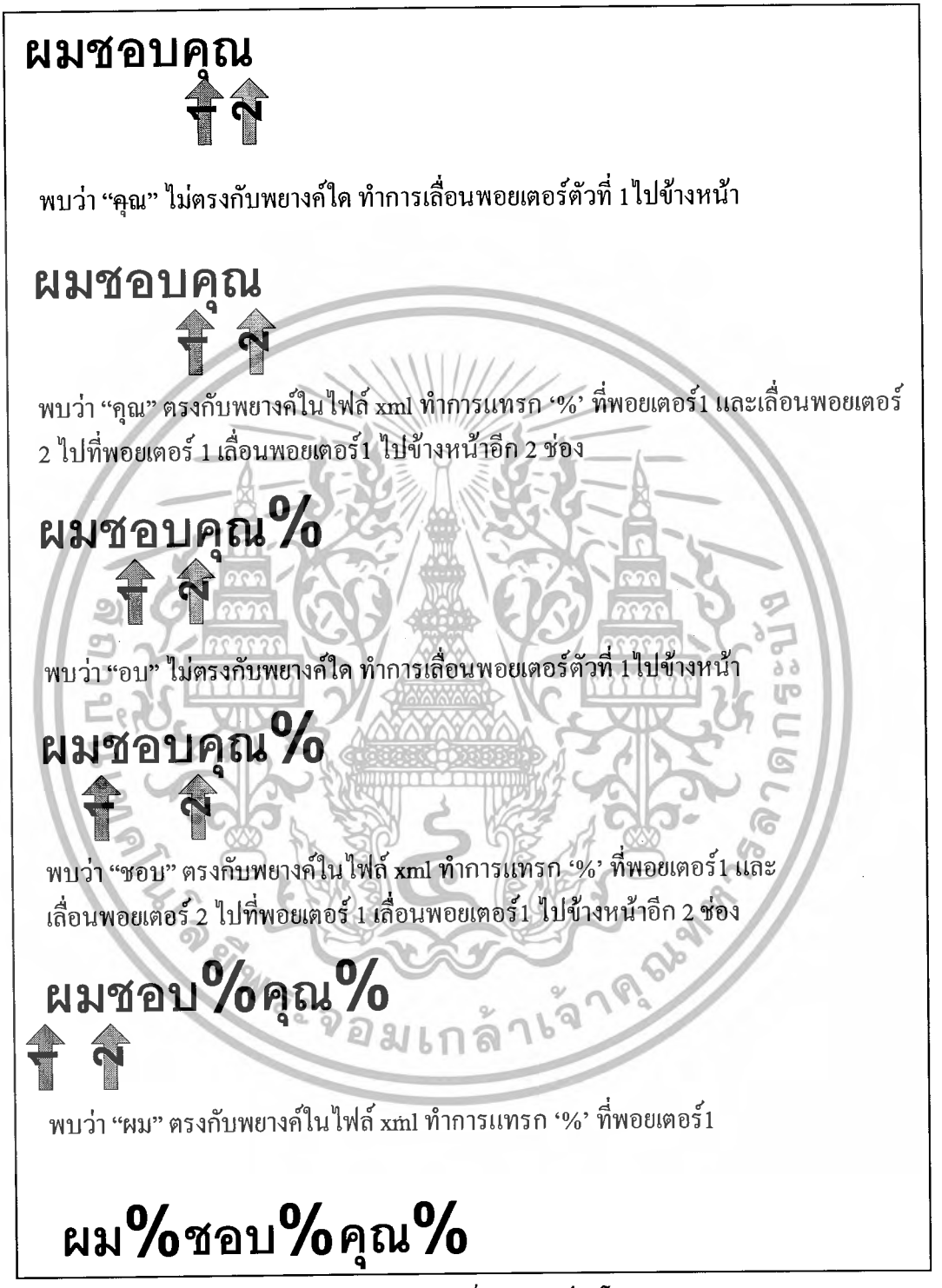
สามารถเขียนโฟลว์ชาร์ตแสดงอัลกอริทึมการวิเคราะห์ข้อความได้ดังนี้



รูป 3.6 โฟลว์ชาร์ตอัลกอริทึมการวิเคราะห์ข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

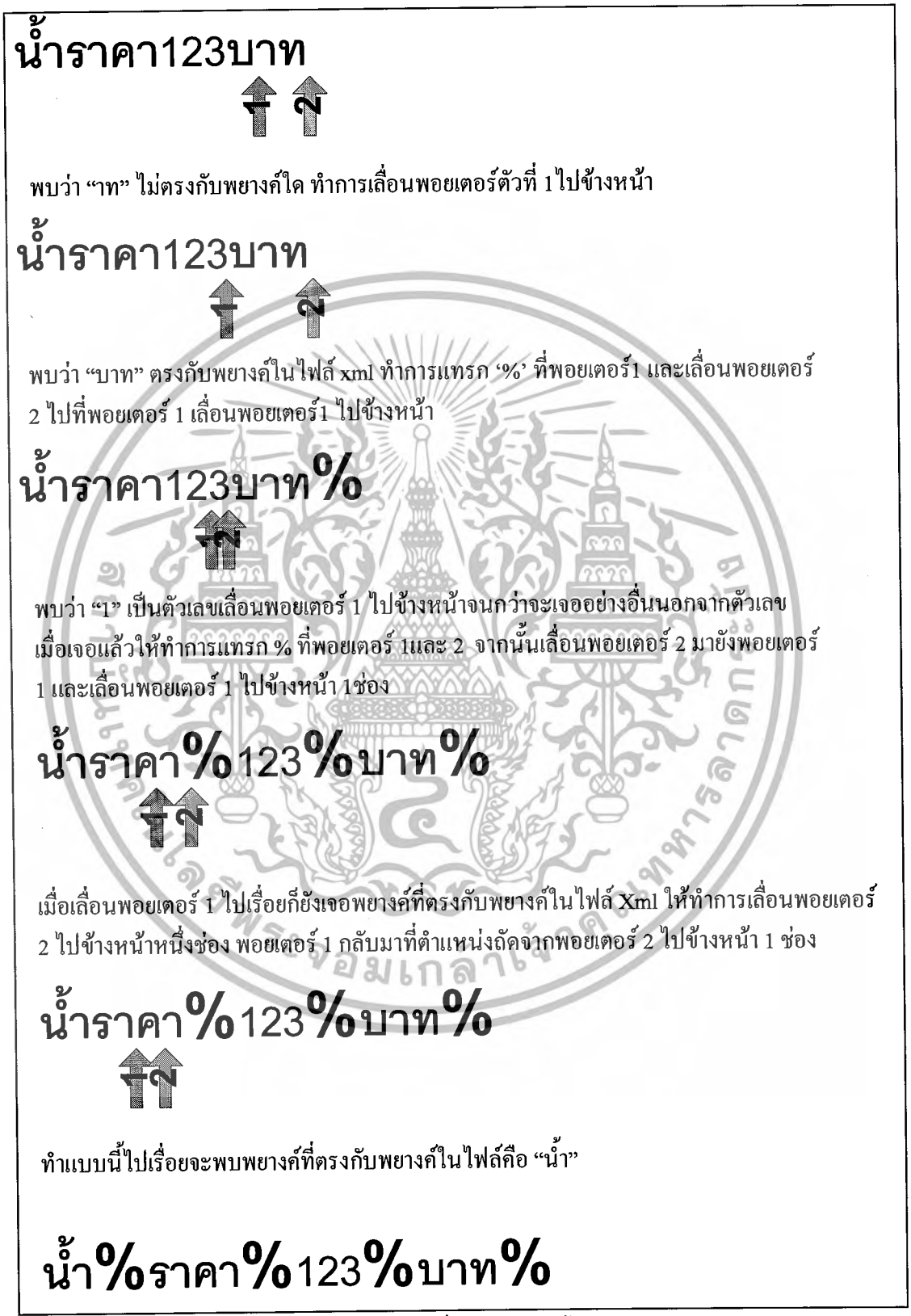
ตัวอย่างกระบวนการวิเคราะห์ข้อความ “ผมชอบคุณ”



รูป 3.7 กระบวนการแยกพยางค์ออกจากประโยค “ผมชอบคุณ”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างการวิเคราะห์ข้อความ “ราคา1234บาท”



รูป 3.8 กระบวนการแยกพยางค์ออกจากประโยค “ราคา123บาท”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2.3 การออกแบบส่วนการสังเคราะห์คำพูด

สำหรับโครงงานนี้เลือกใช้วิธีคอนแคททิเนชัน ซินเทอะซิส ซึ่งเป็นวิธีที่ได้รับความนิยมมากที่สุดในปัจจุบัน โดยเสียงที่ทำการสังเคราะห์ขึ้นนี้เกิดจากการนำหน่วยเสียงย่อย ที่ทำการเก็บไว้ก่อนมาต่อกันเป็นเสียงพูดที่ต้องการ สำหรับการจับคู่ไฟล์เสียงกับพยางค์แต่ละตัวใช้เอ็กซ์เอ็มแอล ไฟล์เป็นตัวเก็บข้อมูลดังกล่าวเพื่อให้ง่ายต่อการเพิ่มเสียงพยางค์ โดยมีการกำหนดแท็ค (Tag) ต่างๆไว้ดังนี้

<Root>
<ชื่อพยางค์>
<male> พาธ (path) ไฟล์เสียงผู้ชายของพยางค์นั้น </male>
<female> พาธไฟล์เสียงผู้หญิงของพยางค์นั้น </female>
</ชื่อพยางค์>

<ชื่อพยางค์>
<male> พาธไฟล์เสียงผู้ชายของพยางค์นั้น </male>
<female> พาธไฟล์เสียงผู้หญิงของพยางค์นั้น </female>
</ชื่อพยางค์>

<ชื่อพยางค์>
<male> พาธไฟล์เสียงผู้ชายของพยางค์นั้น </male>
<female> พาธไฟล์เสียงผู้หญิงของพยางค์นั้น </female>
</ชื่อพยางค์>
.
.
.
</Root>

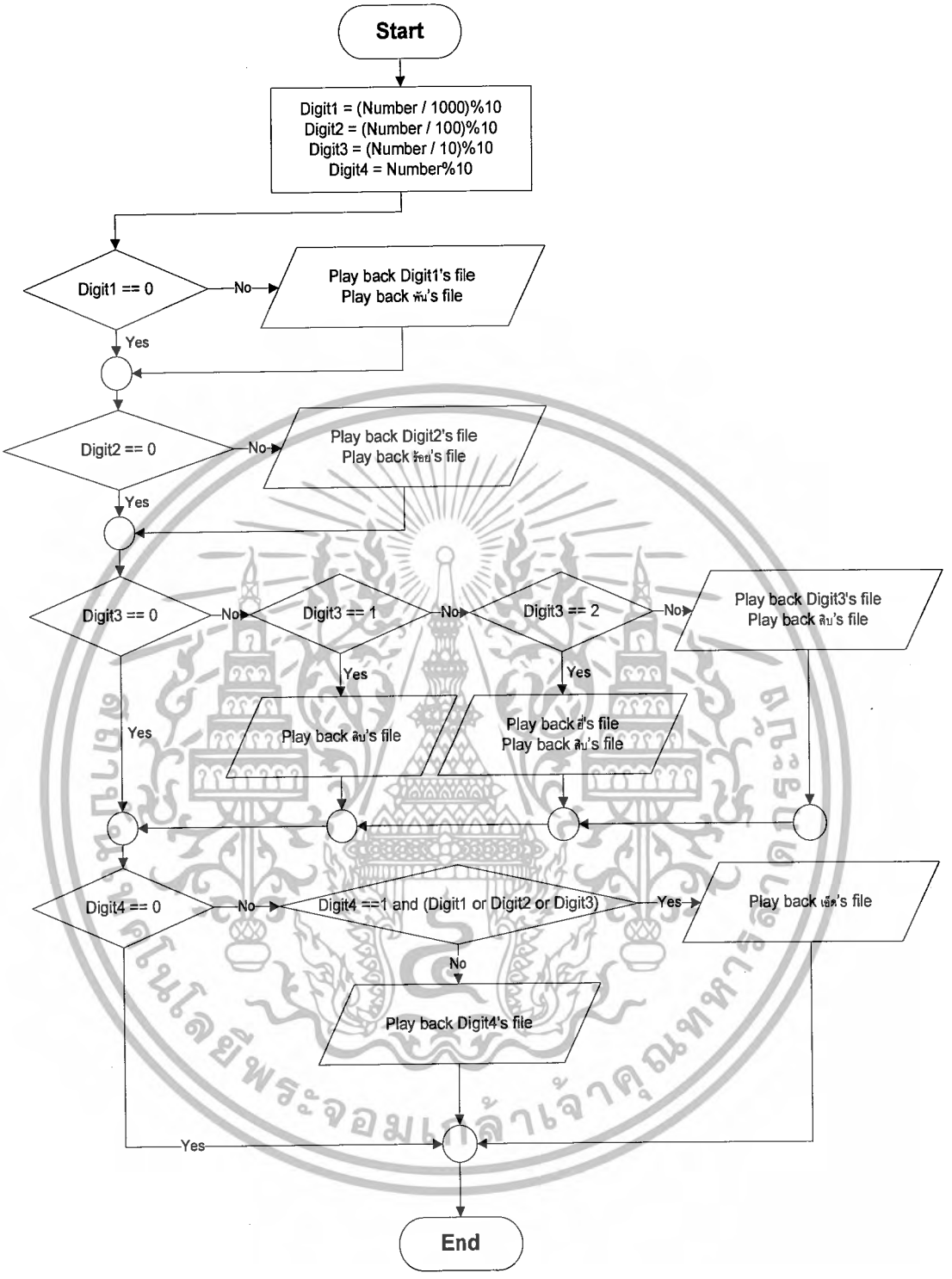
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติว่าหากมีการเก็บพยางค์ “กิน” “กลับ” โดยคำว่า “กิน” มีไฟล์เสียงผู้ชายอยู่ที่พาธ /sound/male/kin.wav ไฟล์เสียงผู้หญิงอยู่ที่พาธ /sound/female/kin.wav คำว่า “กลับ” มีไฟล์เสียงผู้ชายอยู่ที่พาธ /sound/glub.wav ไฟล์เสียงผู้หญิงอยู่ที่พาธ /wav/glub.wav สามารถเขียนเอ็กซ์เอ็มแอล ไฟล์ได้ดังนี้

```
<Root>
<กิน>
<male>/sound/male/kin.wav</male>
<female>/sound/female/kin.wav</female>
</กิน>
<กลับ>
<male>/sound/glub.wav</male>
<female>/wav/glub.wav</female>
</กลับ>
</Root>
```

นอกจากพยางค์ที่ได้จากส่วนการวิเคราะห์ข้อความที่สามารถนำมาเล่นไฟล์เสียงได้เลย ยังมีตัวเลขที่ไม่สามารถนำมาจับคู่กับไฟล์เอ็กซ์เอ็มแอล ได้ ทั้งนี้จึงจำเป็นที่จะต้องมีการวิเคราะห์ว่าควรอ่านตัวเลขอย่างไร ซึ่งในโครงงานนี้ใช้อัลกอริทึมในการอ่านตัวเลขโดยอ่านได้ตั้งแต่ 0 – 9999 ดังนี้

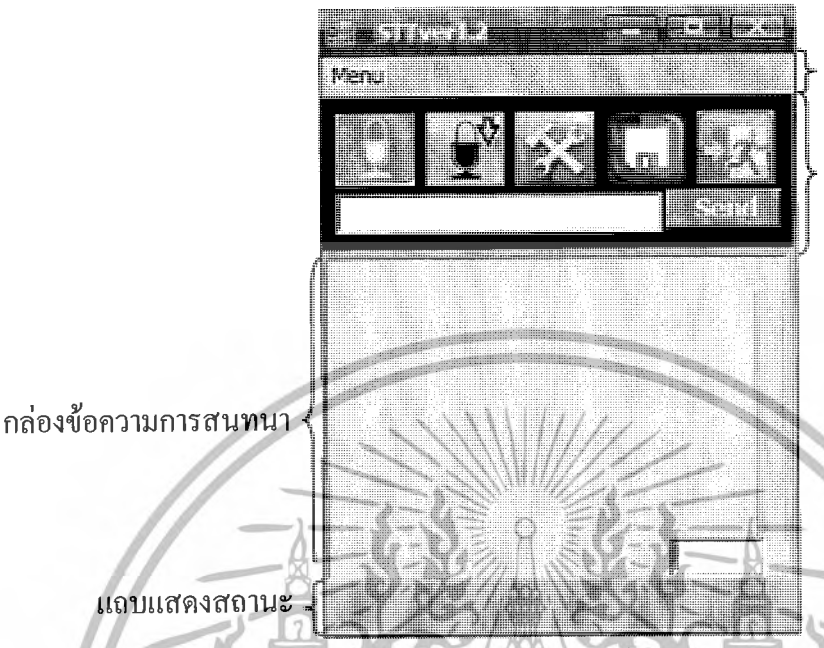
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.9 โฟลว์ชาร์ตอัลกอริทึมการอ่านตัวเลข

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.3 การออกแบบส่วนติดต่อผู้ใช้งาน



รูป 3.10 หน้าหลักของโปรแกรม

ส่วนประกอบหลัก ได้แก่

1) แถบเครื่องมือ เป็นส่วนที่ใช้เก็บเครื่องมือหรือคำสั่งหลักๆในโปรแกรม

ปุ่มบันทึกเสียง ใช้สำหรับเริ่มและหยุดการรับเสียง

ปุ่มบันทึกเสียงพื้นหลัง ใช้สำหรับเก็บเสียงพื้นหลัง

ปุ่มตั้งค่า ใช้สำหรับตั้งค่าต่างๆในโปรแกรม เช่น ชื่อ เพศของผู้ใช้งาน

ปุ่มบันทึก ใช้สำหรับบันทึกบทสนทนา

ปุ่มออก ใช้เมื่อต้องการจบการทำงานของโปรแกรม

แถบการอ่าน เมื่อพิมพ์ข้อความแล้วกดปุ่ม Send โปรแกรมจะทำการอ่านข้อความดังกล่าว

2) กล่องข้อความการสนทนา เป็นกล่องข้อความแสดงข้อมูลการสนทนา

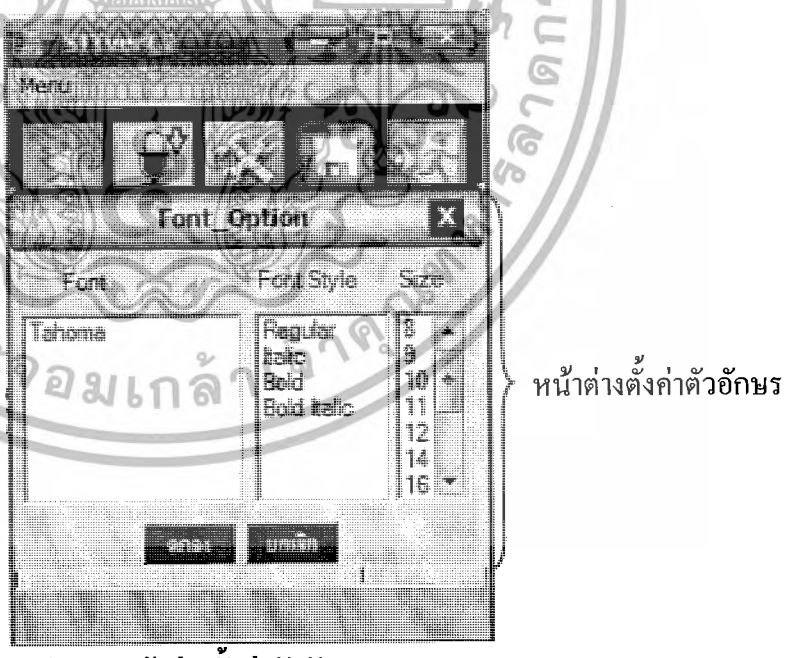
3) แถบแสดงสถานะ เป็นแถบแสดงสถานะของโปรแกรม

4) เมนู เป็นส่วนที่เก็บคำสั่งต่างๆไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูป 3.11 หน้าต่างการตั้งค่า

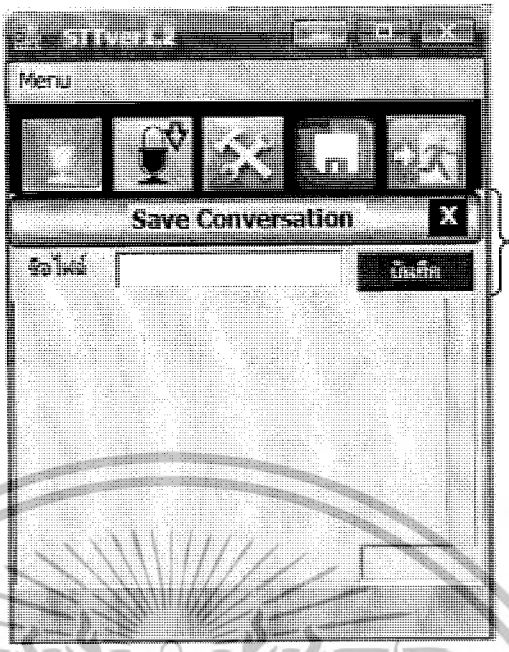
6) หน้าต่างการตั้งค่า จะปรากฏเมื่อมีการคลิกที่ปุ่ม Option หรือเลือกใช้เมนู Option ผู้ใช้งานสามารถกำหนดค่าต่างๆของโปรแกรมได้จากส่วนนี้



รูป 3.12 หน้าต่างตั้งค่าตัวอักษร

7) หน้าต่างตั้งค่าตัวอักษร ปรากฏขึ้นเมื่อมีการคลิกที่ปุ่ม set ของ Font ใน หน้าต่างการตั้งค่า โดยในส่วนนี้ใช้สำหรับเลือกรูปแบบตัวอักษรที่จะแสดงใน กล่องข้อความการสนทนา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



หน้าต่างบันทึกการสนทนา

รูป 3.13 หน้าต่างบันทึกการสนทนา

8) หน้าต่างบันทึกการสนทนา ปรากฏขึ้นเมื่อมีการคลิกที่ปุ่ม Save หรือ เลือกเมนู Save ใช้สำหรับบันทึกการสนทนา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

การทดลอง

การทดลองแบ่งเป็น 2 ส่วน คือ ส่วนการแปลงเสียงเป็นตัวอักษร และส่วนการแปลงตัวอักษรเป็นเสียง

4.1 การทดลองส่วนการแปลงเสียงเป็นตัวอักษร ประกอบด้วย 3 การทดลองดังนี้

4.1.1 การทดลองความถูกต้องของการตัดพยางค์

1) วิธีการทดลอง : อัดเสียงลงพีดีเอโดยใช้ประโยคทดลอง 5 ประโยค แต่ละประโยคมี 10 พยางค์ อัดประโยคละ 5 ครั้ง

ประโยคที่ 1 : ผม ชื่อ เล็ก เป็น คน กรุง เทพ โดย กำ เนิด

ประโยคที่ 2 : ฉัน ชอบ วิ ชา โปร แกรม มิ่ง มาก เลย ค่ะ

ประโยคที่ 3 : เครื่อง นี้ ใช้ สำ หรับ ผู้ พิ การ ทาง หู

ประโยคที่ 4 : นี่ คือ การ ทด ลอง โปร เจ็ก ของ พวก เรา

ประโยคที่ 5 : ผู้ คุม ลอง ให้ คะ แนน เต็ม ด้วย นะ ครับ

2) จำนวนผู้ทดลอง : 3 คน

3) สถานที่ : ห้องฮาร์ดแวร์

4) สภาพแวดล้อม : เงียบ (ไม่มีคน)

5) ขั้นตอนการทดลอง :

5.1) อัดเสียงผ่านพีดีเอโดยใช้ประโยคที่กำหนด แล้วบันทึกลงในพีดีเอ

5.2) ทำการทดลอง 5 ประโยค ประโยคละ 5 ครั้ง ผู้ทำการทดลอง 3 คน จึงได้ ประโยค ทั้งหมด 75 ประโยค เท่ากับ 750 พยางค์

5.3) นำไฟล์เสียงแต่ละประโยคมาเข้ากระบวนการตัดเสียงด้วยโปรแกรม

5.4) ฟังเสียงแต่ละไฟล์ว่าเป็นเสียงที่ถูกต้องหรือไม่ แล้วบันทึกผล โดยจำแนกชนิดของเสียงที่ตัดได้ออกเป็น 3 ประเภทคือ

1) เสียงที่ตัดได้ถูกต้อง คือ เสียงที่ตัดได้ตามที่คาดหวังไว้

2) เสียงที่ตัดผิด แบ่งได้ 3 ประเภท คือ

2.1) เสียงพื้นหลัง

2.2) เสียงไม่จบพยางค์

2.3) เสียงมากกว่า 1 พยางค์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) เสียงที่ตัดพลาด คือ เสียงที่หายไปเมื่อผ่านกระบวนการตัดเสียง

พบว่าอัลกอริทึมที่ใช้ในโครงงานนี้ ได้ผลการทดลองดังตาราง 4.1, 4.2, 4.3, 4.4, 4.5

ตาราง 4.1 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ

ประโยคที่	ครั้งที่	จำนวนพยางค์ทั้งหมดที่ได้	พยางค์ที่ตัดได้ถูกต้อง	พยางค์ที่ตัดผิด			พยางค์ที่ตัดพลาด
				เสียงพื้นหลัง	พยางค์ที่พลาด	ตัดได้มากกว่า 1 พยางค์	
1	1	10	10				
	2	10	9		1		
	3	10	9		1		
	4	10	9		1		
	5	10	10				
2	1	10	10				
	2	10	10				
	3	10	10				
	4	10	10				
	5	12	10	2			
3	1	9	9				1
	2	10	9		1		
	3	9	9				1
	4	10	10				
	5	9	9				1
4	1	10	10				
	2	10	10				
	3	10	10				
	4	10	10				
	5	10	10				
5	1	10	10				
	2	10	10				
	3	10	10				
	4	10	10				
	5	10	10				
รวม		249	243	2	4	0	3

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.2 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น

ประโยคที่	ครั้งที่	จำนวนพยางค์ทั้งหมดที่ได้	พยางค์ที่ตัดได้ถูกต้อง	พยางค์ที่ตัดผิด			พยางค์ที่ตัดพลาด
				เสียงพื้นหลัง	ตัดได้ไม่จบพยางค์	ตัดได้มากกว่า 1 พยางค์	
1	1	9	8		1		1
	2	12	9	2	1		
	3	11	10	1			
	4	10	10				
	5	10	9		1		
2	1	11	9	1	1		
	2	9	8		1		1
	3	10	9		1		
	4	9	8		1		1
	5	9	8		1		1
3	1	8	7			1	1
	2	10	8	1		1	
	3	7	5			2	1
	4	8	7			1	1
	5	9	7	1		1	1
4	1	10	10				
	2	7	3	1		3	
	3	7	4	1		2	
	4	6	1	2		3	
	5	6	2	1		3	
5	1	10	9		1		
	2	10	8		2		
	3	10	9		1		
	4	10	9		1		
	5	9	8		1		1
รวม		227	185	11	14	17	9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.3 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล

ประโยคที่	ครั้งที่	จำนวนพยางค์ทั้งหมดที่ได้	พยางค์ที่ตัดได้ถูกต้อง	พยางค์ที่ตัดผิด			พยางค์ที่ตัดพลาด
				เสียงพื้นหลัง	ตัดได้ไม่จบพยางค์	ตัดได้มากกว่า 1 พยางค์	
1	1	10	10				
	2	9	8			1	
	3	10	10				
	4	11	10	1			
	5	10	10				
2	1	8	6			2	
	2	8	6			2	
	3	10	8	1		1	
	4	9	8			1	
	5	8	6			2	
3	1	10	10				
	2	10	10				
	3	10	10				
	4	10	10				
	5	10	10				
4	1	9	8			1	
	2	9	8			1	
	3	10	10				
	4	11	10	1			
	5	10	10				
5	1	10	10				
	2	10	10				
	3	9	8			1	
	4	10	10				
	5	10	10				
รวม		241	226	3	0	12	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.4 สรุปผลการทดลอง

ชื่อ	จำนวนพยางค์ทั้งหมดที่ได้	พยางค์ที่ตัดได้ถูกต้อง	พยางค์ที่ตัดผิด			พยางค์ที่ตัดพลาด
			เสียงพื้นหลัง	ตัดได้ไม่จบพยางค์	ตัดได้มากกว่า 1 พยางค์	
ปัทมา	249	243	2	4	0	3
รุ่งโรจน์	227	185	11	14	17	9
ชาญวิศ	241	226	3	0	12	0
รวม	717	654	16	18	29	12

ตาราง 4.5 สรุปผลการทดลองเป็นเปอร์เซ็นต์ (percent)

พยางค์ที่ตัดได้ถูกต้อง	พยางค์ที่ตัดผิด			พยางค์ที่ตัดพลาด
	ตัดได้เสียงพื้นหลัง	ตัดได้ไม่จบพยางค์	ตัดได้มากกว่า 1 พยางค์	
91.21 %	2.23 %	2.51 %	4.05 %	1.60 %

4.1.2 การทดลองความถูกต้องของการรู้จำเสียง

1) อุปกรณ์ที่ใช้ทดลอง คือ พีดีเอ เอชพี ไอแพ็ค มาเวล อาร์ พีเอ็กซ์เอ310 (PDA HP iPAQ Marvell (R) PXA310) 57.86 เมกกะ ไบท์ (MB) วินโดว์ โมบาย 6 (Window Mobile 6), โปรแกรมรู้จำเสียงบนพีดีเอซึ่งผ่านการฝึกสอนด้วยโครงข่ายประสาทเทียม มีโครงสร้างแบบมัลติเลเยอร์ เพอร์เซปตรอน โดยใช้พยางค์ในภาษาไทยในการทดสอบจำนวน 100 พยางค์ จนกระทั่งได้ค่ามีนสแควร์ เออเรอร์ (Mean Square Error) = 0.00055

2) วิธีการทดลอง: ทดลองคนละ 100 พยางค์ พยางค์ละ 10 ครั้ง

3) จำนวนผู้ทดลอง 3 คน (รวม 30ตัวอย่าง/1พยางค์)

4) สถานที่: ห้องฮาร์ดแวร์

5) สภาพแวดล้อม: เงียบ (ไม่มีคน)

6) ขั้นตอนการทดลอง:

6.1) อัดเสียงใส่พีดีเอ 100 พยางค์ โดยแต่ละพยางค์พูด 15 ครั้ง

6.2) นำไฟล์เสียงที่ได้ไปเข้าโปรแกรมตัดเสียง แล้วตรวจลองไฟล์ที่ตัดได้ จากนั้นคัดเลือกเฉพาะไฟล์เสียงที่สมบูรณ์มา 10 ไฟล์/พยางค์

6.3) นำไฟล์ที่ได้มาเข้าโปรแกรมรู้จำเสียง

พบว่าได้ผลการทดลองดังตาราง 4.9, 4.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	10
การ	การ	การ	การ	การ	การ	การ	การ	การ	การ	การ	10
กิน	กิน	กิน	กิน	กิน	*	กิน	กิน	กิน	กิน	กิน	9
เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	10
เก้า	เก้า	เก้า	เก้า	*	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	9
โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	10
ขอ	ขอบ	สอง	ขอบ	ท้อง	ขอ	ท้อง	ท้อง	ท้อง	ท้อง	ขอ	2
ขอบ	พ่อ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	9
ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	10
ข้าว	ข้าว	ข้าว	ข้าว	บาท	ข้าว	พ่อ	ข้าว	ข้าว	ข้าว	พ่อ	7
ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	เป็น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	9
เขา	เท่า	เขา	เขา	เขา	เขา	เขา	เขา	เขา	เขา	เขา	9
แขน	แขน	แขน	แขน	แขน	แขน	แขน	แขน	แขน	แขน	แขน	10
คน	รถ	คน	คน	โกรธ	คน	คน	โกรธ	คน	คน	โกรธ	6
ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	10
ค่า	*	ค่า	*	*	*	*	ขา	*	ขา	*	1
คิด	คิด	*	คิด	คิด	คิด	คิด	คิด	กิน	คิด	คิด	8
คุณ	คุณ	คุณ	คุณ	คุณ	*	ผม	โกรธ	คุณ	*	คุณ	6

ตาราง 4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
ใคร	ชาย	ไฟ	ไฟ	ใจ	*	*	ไฟ	ไฟ	ใจ	*	0
เจ็ด	เจ็ด	เจ็ด	เจ็บ	*	เจ็ด	*	เมล์	*	เจ็ด	*	4
เจ็บ	*	เกลียด	เจ็บ	เจ็บ	เจ็บ	เจ็บ	*	เจ็ด	หญิง	เจ็บ	5
ใจ	ไฟ	*	ใจ	ใจ	ใจ	ใจ	ใจ	ใจ	ใจ	*	7
ฉัน	สาม	ฉัน	เรียน	ฉัน	ฉัน	กลับ	ฉัน	*	*	ทำ	4
ชอบ	ชอบ	ชอบ	ชอบ	*	*	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	8
ชาย	การ	ชาย	ชาย	*	*	*	ชาย	*	*	*	3
ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	10
เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	10
ดี	ดี	ดี	ดี	ดี	ดี	ดี	*	ดี	ดี	ดี	9
เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	10
ตา	ปาก	ปาก	ตา	ปาก	ตา	ตา	*	ปาก	ปาก	ตา	4
ถึง	หนึ่ง	ถึง	หนึ่ง	*	หญิง	ถึง	หนึ่ง	ถึง	ถึง	*	4
ท้อง	พบ	ท้อง	ท้อง	สอง	ท้อง	ท้อง	ท้อง	*	ท้อง	สอง	6
ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	สาม	ทำ	ทำ	ทำ	ทำ	9
ที่	กิน	ที่	ที่	ที่	ที่	ที่	ที่	ที่	ที่	ที่	9
เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	10
เที่ยง	เที่ยง	เที่ยง	เที่ยง	*	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เรียน	*	7

ตาราง 4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	*	โทษ	โทษ	โทษ	9
เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	10
น้ำ	ยา	น้ำ	น้ำ	น้ำ	น้ำ	น้ำ	น้ำ	น้ำ	น้ำ	*	8
บาท	บ้าน	บาท	บาท	บาท	บาท	บาท	บาท	บาท	บาท	บาท	9
บ้าน	ทำ	บ้าน	*	อยาก	บาท	บ้าน	บ้าน	บ้าน	บ้าน	บ้าน	6
บ่าย	บ่าย	บ่าย	คน	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	9
เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	10
ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	10
ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	10
เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	*	เป็น	เป็น	เป็น	เป็น	9
แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	10
ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	10
ผม	หก	ผม	ผม	ผม	ผม	ผม	ผม	ผม	ผม	ผม	9
พบ	หก	พบ	ผม	พบ	พบ	ผม	ผม	หก	ผม	ผม	3
พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	10
พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	10
พัน	พัก	ข้าว	*	พัน	พัก	พัน	พัน	พัน	พัน	*	5
เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	10
ไฟ	ไฟ	ไฟ	ไฟ	ไฟ	ไฟ	ไฟ	ไป	ไฟ	ไฟ	ไฟ	9

ตาราง 4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
มา	บ้าน	บ้าน	บาท	บาท	มา	บาท	มา	บ้าน	มา	บาท	3
มาก	ตา	มา	บ้าน	บ้าน	มา	มา	มาก	บ้าน	มา	มา	1
มือ	มือ	มือ	มือ	มือ	หนึ่ง	มือ	มือ	มือ	มือ	มือ	9
เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	10
เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	10
แม่	*	แม่	*	มา	เมล์	เมล์	เมล์	แม่	*	*	2
ไม	ไป	ไม	ไม	ไป	มา	ไม	มา	ไม	ไม	*	5
ไม่	ไม่	ไม	มา	มา	มา	ไม่	มา	ไม่	มา	*	3
ยา	ยา	ยา	ยา	ยา	ยา	ยา	ยา	ยา	ยา	ยา	10
ยี่	ดี	ดี	ยี่	ยี่	ยี่	ดี	กิน	ยี่	*	ยี่	5
เย็น	หญิง	เย็น	เย็น	เด็ก	เย็น	*	หญิง	เมล์	เมล์	เย็น	4
รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	10
ร้อน	ร้อน	ร้อน	ร้อน	ร้อน	ร้อน	บาท	ร้อน	ร้อน	ร้อน	ร้อน	9
ร้อย	ร้อน	ร้อย	ร้อย	ร้อย	ร้อย	ร้อย	ร้อย	ร้อย	ร้อย	ร้อย	9
เรียน	เบื่อ	เรียน	เรียน	เรียน	*	เรียน	เรียน	เบื่อ	เรียน	เรียน	7
โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	10
ไร	ไร	ไร	ไร	ไร	*	ไร	ไร	บ้าน	ไร	ไร	8
ลง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	0
ลืม	ลืม	ลืม	ชื่อ	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	9

ตาราง 4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
เลือด	เลือด	เด็ก	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	9
วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	10
วัน	วัด	วัด	วัด	วัด	วัน	วัน	วัน	วัน	วัน	วัด	5
วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	10
สอง	สอง	สอง	สอง	สอง	สอง	สุข	สอง	สอง	ขอบ	สอง	8
สาม	สาม	สาม	สาม	สาม	สาม	*	*	สาม	สาม	สาม	8
สิบ	สิบ	สิบ	สิบ	สิบ	สิบ	สิบ	คิด	สิบ	สิบ	สิบ	9
สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	10
สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	10
เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	10
หก	หก	หก	ผม	หก	หก	หก	ผม	ผม	หก	หก	7
หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	ดี	หญิง	หญิง	หญิง	9
หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	10
หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	10
หนึ่ง	ลืม	ลืม	ลืม	ลืม	มือ	หนึ่ง	เรียน	หนึ่ง	ลืม	ลืม	2
หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	10
หลัง	วัด	บาท	*	บาท	อยาก	บาท	อยาก	*	หลัง	หลัง	2
หัว	หัว	หัว	หัว	หัว	หัว	หัว	หัว	หัว	หัว	หัว	10

ตาราง 4.6 ผลการทดลองของ นายรุ่งโรจน์ ศิริปิ่น (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
ห้า	หน้า	ห้า	ห้า	ห้า	*	ห้า	*	ห้า	ห้า	ห้า	7
หาร	ห้า	อา	หาร	หาร	หาร	*	หาร	อา	*	พัน	4
ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	10
อยาก	อยาก	ยา	อยาก	ยา	อย่าง	อยาก	อย่าง	อย่าง	ยา	อยาก	4
อย่าง	อยาก	อยาก	อยาก	อยาก	อยาก	อย่าง	อยาก	อยาก	อย่าง	อยาก	2
อะ	อะ	อะ	อะ	อะ	อะ	อะ	อะ	*	อะ	มาก	8
อา	อา	อา	อา	มา	อา	อา	*	อา	อา	อา	8
เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เย็น	เอ็ด	เอ็ด	เอ็ด	9

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ตาราง 4.7 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	10
การ	พัน	การ	การ	การ	การ	การ	การ	การ	การ	การ	9
กิน	ที่	กิน	กิน	สี่	กิน	กิน	กิน	กิน	กิน	กิน	8
เกลียด	เพื่อน	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	เกลียด	9
เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	10
โกรธ	โกรธ	โกรธ	ขอบ	โกรธ	โกรธ	โรง	โกรธ	*	โกรธ	*	6
ขอ	ขา	ขา	ขา	ขา	ขอ	ขอ	ขอ	ขอ	ขอ	ขา	5
ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	อะ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	9
ขา	ขา	*	ขา	ขา	ขา	ขา	*	*	ขา	อะ	6
ข้าว	เท่า	เท่า	ข้าว	ข้าว	เท่า	พ่อ	เท่า	เท่า	ข้าว	ข้าว	4
ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	10
เขา	เช้า	เขา	เช้า	เช้า	เขา	เช้า	เช้า	เช้า	เช้า	เขา	3
แขน	ขอบ	แขน	ขอบ	ขอบ	ท้อง	ขอบ	ขอบ	แขน	ขอบ	แขน	3
คน	คุณ	ตา	ผม	ทำ	*	ผม	*	ไป	คุณ	ผม	0
ครู	ครู	ครู	ไป	*	*	*	*	ครู	เธอ	ครู	4
ค่า	พัน	*	เท่า	พ่อ	ค่า	หน้า	*	ค่า	ค่า	ขา	3
คิด	คิด	คิด	สิบ	สิบ	คิด	คิด	สี่	คิด	สิบ	คิด	6
คุณ	ไป	คุณ	คุณ	คุณ	คุณ	คุณ	คุณ	ตา	คุณ	เธอ	7

ตาราง 4.7 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
ใคร	ไป	คน	ขอบ	*	*	ขอบ	ไป	*	คน	ขอบ	0
เจ็ด	เจ็ด	เจ็ด	เจ็ด	เจ็ด	เจ็ด	เจ็ด	เอ็ด	เป็น	เจ็ด	เอ็ด	7
เจ็บ	เจ็ด	เอ็ด	เอ็ด	*	เอ็ด	เอ็ด	*	*	เจ็ด	เอ็ด	0
ใจ	ไป	ใจ	ใจ	ใจ	ใจ	ใจ	ไฟ	ไป	ใจ	ใจ	7
ฉัน	ฉัน	ฉัน	*	ฉัน	ฉัน	ฉัน	ฉัน	ฉัน	กลับ	ฉัน	8
ชอบ	พ่อ	เท่า	ข้าว	กลับ	*	กลับ	*	ข้าว	กลับ	*	0
ชาย	เช้า	ชาย	ชาย	ชาย	ไฟ	ชาย	*	ชาย	ชาย	ชาย	7
ชื่อ	ชื่อ	ชื่อ	ชื่อ	ที่	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	9
เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	10
ดี	เที่ยง	ดี	ดี	ดี	ดี	ดี	ดี	ดี	ดี	ดี	9
เด็ก	*	เด็ก	เด็ก	ดี	เด็ก	เด็ก	เอ็ด	*	เด็ก	เอ็ด	5
ตา	*	การ	ตา	ตา	ไป	ไป	ไป	ปาก	การ	ตา	3
ถึง	*	ถึง	ถึง	ถึง	*	ถึง	ถึง	*	ถึง	ถึง	7
ท้อง	เช้า	ท้อง	เช้า	ทำ	ทำ	เช้า	เช้า	เท่า	เช้า	เช้า	1
ทำ	อย่าง	หาร	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	กลับ	7
ที่	*	ที่	สิบ	สิบ	สิบ	ที่	ยี่	ที่	ที่	ที่	5
เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	10
เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	10

ตาราง 4.7 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	10
เธอ	เธอ	เช้า	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	*	เธอ	8
น้ำ	ดี	ร้อย	มา	ดี	เช้า	น้ำ	เช้า	น้ำ	มา	น้ำ	3
บาท	บาท	ปาก	บ่าย	ปาก	ปาก	ไป	*	ปาก	ปาก	*	1
บ้าน	เท่า	บ้าน	บ้าน	บ้าน	บ้าน	บ้าน	บ้าน	บ้าน	บ้าน	บ้าน	9
บ่าย	บ่าย	บ่าย	บ่าย	*	*	*	บ่าย	บ่าย	หนาว	สอง	5
เบื่อ	มา	เบื่อ	เรียน	เพื่อน	เพื่อน	เสีย	เพื่อน	เพื่อน	เพื่อน	เพื่อน	1
ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	10
ปาก	ปาก	ปาก	อยาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ตา	8
เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	เป็น	10
แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	10
ไป	ไป	ไป	ไป	ไป	ไป	*	ไป	ไป	ไป	*	8
ผม	ผม	ผม	ผม	ผม	ตา	ผม	เมล์	หมอ	ผม	ไป	6
พบ	*	พบ	หก	พบ	พบ	พบ	พบ	พบ	หก	พบ	7
พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	พ่อ	10
พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	10
พัน	ทำ	พัน	พัน	พัน	พัก	พัก	พัน	พัน	*	พัน	6
เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เรียน	เพื่อน	9

ตาราง 4.7 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
ไฟ	ไฟ	เก้า	ไฟ	หาร	ไฟ	ไฟ	เค้า	หาร	หาร	ไฟ	5
มา	มา	มา	*	เบื่อ	*	เบื่อ	มา	มา	มา	มา	6
มาก	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	0
มือ	หนึ่ง	มือ	มือ	*	เมล์	*	มือ	หนึ่ง	มือ	มือ	5
เมล์	เมล์	เมล์	เมล์	เมล์	เมล์	*	เมล์	เมล์	เมล์	เมล์	9
เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เลือด	เมื่อ	เมื่อ	เลือด	8
แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	10
ไม	หน้า	ไม	*	ไป	ไป	ไม	ไม่	ไม	ไม่	ไม	4
ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	10
ยา	มา	*	*	*	*	*	ยา	*	ยา	*	2
ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	10
เย็น	เย็น	*	เย็น	เย็น	เย็น	แขน	เมล์	เย็น	เมล์	เย็น	6
รถ	*	รถ	*	สอง	*	สอง	*	รถ	สอง	*	2
ร้อน	ร้อย	*	ร้อน	บ้าน	*	*	*	ร้อน	ตา	ร้อน	3
ร้อย	หมอ	ร้อย	เมล์	ร้อย	*	ร้อย	ร้อย	ร้อย	ร้อย	ร้อย	7
เรียน	เรียน	เรียน	เรียน	เรียน	เรียน	เรียน	เรียน	เรียน	เรียน	เรียน	10
โรง	เธอ	ลง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	โรง	8
ไร	ไม่	ไร	ไร	ไร	ไป	ไร	ไร	ไร	ไร	วัน	7
ลง	เธอ	โรง	เธอ	เธอ	โรง	เธอ	เธอ	โรง	เธอ	เธอ	0

ตาราง 4.7 ผลการทดลองของ ปัทมา พรสุขศิริ (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	10
เลือด	เลือด	เลือด	ยี่	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	9
วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	10
วัน	วัน	วัน	วัน	วัน	วัน	วัน	วัน	วัน	วัน	วัน	10
วิ่ง	เรียน	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	วิ่ง	9
สอง	ทำ	การ	สอง	สอง	สอง	สอง	สอง	สอง	สอง	สอง	8
สาม	อะ	ค่า	อยาก	*	อยาก	ขา	สาม	อะ	กลับ	สาม	2
สิบ	สุข	สิบ	*	สิบ	ชื่อ	สิบ	สุข	สิบ	สิบ	สิบ	6
สี่	คิด	สี่	สี่	*	สี่	สี่	สี่	*	สี่	สี่	7
สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	สุข	10
เสีย	เสีย	เสีย	หัว	เสีย	หัว	เสีย	เสีย	เสีย	เสีย	เสีย	8
หก	หก	หก	หก	หก	หก	หก	หก	หก	หก	หก	10
หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	หญิง	10
หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	10
หนาว	เช้า	เช้า	น้ำ	หนาว	ผม	การ	หนาว	หนาว	การ	หนาว	4
หนึ่ง	ยี่	*	หนึ่ง	หนึ่ง	*	ดี	ดี	หนึ่ง	หนึ่ง	หนึ่ง	5
หมอ	*	หัว	หัว	หมอ	หัว	หมอ	หมอ	*	หมอ	หัว	4
หลัง	การ	*	*	*	*	เบื่อ	*	*	ไร	บ่าย	0

ตาราง 4.7 ผลการทดลองของ นางสาวปัทมา พรสุขศิริ (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
หัว	หัว	หัว	หัว	หัว	หัว	มา	หัว	หัว	หัว	หัว	9
ห้า	ห้า	ห้า	เท่า	ห้า	*	ห้า	ที่	ค่า	ห้า	ห้า	6
หาร	การ	การ	ห้า	ขา	หาร	การ	หาร	แปด	หาร	*	3
ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	10
อยาก	อย่าง	อย่าง	อย่าง	อย่าง	อย่าง	อย่าง	หัว	อย่าง	อย่าง	*	0
อย่าง	การ	อย่าง	หัว	*	*	*	การ	อย่าง	*	การ	2
อะ	กลับ	อะ	อะ	อะ	อะ	อะ	อะ	อะ	ตา	อะ	8
อา	การ	อา	การ	การ	การ	การ	การ	*	ตา	การ	1
เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	10

ตาราง 4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
กลับ	กลับ	อะ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	กลับ	9
การ	การ	การ	การ	การ	โกรธ	การ	การ	การ	การ	การ	9
กิน	มา	เที่ยง	*	หญิง	หญิง	หญิง	กิน	หญิง	กิน	กิน	3
เกลียด	เกลียด	เกลียด	*	เกลียด	เกลียด	ครู	ครู	เกลียด	ไป	เกลียด	6
เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	เก้า	10
โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โกรธ	โรง	โกรธ	โกรธ	9
ขอ	ขอ	ขอ	ขอ	ขอ	ขอ	ขอ	ขอ	*	ขอ	ขอ	9
ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	ขอบ	10
ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	ขา	10
ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	ข้าว	10
ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	ขึ้น	10
เขา	เขา	เขา	เขา	เขา	เขา	เขา	เขา	เขา	เขา	เขา	10
แขน	แขน	แขน	แขน	*	*	แขน	แขน	แขน	แขน	แขน	8
คน	คน	คน	คน	คน	คน	ผม	ผม	ผม	ผม	คน	6
ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	ครู	10
ค่า	ค่า	ค่า	ค่า	ค่า	ค่า	ค่า	ค่า	ค่า	ค่า	ห้า	9
คิด	คิด	คิด	คิด	คิด	ชื่อ	กิน	คิด	คิด	คิด	คิด	8
คุณ	มือ	วัน	หนึ่ง	คุณ	วัน	*	คุณ	วัน	*	คุณ	3
ใคร	*	ไป	ใคร	ใคร	ใคร	ใคร	ไป	ไป	*	ใคร	5

ตาราง 4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
เจ็ด	เจ็บ	ไป	โกรธ	เอ็ด	เอ็ด	เอ็ด	เจ็ด	เจ็ด	*	*	2
เจ็บ	โกรธ	เจ็บ	เอ็ด	โกรธ	*	เจ็บ	เจ็บ	เอ็ด	เอ็ด	เจ็บ	4
ใจ	ใจ	ใจ	ใจ	ใจ	ใจ	ไป	ใจ	ไป	ใจ	ใจ	8
ฉัน	เขา	ทำ	ฉัน	ฉัน	พัน	ฉัน	ฉัน	ฉัน	ผม	ขา	5
ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	ชอบ	10
ชาย	*	การ	ชาย	*	ชาย	ชาย	ไป	การ	ชาย	ชาย	5
ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ชื่อ	ที่	ชื่อ	ชื่อ	ชื่อ	ชื่อ	9
เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	เช้า	10
ดี	ดี	ดี	ดี	หน้า	ดี	ดี	ดี	ดี	ดี	ยี่	8
เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	เด็ก	10
ตา	ปาก	ปาก	อะ	ดา	ตา	ปาก	ปาก	ปาก	ตา	อะ	3
ถึง	ถึง	ถึง	ถึง	ถึง	ถึง	ถึง	ถึง	ถึง	ถึง	ถึง	10
ท้อง	ท้อง	ท้อง	ท้อง	ท้อง	ท้อง	*	ท้อง	ท้อง	ท้อง	ท้อง	9
ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	ทำ	10
ที่	ที่	ที่	ที่	ที่	ที่	ที่	ที่	ที่	ที่	ที่	10
เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	เท่า	10
เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	เที่ยง	10
โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	โทษ	10

ตาราง 4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	เธอ	10
น้ำ	น้ำ	น้ำ	น้ำ	น้ำ	หน้า	การ	น้ำ	น้ำ	น้ำ	น้ำ	8
บาท	บาท	ปาก	บาท	บาท	บาท	บาท	บาท	ปาก	บาท	บาท	8
บ้าน	บ้าน	บ้าน	บ้าน	*	ปาก	ไม่	บ้าน	บ้าน	*	บ้าน	6
บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	บ่าย	10
เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	เบื่อ	10
ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	ปวด	10
ปาก	อะ	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	ปาก	9
เป็น	การ	เป็น	*	เป็น	*	*	เป็น	เป็น	*	เป็น	5
แปด	อะ	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	แปด	9
ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	ไป	10
ผม	ผม	ผม	ผม	ผม	ผม	*	ผม	ผม	ผม	ผม	9
พบ	*	ปวด	*	พบ	พบ	พบ	พบ	พบ	*	พบ	6
พ่อ	*	*	*	พ่อ	พ่อ	ดา	ตา	พ่อ	ตา	พ่อ	4
พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	พัก	10
พัน	พัน	ปาก	ไป	ไป	ไป	พัน	พัน	พัน	ไป	ไป	4
เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	ขึ้น	เพื่อน	เพื่อน	เพื่อน	เพื่อน	เพื่อน	9
ไฟ	ไฟ	ไป	ไฟ	ไฟ	ไฟ	ไฟ	ไฟ	ไฟ	*	ไฟ	8

ตาราง 4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
มา	มา	มา	บาท	ปาก	มา	บาท	ปาก	มา	มา	มา	6
มาก	มา	มา	ปาก	มาก	มา	มาก	มาก	มาก	มาก	มาก	6
มือ	เมื่อ	มือ	มือ	ชื่อ	ชื่อ	มือ	ชื่อ	ชื่อ	เมื่อ	มือ	4
เมล์	เมล์	โรง	เมล์	เรียน	เมล์	เรียน	เมล์	เมล์	เมล์	เมล์	7
เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	เมื่อ	10
แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	แม่	10
ไม	ไป	ไม	ไป	ไป	ไม	ไป	ไป	ไป	ไม	ไม	4
ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	ไม่	*	ไม่	ไม่	9
ยา	ยา	*	*	ยา	อยาก	อยาก	อยาก	อยาก	ยา	ยา	4
ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	ยี่	10
เย็น	ลืม	ขึ้น	เย็น	เย็น	ลืม	ลืม	ลืม	เย็น	เย็น	ลืม	4
รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	รถ	10
ร้อน	ร้อน	ทำ	คน	ร้อน	ทำ	ร้อน	ร้อน	ร้อน	ร้อน	*	6
ร้อย	ร้อย	ร้อย	รถ	ร้อย	ร้อย	*	ร้อย	ร้อย	ร้อย	ร้อย	8
เรียน	เย็น	เย็น	เย็น	เย็น	เรียน	เย็น	เรียน	เย็น	เย็น	เย็น	2
โรง	ลง	*	โรง	ลง	โรง	โรง	โกรธ	โรง	โรง	โรง	6
ไร	ไป	ไร	ไป	ไป	ไป	*	ไป	ไป	ไป	ไป	1

ตาราง 4.6 4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
ลง	เธอ	ลง	ลง	ลง	*	*	ลง	*	*	ลง	5
ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	ลืม	10
เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	เลือด	10
วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	วัด	10
วัน	วัน	วัน	หลัง	วัน	หลัง	หลัง	วัน	วัน	ไร	วัน	6
วิ่ง	วิ่ง	วิ่ง	วิ่ง	ไม่	วิ่ง	วิ่ง	วิ่ง	วิ่ง	ถึง	ไม่	7
สอง	สอง	สอง	สอง	สอง	*	*	สอง	สอง	สอง	สอง	8
สาม	ปาก	ปาก	ปาก	สาม	สาม	ขา	ปาก	ขา	ขา	ขา	2
สิบ	สิบ	คิด	เจ็ด	สิบ	คิด	*	*	สิบ	เอ็ด	สิบ	4
สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	สี่	10
สุข	สุข	สุข	สุข	สุข	สุข	หญิง	สุข	สุข	สุข	สุข	9
เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	เสีย	10
หก	หก	หก	หก	หก	โกรธ	หก	หก	หก	หก	หก	9
หญิง	หญิง	หญิง	*	เมื่อ	หญิง	*	หญิง	หญิง	ยี่	หนึ่ง	5
หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	หน้า	10
หนาว	*	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	หนาว	9
หนึ่ง	*	*	มือ	หนึ่ง	หนึ่ง	*	มือ	หนึ่ง	หนึ่ง	หนึ่ง	5
หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	หมอ	10

ตาราง 4.8 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล (ต่อ)

ครั้งที่ / พยางค์	1	2	3	4	5	6	7	8	9	10	จำนวนพยางค์ที่ถูกต้อง
หลัง	หลัง	หลัง	ปาก	หลัง	หลัง	*	*	*	หลัง	หลัง	6
หัว	หัว	หัว	หัว	หัว	หัว	หัว	หัว	วัด	หัว	หัว	9
ห้า	ค่า	ห้า	ห้า	ห้า	ห้า	ห้า	ค่า	ห้า	หน้า	ห้า	7
หาร	หาร	หาร	หาร	ขา	หาร	คน	ขา	หาร	หาร	หาร	7
ไหน	เรียน	เรียน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	ไหน	8
อยาก	*	อย่าง	อยาก	อย่าง	อย่าง	อย่าง	อยาก	อยาก	อย่าง	อย่าง	3
อย่าง	อย่าง	อย่าง	อย่าง	*	อย่าง	เก้า	ยา	*	อย่าง	*	5
อะ	อะ	แปด	อะ	อะ	อะ	อะ	อะ	อะ	อะ	อะ	9
อา	อา	การ	อา	อา	อา	*	อา	อา	อา	อา	8
เอ็ด	อา	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	เอ็ด	9

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 4.9 สรุปผลการทดลอง

พยางค์ที่	ชื่อ / พยางค์ที่ทดลอง	รุ่งโรจน์	ปัทมา	ชาญวิศ	เฉลี่ย
1	กลับ	10	10	9	9.67
2	การ	10	9	9	9.33
3	กิน	9	8	3	6.67
4	เกลียด	10	9	6	8.33
5	เก้า	9	10	10	9.67
6	โกรธ	10	6	9	8.33
7	ขอ	2	5	9	5.33
8	ขอบ	9	9	10	9.33
9	ขา	10	6	10	8.67
10	ข้าว	7	4	10	7.00
11	ขึ้น	9	10	10	9.67
12	เขา	9	3	10	7.33
13	แขน	10	3	8	7.00
14	คน	6	0	6	4.00
15	ครู	10	4	10	8.00
16	ค่า	1	3	9	4.33
17	คิด	8	6	8	7.33
18	คุณ	6	7	3	5.33
19	ใคร	0	0	5	1.67
20	เจ็ด	4	7	2	4.33
21	เจ็บ	5	0	4	3.00
22	ใจ	7	7	8	7.33
23	ฉัน	4	8	5	5.67
24	ชอบ	8	0	10	6.00
25	ชาย	3	7	5	5.00
26	ชื่อ	10	9	9	9.33
27	เช้า	10	10	10	10.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.9 สรุปผลการทดลอง (ต่อ)

พยางค์ที่	ชื่อ / พยางค์ที่ทดลอง	รุ่งโรจน์	ปัทมา	ชาญวิศ	เฉลี่ย
28	ดี	9	9	8	8.67
29	เด็ก	10	5	10	8.33
30	ตา	4	3	3	3.33
31	ถึง	4	7	10	7.00
32	ท้อง	6	1	9	5.33
33	ทำ	9	7	10	8.67
34	ที่	9	5	10	8.00
35	เท่า	10	10	10	10.00
36	เที่ยง	7	10	10	9.00
37	โทษ	9	10	10	9.67
38	เธอ	10	8	10	9.33
39	น้ำ	8	3	8	6.33
40	บาท	9	1	8	6.00
41	บ้าน	6	9	6	7.00
42	บ่าย	9	5	10	8.00
43	เบื่อ	10	1	10	7.00
44	ปวด	10	10	10	10.00
45	ปาก	10	8	9	9.00
46	เป็น	9	10	5	8.00
47	แปด	10	10	9	9.67
48	ไป	10	8	10	9.33
49	ผม	9	6	9	8.00
50	พบ	3	7	6	5.33
51	พ่อ	10	10	4	8.00
52	พัก	10	10	10	10.00
53	พัน	5	6	4	5.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.9 สรุปผลการทดลอง (ต่อ)

พยางค์ที่	ชื่อ / พยางค์ที่ทดลอง	รุ่งโรจน์	ปัทมา	ชาญวิศ	เฉลี่ย
54	เพื่อน	10	9	9	9.33
55	ไฟ	9	5	8	7.33
56	มา	3	6	6	5.00
57	มาก	1	0	6	2.33
58	มือ	9	5	4	6.00
59	เมล็	10	9	7	8.67
60	เมื่อ	10	8	10	9.33
61	แม่	2	10	10	7.33
62	ไม	5	4	4	4.33
63	ไม่	3	10	9	7.33
64	ยา	10	2	4	5.33
65	ยี่	5	10	10	8.33
66	เย็น	4	6	4	4.67
67	รถ	10	2	10	7.33
68	ร้อน	9	3	6	6.00
69	ร้อย	9	7	8	8.00
70	เรียน	7	10	2	6.33
71	โรง	10	8	6	8.00
72	ไร	8	7	1	5.33
73	ลง	0	0	5	1.67
74	ลืม	9	10	10	9.67
75	เลือด	9	9	10	9.33
76	วัด	10	10	10	10.00
77	วัน	5	10	6	7.00
78	วิ่ง	10	9	7	8.67
79	สอง	8	8	8	8.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.9 สรุปผลการทดลอง (ต่อ)

พยางค์ที่	ชื่อ / พยางค์ที่ทดลอง	รุ่งโรจน์	ปัทมา	ชาญวิศ	เฉลี่ย
80	สาม	8	2	2	4.00
81	สิบ	9	6	4	6.33
82	สี่	10	7	10	9.00
83	สุข	10	10	9	9.67
84	เสีย	10	8	10	9.33
85	หก	7	10	9	8.67
86	หญิง	9	10	5	8.00
87	หน้า	10	10	10	10.00
88	หนาว	10	4	9	7.67
89	หนึ่ง	2	5	5	4.00
90	หมอ	10	4	10	8.00
91	หลัง	2	0	6	2.67
92	หัว	10	9	9	9.33
93	ห้า	7	6	7	6.67
94	หาร	4	3	7	4.67
95	ไหน	10	10	8	9.33
96	อยาก	4	0	3	2.33
97	อย่าง	2	2	5	3.00
98	อะ	8	8	9	8.33
99	อา	8	1	8	5.67
100	เอ็ด	9	10	9	9.33
	ค่าเฉลี่ย	7.55	6.41	7.59	7.18
	%ความถูกต้อง	75.50	64.10	75.90	71.83

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.10 จำนวนพยางค์ที่อยู่ในแต่ละช่วงเปอร์เซ็นต์ความถูกต้อง

% ความถูกต้อง	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
จำนวนพยางค์	0	2	5	4	8	12	11	19	14	25

1) ค่าเปอร์เซ็นต์ความถูกต้องเฉลี่ย เท่ากับ 71.83

2) พยางค์ที่รู้จำได้ถูกต้อง 100% ได้แก่ เช้า, เท่า, ปวด, พัก, วัด, หน้า

3) พยางค์ที่รู้จำได้ถูกต้องต่ำกว่า 50% ได้แก่ คน, ค่า, ใคร, เจ็ด, เจ็บ, ตา, มาก, ไม, เย็น, ลง, สาม, หนึ่ง, หลัง, หาร, อยาก, อย่าง

4) พยางค์ที่รู้จำได้ถูกต้องต่ำที่สุด คือ ใคร(1.67%), ลง (1.67%)

4.1.3 การทดลองระยะเวลาในการทำงานของโปรแกรมบนพีดีเอ

ทดลองโปรแกรมจับเวลาส่วนแปลงเสียงเป็นตัวอักษร โดยแบ่งการทดลองออกเป็น 5 ส่วนย่อย ได้แก่

ส่วนที่ 1 พูด(หนึ่ง สอง สาม) ทดลอง 5 ครั้ง

ส่วนที่ 2 พูด(หนึ่ง สอง สาม สี่ ห้า) ทดลอง 5 ครั้ง

ส่วนที่ 3 พูด(หนึ่ง สอง สาม สี่ ห้า หก เจ็ด) ทดลอง 5 ครั้ง

ส่วนที่ 4 พูด(หนึ่ง สอง สาม สี่ ห้า หก เจ็ด แปด เก้า สิบ) ทดลอง 5 ครั้ง

ส่วนที่ 5 พูด(หนึ่ง สอง สาม สี่ ห้า หก เจ็ด แปด เก้า สิบ หนึ่ง สอง สาม สี่ ห้า) ทดลอง 5 ครั้ง

จำนวนผู้ทดลอง 1 คน

สถานที่ : ห้องฮาร์ดแวร์

สภาพแวดล้อม : เงียบ (ไม่มีคน)

ขั้นตอนการทดลอง :

1) เปิดโปรแกรมจับเวลาส่วนการแปลงเสียงเป็นตัวอักษรบนพีดีเอ

2) คลิกที่ปุ่ม อัดเสียง ของโปรแกรม แล้วพูดตามพยางค์ในแต่ละส่วนย่อย จากนั้น คลิกที่ปุ่ม หยุด แล้วบันทึกค่าเวลาที่ใช้

พบว่าได้ผลการทดลองดังตาราง 4.11 และรูป 4.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.11 ผลการทดลองของ นายชาญวิศ ประสานเหลืองวิไล

ครั้งที่ / จำนวนพยางค์	1	2	3	4	5	เฉลี่ย	เฉลี่ย/พยางค์
3	3.048	2.875	2.777	2.979	2.904	2.916	0.97
5	3.016	3.056	3.273	3270	3.328	3.188	0.64
7	5.041	5.414	4.998	5142	5.140	5.147	0.74
10	5.646	5.262	5.302	5.588	5.555	5.470	0.55
15	6.856	5.669	7.239	7.008	5.706	6.495	0.43

รูป 4.1 กราฟความสัมพันธ์ระหว่างระยะเวลาที่พีดีเอประมวลผลต่อ1คำ กับจำนวนคำ

จากการทดลองระยะเวลาที่ใช้ในการทำงานของโปรแกรมแปลงเสียงเป็นตัวอักษรบนพีดีเอ พบว่า ระยะเวลาที่ใช้โดยเฉลี่ยต่อ 1 พยางค์ จะไม่เกิน 1 วินาที และยิ่งจำนวนพยางค์มาก ระยะเวลาที่ใช้โดยเฉลี่ยต่อ 1 พยางค์ จะลดลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 การทดลองส่วนการแปลงตัวอักษรเป็นเสียง

ทดลองจับเวลาที่ใช้ในการทำงานส่วนวิเคราะห์ข้อความ (Text Analysis) เพื่อแยกประโยคออกเป็นแต่ละพยางค์ โดยใช้ประโยคทดลอง 5 ประโยค แต่ละประโยคมี 10 พยางค์ ทดลองประโยคละ 5 ครั้ง

1) ประโยคที่ใช้ทดลอง 1 : ไปขึ้นรถเมล์ที่วัดหน้าบ้านเมื่อเช้า

2) ประโยคที่ใช้ทดลอง 2 : พ่อแม่พี่น้อง 321 คน มาหาข้าวกิน

3) ประโยคที่ใช้ทดลอง 3 : น้ำราคา 123 บาท โรงเรียนที่คุณรัก

4) ประโยคที่ใช้ทดลอง 4 : ผมชอบคุณแล้วคุณชอบผมบ้างรึเปล่า

5) ประโยคที่ใช้ทดลอง 5 : ปวดท้องมากเลยขอกินยาหน่อยครับ

จำนวนผู้ทดลอง : 1 คน

สถานที่ : ห้องฮาร์ดแวร์

ขั้นตอนการทดลอง :

1) พิมพ์ข้อความลงบนพีดีเอ

2) กดปุ่ม Send จากนั้นโปรแกรมจะทำการจับเวลาในขั้นตอนการวิเคราะห์ข้อความ (Text Analysis)

3) บันทึกผลการทดลอง

พบว่าได้ผลการทดลองดังตาราง 4.10

ตาราง 4.10 ผลการทดลองส่วนการแปลงตัวอักษรเป็นเสียง

ประโยคที่	ครั้งที่	เวลาที่ใช้ (มิลลิวินาที)
1	1	4
	2	4
	3	4
	4	4
	5	5
2	1	11
	2	11
	3	12
	4	12
	5	11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตาราง 4.10 ผลการทดลองส่วนการแปลงตัวอักษรเป็นเสียง (ต่อ)

ประโยคที่	ครั้งที่	เวลาที่ใช้ (มิลลิวินาที)
3	1	7
	2	6
	3	7
	4	7
	5	7
4	1	12
	2	12
	3	12
	4	12
	5	12
5	1	13
	2	13
	3	13
	4	13
	5	13
	เฉลี่ย	9.48

การทดลองระยะเวลาที่ใช้ในการทำงานส่วนวิเคราะห์ข้อความ (Text Analysis) โดยเฉลี่ยเท่ากับ 9.48 มิลลิวินาที

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5
สรุป

5.1 สรุป

สิ่งที่ทำในโครงงานนี้ ได้แก่

1) สำรวจความต้องการของของผู้ใช้งาน

2) คัดเลือกคำไทยจำนวน 100 พยางค์ ที่จะใช้ฝึกสอนจากหนังสือภาษามือไทย และบันทึกเสียงจากสมาชิกทั้งสามคน พยางค์ละ 50 ครั้ง

3) ศึกษาทฤษฎีต่างๆที่จะใช้ในการจัดการกับคลื่นเสียงก่อนจะนำไปฝึกสอน ดังนี้

3.1) ซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอจี

3.2) รีแซมปลิง

3.3) นอร์มอลไลเซชัน

3.4) สัมประสิทธิ์เซปสตรอลบนความถี่เมล

4) เขียนโปรแกรมจากทฤษฎี ซิลลาเบล เซกเมนเตชัน เบส ออน เอ็นเนอจี เพื่อตรวจจับคำพูดจากไฟล์เสียง และนำไฟล์เสียงทั้งหมดมาทดลองตัดคำ

5) ศึกษาการทำงานของโครงข่ายประสาทเทียม และหารูปแบบที่เหมาะสมในการรู้จำเสียงภาษาไทยโดยไม่ขึ้นกับผู้พูด ซึ่งโครงงานนี้ได้ใช้โครงข่ายประสาทเทียมโดยใช้กฎการเรียนรู้แบบแบคพรอพพาเกชั่น ซึ่งมี อินพุทโหนด 280โหนด ฮิดเดนเลเยอร์ 560 โหนด และเอาท์พุทเลเยอร์ 100 โหนด

6) ทำการฝึกสอนนิวรอนเนตเวิร์ค 100 พยางค์ จำนวน 30000 รอบ จนกระทั่งได้ค่าความผิดพลาด 0.00055 %

7) ประสิทธิภาพของระบบ

7.1) การตัดพยางค์ที่เป็นเนื้อเสียง ตัดได้ถูกต้อง 91.21%

7.2) การรู้จำเสียงรู้จำได้ถูกต้อง 71.83%

7.3) ความเร็วในการทำงานของโปรแกรมบนพีดีเอ เท่ากับ 0.67 วินาที/คำ

5.2 ปัญหาและอุปสรรค

1) เนื่องจากเป็นโปรแกรมที่พัฒนาบนพีดีเอ จึงไม่สามารถใช้โปรแกรมแมทแลป (MATLAB) ช่วยประมวลผลได้ และต้องเขียนโปรแกรมประมวลผลในกระบวนการต่างๆเองเกือบทั้งหมด

2) การเก็บตัวอย่างทำได้ล่าช้าเนื่องจากต้องอาศัยความร่วมมือจากบุคคลอื่น และต้องเก็บ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างจำนวนมากซึ่งต้องใช้ระยะเวลานาน

3) ไม่มีลิขสิทธิ์ของโปรแกรมนิวโรโซลูชัน รุ่นล่าสุด ซึ่งสามารถฝึกสอนได้เร็วกว่า

4) โปรแกรมและไลบารี่บางอย่างไม่สามารถนำมาใช้งานบนพีดีเอได้ เช่น ดีแอลแอล (DLL) ของวาจา

5) ผู้ทำการทดลองไม่มีพื้นฐานการทำงานเรื่องเสียงมาก่อนจึงต้องใช้เวลานานในการทำความเข้าใจข้อมูลต่างๆ

6) การรู้จำเสียงภาษาไทยอย่างมีประสิทธิภาพ ต้องผ่านหลายกระบวนการที่ไม่เคยเรียนรู้มาก่อน

5.3 แนวทางการแก้ไข

1) เราจะใช้แมทแลปในการทดลองบนเครื่องคอมพิวเตอร์ แต่ต้องศึกษาการเขียนโปรแกรมบนพีดีเอเพิ่มเติม เพื่อใช้เขียนแอพพลิเคชั่นบนพีดีเอ

2) เก็บตัวอย่างเสียงจากสมาชิกในกลุ่ม แต่เพิ่มจำนวนครั้งเป็นพยางค์ละ 50 ครั้งต่อคน

3) ใช้โปรแกรมนิวโรโซลูชันรุ่นเก่าในการฝึกสอนแทน

4) เขียนโปรแกรมขึ้นมาใช้บนพีดีเอเองโดยใช้ ซีชาร์ป (C#)

5) ศึกษาข้อมูลเพิ่มเติมจากหลายแหล่งข้อมูล เพื่อเพิ่มพูนความชำนาญและความเข้าใจให้มากขึ้น

5.4 แนวทางการพัฒนาต่อ

1) เก็บเสียงตัวอย่างจำนวนมากขึ้น จากบุคคลทั่วไปมากขึ้น และเพิ่มจำนวนครั้งของแต่ละพยางค์มากขึ้น

2) ศึกษาโครงสร้างที่เหมาะสมในการรู้จำเสียงมากกว่าโครงข่ายประสาทเทียม เช่น ฮิดเดนมาคอฟ (Hidden Makov Model) เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

Michael Negnevitsky. 2005. **Artificial Intelligence A Guide to Intelligent Systems.** 2nd ed. Great Britain : Biddles Ltd, King's Lynn.

Nutthacha Jittiwarangkul, Somchai Jitapunkul, Sudaporn Luksaneeyanawin, Visarut Ahkuputra, Chai Wutiwiwatchai. 1998. "Thai Syllable Segmentation for Connected Speech Based on Energy." 169-172. **in Digital Signal Processing Research Laboratory, Department of Electrical Engineering, Faculty of Engineering**, Bangkok: Chulalongkorn University.

Prathumthan, T. 1987. "Thai Speech Recognition using Syllable Unit." Master's thesis of Chulalongkom University.

เสรี ปานชาง. 2540. **"การรู้จำเสียงพูดคำไทยแบบไม่ขึ้นกับผู้พูดด้วยนิวรัลเน็ตเวอร์ค."** วิทยานิพนธ์ วิศวกรรมศาสตรมหาบัณฑิต สาขาวิศวกรรมไฟฟ้าบัณฑิตวิทยาลัย, สถาบันเทคโนโลยี พระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ไชยันต์ สุวรรณชีวะศิริ. 2549. **การรู้จำเสียงพูดตัวเลขภาษาไทยแบบหลายผู้พูดด้วยนิวรัลเน็ตเวอร์ค.** [Online].Available : http://202.28.17.1/article/atc41/atc00233.html.

Julius O. Smith III. 2009. "The Digital Audio Resampling Home Page." Center for Computer Research in Music and Acoustics (CCRMA), Stanford University.

Shutimun Limpanakorn,Duangkaew Sawamiphakdi,Chularat Tanprasert. 2001. "Voice Articulator for Thai Speaker Recognition." 46-54. **in Thammasat Int.J.Sc.Tech.,Vol.6**.Pathum-thani: ThamrnasaUt niversity.

Ulla Yadhunandan, S. and Goutam Saha. 2003. **Modified Mel-Frequency Cepstral Coefficient.** West Bengal: Kharagpur.

Sigurdur Sigurdsson, Kaare Brandt Petersen,Tue Lehn-Schiøler. 2006. "Mel Frequency Cepstral Coefficients: An Evaluation of Robustness of MP3 Encoded Music." Richard Petersens Plads, Informatics and Mathematical Modelling Technical University of denmark

Rashidul Hasa ,Md. Mustafa Jamil,Golam Rabbani, Md. and Saifur Rahman, Md. 2004. "Speaker Identification Using Mel Frequency Cepstral Coefficients." 565-568. **In 3rd International Conference on Electrical & Computer Engineering**, Bangladesh: Dhaka

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

วิกิพิเดีย.2552.**การแปลงฟูริเยร์**.[Online].Available : http://th.wikipedia.org/wiki/การแปลงฟูริเยร์

วิกิพีเดีย.2552.**การแปลงโคไซน์ไม่ต่อเนื่อง**.[Online].Available : http://th.wikipedia.org/wiki/การแปลงโคไซน์ไม่ต่อเนื่อง

เอกชัย เนาวนิช. 2550. "โปรแกรมฝึกออกเสียงพยัญชนะไทยสำหรับผู้บกพร่องทางการได้ยินโดยใช้ โครงข่ายประสาทเทียม." วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาวิทยาการคอมพิวเตอร์และสารสนเทศ บัณฑิตวิทยาลัย, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

ณัฐวุฒิ ยิ้มสวัสดิ์ และ วิษณุ ลิ้ทวีสุข. 2548. **การสังเคราะห์เสียงพูดจากข้อความภาษาไทย.** [Online].Available : 202.28.94.55/web/320491/2548/web1/g16/word.doc

สมาคมหูหนวกแห่งประเทศไทย. 2547. **หนังสือภาษามือไทย**.5 เล่ม. พิมพ์ครั้งที่ 2. กรุงเทพฯ: บริษัท บพิธการพิมพ์ จำกัด

เสาวนีย์ เจริญพัฒนานุกูล ให้สัมภาษณ์, 20 กรกฎาคม 2552. ปัทมา พรสุขศิริ ผู้สัมภาษณ์.**การใช้ชีวิตประจำวันของผู้พิการทางหู.** สมาคมคนหูหนวกแห่งประเทศไทย.

เกียรติศักดิ์ แพ้วมานะกูล ให้สัมภาษณ์, 20 กรกฎาคม 2552. ปัทมา พรสุขศิริ ผู้สัมภาษณ์.**การใช้ชีวิตประจำวันของผู้พิการทางหู.** สมาคมคนหูหนวกแห่งประเทศไทย.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

หลักการตั้งชื่อไฟล์เสียง

ไฟล์เสียงจะตั้งชื่อนำหน้าด้วยอักษร S และตามด้วยตัวเลขหกหลัก ซึ่งสามหลักแรก เป็นการกำหนดหมายเลขให้กับเสียงของแต่ละคน ส่วนสามหลักหลัง คือ หมายเลขของไฟล์เสียงเรียงตามลำดับตัวอักษร เพื่อกันการสับสนว่าเลขไฟล์ใดเป็นคำอะไร เจ้าของเสียงคือใคร จะมีการเก็บlist ของเลขเสียงและเลขที่ของเจ้าของเสียงไว้

ตัวอย่างการตั้งชื่อไฟล์

S001001 หมายถึง เสียงของผู้บันทึกคนที่ 1	บันทึกคำว่า กลับ
S001002 หมายถึง เสียงของผู้บันทึกคนที่ 1	บันทึกคำว่า การ
S002001 หมายถึง เสียงของผู้บันทึกคนที่ 2	บันทึกคำว่า กลับ
S002002 หมายถึง เสียงของผู้บันทึกคนที่ 2	บันทึกคำว่า การ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

คำไทยจำนวน 100 พยางค์ที่ใช้ในการทดลอง

ข.1 คำสรรพนาม

ฉัน คุณ ผม เขา เธอ

ข.2 คำนาม

ขอบ	โทษ	ชื่อ	หนึ่ง	สอง	สาม	สี่	ห้า	หก	เจ็ด	แปด
เก้า	สิบ	ยี่	ร้อย	พัน	เอ็ด	เช้า	เที่ยง	บ่าย	คน	ชาย
หญิง	เด็ก	พ่อ	แม่	เพื่อน	ข้าว	หัว	หน้า	ตา	ปาก	แขน
มือ	ท้อง	หลัง	ขา	รถ	เมล์	ไฟ	วัด	โรง	บ้าน	อา
หาร	น้ำ	หมอ	ครู	การ	ใจ	เลือด	ยา	บาท	วัน	ค่า

ข.3 คำกริยา

เรียน	พบ	เกลียด	ชอบ	โกรธ	คิด	ลืม	ขอ	พัก	ถึง	ไป
มา	กลับ	วิ่ง	กิน	เป็น	ขึ้น	ลง				

ข.4 คำวิเศษณ์

อยาก	สุข	ร้อน	เย็น	ดี	เสีย	เบื่อ	หนาว	ไม่	เจ็บ	ปวด
มาก										

ข.5 คำแสดงคำถาม

อะ ใคร ที่ เมื่อ อย่าง ทำ เท่า ไร ไหน ไม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

บทสัมภาษณ์ผู้พิการทางหู

ค.1 บทสัมภาษณ์ผู้พิการคนที่1

ค.1.1 ประวัติส่วนตัว

ชื่อ นางสาวเสาวนีย์ เจริญพัฒนานุกูล

เพศ หญิง

อายุ 38 ปี

วันเกิด 16 ตุลาคม พ.ศ.2514

อาชีพ รับจ้าง

จำนวนพี่น้อง 3 คน

ที่อยู่ 66/568 หมู่ 5 ถนนเทพารักษ์ ตำบลบางเมือง อำเภอเมือง จังหวัดสมุทรปราการ 10270

การศึกษา ปริญญา

รายได้ 5,000-10,000 บาท

ค.1.2 บทสัมภาษณ์

ถาม : คนในครอบครัวหูหนวกด้วยหรือไม่

ตอบ : ไม่มี

ถาม : มีเพื่อนเยอะหรือไม่ เป็นคนหูหนวกหรือปกติ

ตอบ : เยอะ แต่ส่วนใหญ่เป็นคนหูหนวกด้วยกัน รู้จักกันทาง MSN และสื่อสารกันด้วยภาษามือผ่านทาง เว็บแคม (Webcam) มีเพื่อนชาวต่างชาติด้วย สื่อสารกันด้วยภาษามืออเมริกัน ซึ่งศึกษาจากโปรแกรมสอนภาษามืออเมริกัน

ถาม : ถ้าต้องสื่อสารกับคนที่ไม่เป็นภาษามือ สื่อสารกันอย่างไร มีปัญหาอย่างไรบ้าง

ตอบ : เขียนโต้ตอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถาม : เมื่อต้องออกไปข้างนอกแล้วหลงทาง จะทำอย่างไร

ตอบ : เขียนถาม ดังนั้นจึงต้องพกสมุดและดินสอติดตัวเสมอ บางครั้งก็ใช้มือถือพิมพ์ข้อความถาม

ถาม : เมื่อออกไปพบแพทย์ข้างนอกมีวิธีสื่อสารกับแพทย์อย่างไร

ตอบ : กรณีอาการไม่หนัก ก็ไปเอง แต่ถ้าอาการหนัก ต้องพาล่ามไปด้วย

ถาม : ในปัจจุบันมีเครื่องมือช่วยในการสื่อสารอะไรบ้าง

ตอบ : เครื่องช่วยฟัง พกติดตัวเสมอ เพื่อช่วยฟังเสียง เช่นเสียงเคาะโต๊ะ หรือเสียงแตร แต่เสียงพูดได้ยินไม่ชัด ราคาประมาณ หมื่นกว่าบาท แต่ได้ฟรีมาจากโครงการหนึ่ง

ถาม : เรียนภาษามือมาจากที่ไหน

ตอบ : จากเรียนที่โรงเรียนโสตศึกษา จ.ชลบุรี

ถาม : ตอนนี้ทำงานที่ไหน เคยถูกกีดกันเวลาสมัครงานไหม

ตอบ : ทำงานที่สมาคมคนหูหนวกแห่งประเทศไทย มีหน้าที่บริการสมาชิกและธุรการ

ถาม : คุณมีความรู้ในการใช้เทคโนโลยีมากน้อยแค่ไหน

ตอบ : ใช้มือถือในการส่งเอสเอ็มเอส (SMS) คุยกับเพื่อนๆ นอกจากนี้ยังใช้ใช้คอมพิวเตอร์อินเตอร์เน็ต และแคมฟอร์ก (Cam frog) ในการสื่อสารกับเพื่อนๆ

ถาม : หากมีอุปกรณ์พกพาที่สามารถแปลงเสียงเป็นตัวอักษร จะช่วยให้ชีวิตประจำวันสะดวกขึ้นหรือไม่

ตอบ : ช่วยให้สะดวกขึ้น เพราะปกติเวลาดูโทรทัศน์จะไม่สามารถเข้าใจช่องที่ไม่มีภาษามือได้

ถาม : หากมีระบบส่งเสียงกลับไปยังผู้พูด คิดว่าจะทำให้การสื่อสารสะดวกขึ้นหรือไม่

ตอบ : สะดวกขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค.2 บทสัมภาษณ์ผู้พิการคนที่2

ค.2.1 ประวัติส่วนตัว

ชื่อ นายเกียรติศักดิ์ แพ้วมานะกุล

เพศ ชาย

อายุ 44 ปี

วันเกิด 23 มกราคม พ.ศ.2508

อาชีพ งานด้านคอมพิวเตอร์

จำนวนพี่น้อง 4 คน

ที่อยู่ 139/121 ถนนประชาอุทิศ อำเภอทุ่งครุ กรุงเทพฯ

การศึกษา อนุปริญญา

รายได้ 5,000-10,000 บาท

ค.2.2 บทสัมภาษณ์

ถาม : คนในครอบครัวหูหนวกด้วยหรือไม่

ตอบ : ไม่มี

ถาม : พ่อแม่ทราบได้อย่างไรว่าหูหนวก

ตอบ : ไม่ได้หูหนวกแต่กำเนิด เพิ่งหูหนวกเพราะโดนรถชนตอนอายุสี่ขวบ สื่อสารกับคนในบ้านด้วยการอ่านปากและใช้ภาษามือที่คิดขึ้นเอง เพราะคนในบ้านไม่รู้ภาษามือ

ถาม : ถ้าต้องสื่อสารกับคนที่ไม่รู้ภาษามือ สื่อสารกันอย่างไร มีปัญหาอย่างไรบ้าง

ตอบ : เขียนโต้ตอบ ปัญหาคือกว่าจะสื่อสารรู้เรื่องค่อนข้างล่าช้า

ถาม : เมื่อออกไปข้างนอกแล้วหลงทาง ทำอย่างไร

ตอบ : ส่งข้อความถามเพื่อน

ถาม : เมื่อออกไปพบแพทย์ข้างนอกมีวิธีสื่อสารกับแพทย์อย่างไร

ตอบ : ไปกับคุณแม่และใช้ภาษาร่างกาย ในการอธิบาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถาม : ในปัจจุบันมีเครื่องมือช่วยในการสื่อสารอะไรบ้าง

ตอบ : มีเครื่องช่วยฟัง แต่ไม่ค่อยใช้

ถาม : เรียนภาษามือมาจากที่ไหน

ตอบ : เรียนที่โรงเรียนโสตศึกษาทุ่งมหาเมฆ ตั้งแต่ ป.1- ม.6

เรียนคอมพิวเตอร์จากโรงเรียน เทคนิคธนบุรี (เรียนร่วมกับคนปกติ)

ถาม : ทำงานที่ไหน เคยถูกกีดกันเวลาสมัครงานหรือไม่

ตอบ:

ประวัติการทำงาน

-ทำงานที่สมาคมคนหูหนวกแห่งประเทศไทย หน้าที่ดูแลงานด้านคอมพิวเตอร์ 1 ปี

-สอนภาษามือ 2ปี

-ปัจจุบันทำงานที่สมาคมคนหูหนวกแห่งประเทศไทย หน้าที่ดูแลงานด้านคอมพิวเตอร์

-เคยเป็นวิทยากรสอนคอมพิวเตอร์ให้คนหูหนวก

-เคยทำงานเป็นพนักงานเย็บกระเป๋าในโรงงาน ซึ่งตอนสัมภาษณ์งานถูกกีดกันเมื่อรู้ว่าเป็นคนหูหนวกแต่มีคนมาช่วยรับรองให้ว่าทำงานได้ จึงได้เข้าทำงาน

ถาม : คุณมีความรู้ในการใช้เทคโนโลยีมากน้อยแค่ไหน

ตอบ : รู้จักโปรแกรมคอมพิวเตอร์ต่างๆและสามารถใช้งานได้

ถาม : หากมีอุปกรณ์พกพาที่สามารถแปลงเสียงเป็นตัวอักษร จะช่วยให้ชีวิตประจำวันสะดวกขึ้นหรือไม่

ตอบ : ช่วยให้สะดวกขึ้น เพราะปกติเวลาดูโทรทัศน์จะไม่สามารถเข้าใจช่องที่ไม่มีภาษามือได้

ถาม : หากมีระบบส่งเสียงกลับไปยังผู้พูด คิดว่าจะทำให้การสื่อสารสะดวกขึ้นหรือไม่

ตอบ : สะดวกขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ง
ข้อมูลเพิ่มเติมเกี่ยวกับผู้พิการทางหู

- ปัจจุบันมีเครื่องช่วยผู้พิการทางหู เป็นเครื่องขนาดพกพา บนหน้าจอจะแสดงรูปภาพต่างๆ เมื่อผู้ใช้งานกดก็จะมีเสียงพูดออกมา เช่น กดที่รูปข้าว ก็จะมีเสียง 'ข้าว'
- ความพิการทางหูแบ่งเป็น 5 ระดับ ตามการได้ยินของหู ตั้งแต่หูตึงน้อย ไปจนถึงหูหนวก คนพิการทางหูจดจำตัวอักษร เสมือนภาพๆหนึ่ง ดังนั้น พวกเขาจึงมักเขียนคำในประโยคสลับตำแหน่งกัน เช่น กินข้าวหรือยัง อาจเขียนเป็น ข้าวกินหรือยัง นอกจากนี้ อักษรที่มีหัวเช่น ด กับ ค อาจจำสลับกัน เป็นเรื่องปกติ
- ภาษาสำหรับผู้พิการทางหู จะไม่มีคำที่เป็นนามธรรม และคำเชื่อมต่างๆ เช่น จะ ก็ เป็นต้น และผู้พิการแต่ละคนจะรู้ศัพท์ไม่เท่ากัน คนที่ไม่ได้เรียนหนังสือ อาจใช้ภาษามือที่คิดขึ้นเองใช้สื่อสารเฉพาะกับคนที่บ้านเท่านั้น ส่วนคนที่เรียนหนังสือ ถ้าเรียนในโรงเรียนที่ใช้หนังสือคนละเล่มกัน ก็อาจใช้ภาษามือที่ต่างกันไปบ้างเล็กน้อย ภาษามือนั้นมีคนคิดขึ้นใหม่ทุกๆวัน แต่จะแพร่หลายหรือไม่ต้องขึ้นอยู่กับความนิยมของผู้ใช้ ถ้าแพร่หลายก็อาจถูกนำมาตีพิมพ์ในหนังสือเรียนภาษามือได้
- ภาษามือของแต่ละประเทศจะแตกต่างกัน แต่ก็มีภาษามือสากลด้วย ใช้ในการประชุมสากลของคนหูหนวก แต่จะไม่ใช้แพร่หลายเท่าภาษามือของแต่ละประเทศ

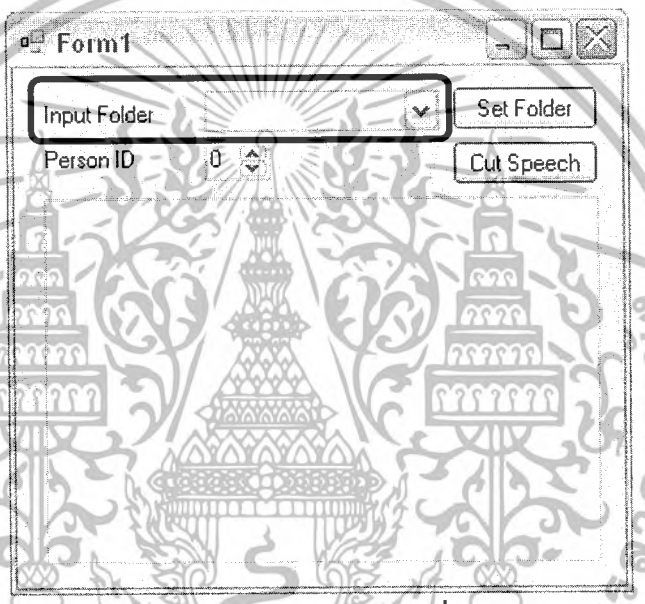
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก จ

การเตรียมข้อมูลเสียงก่อนนำไปฝึกสอนด้วยนิวรอนเนตเวิร์ค

จ.1 การตัดเสียงส่วนที่เป็นคำพูด ด้วยโปรแกรมดีเทค สปีช (Detect speech)

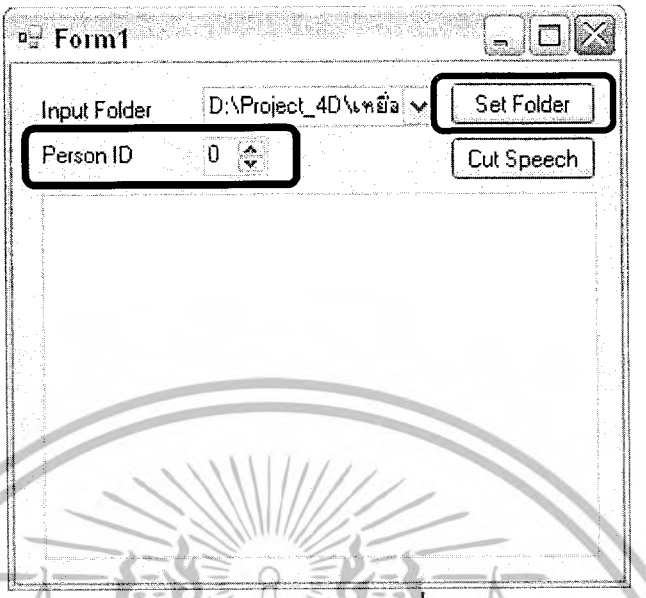
1) ในช่อง Input Folder เลือกโฟลเดอร์ที่ประกอบด้วย ไฟล์เสียง 100 ไฟล์ และไฟล์เสียงพื้นหลัง 100 ไฟล์



รูป จ.1.1 โปรแกรมตัดเสียงส่วนที่เป็นคำพูด

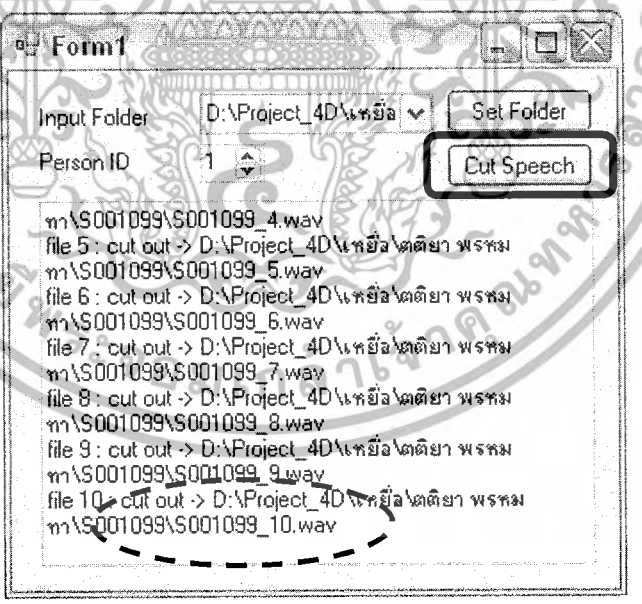
2) กดปุ่ม Set Folder และตั้งหมายเลขประจำตัวสำหรับแต่ละคน (Person ID)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป จ.1.2 โปรแกรมตัดเสียงส่วนที่เป็นคำพูด (2)

3) กดปุ่ม Cut Speech จากนั้นรอโปรแกรมตัดเสียง สังเกตหากโปรแกรมทำงานสมบูรณ์แล้วจะมีโฟลเดอร์ ตั้งแต่ 0-99

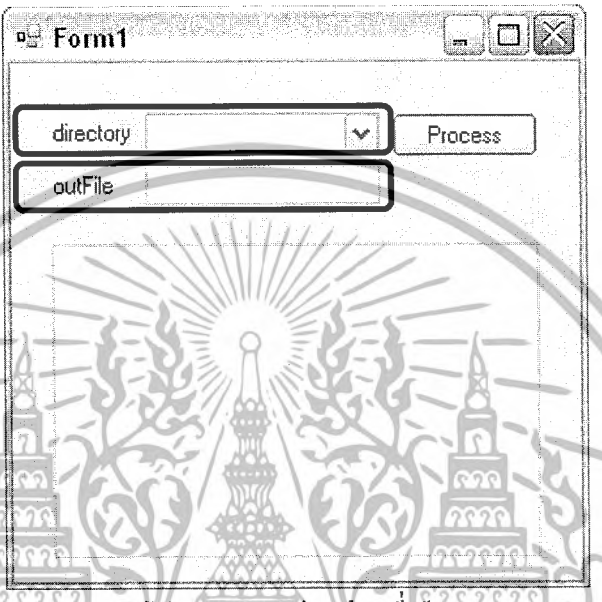


รูป จ.1.3 โปรแกรมตัดเสียงส่วนที่เป็นคำพูด (3)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จ.2 การนำไฟล์เสียง (ที่ผ่านโปรมแกรมข้อ 1 แล้ว) มาทำพรีโปรเซสซิง

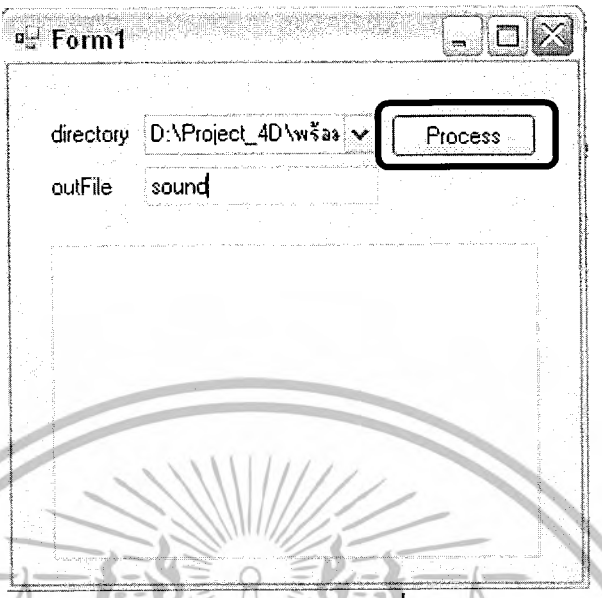
1) ในช่อง directory เลือกโฟลเดอร์ที่มีไฟล์เสียงที่ผ่านกระบวนการข้อ 1 มาแล้ว และตั้งชื่อไฟล์ในช่อง outFile ตามต้องการ



รูป จ.2.1 โปรแกรมตัดเสียงส่วนที่เป็นคำพูด (4)

2) กดปุ่ม Process เมื่อโปรแกรมทำงานเสร็จจะได้ค่า input และ output สำหรับนำไปฝึกสอนนิวรอนเน็ตเวิร์ค ทำเรื่อยไปจนครบทุกโฟลเดอร์เสียงที่ต้องการนำไปฝึกสอนด้วยนิวรอนเน็ตเวิร์ค

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

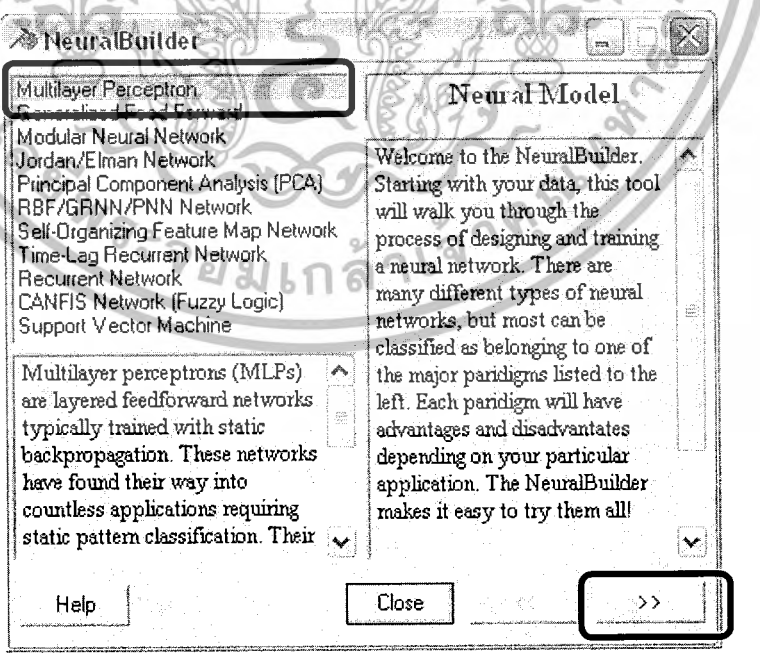


รูป จ.2.2 โปรแกรมตัดเสียงส่วนที่เป็นคำพูด (5)

จ.3 การฝึกสอนด้วยนิวรอน เนตเวิร์ค

1) เปิดโปรแกรมนิวโร โซลูชัน จากนั้นกดปุ่ม

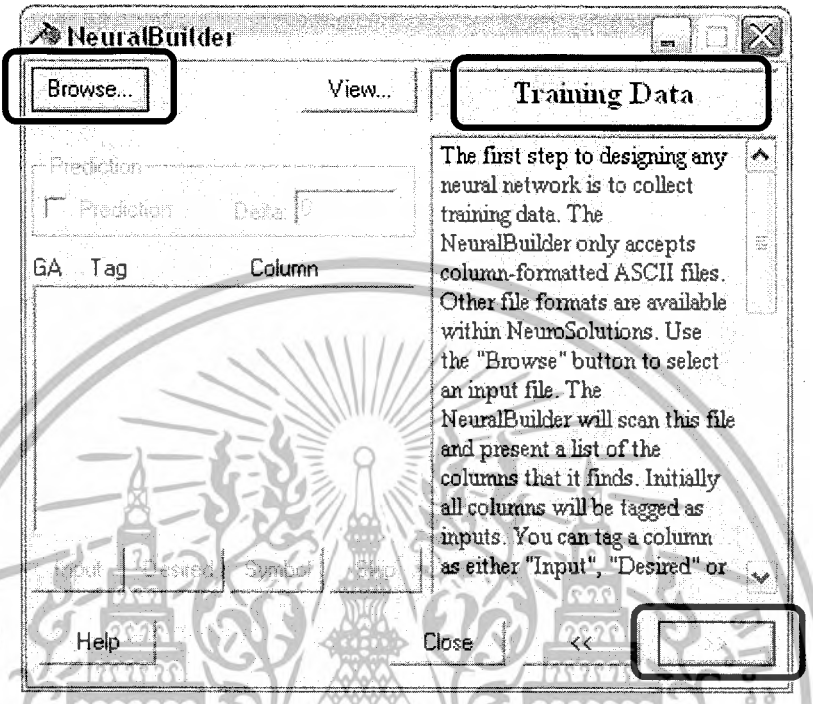
จะปรากฏหน้าต่าง ดังรูป ให้เลือก Multilayer Percepton จากนั้นกด ปุ่ม "ต่อไป"



รูป จ.3.1 โปรแกรมนิวโร โซลูชัน

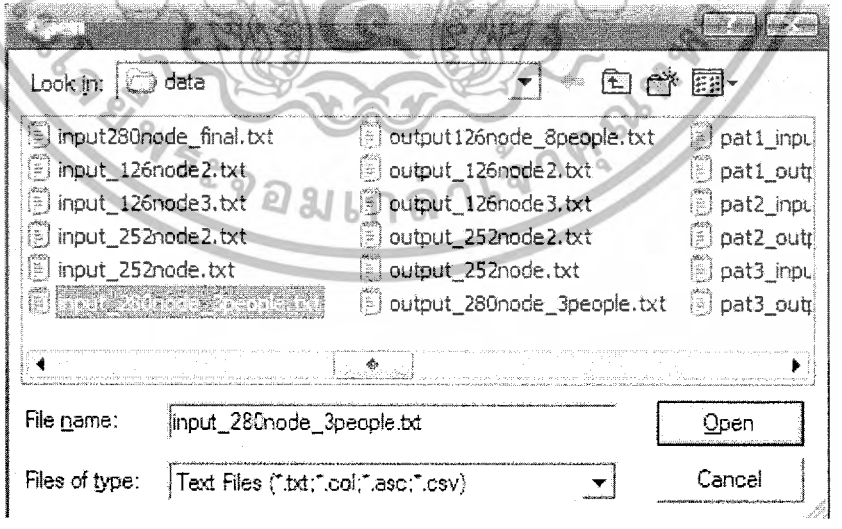
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) เมื่อปรากฏหน้าต่าง Training Data ดังรูป กดปุ่ม Browse



รูป จ.3.2 โปรแกรมนิวโร โซลูชัน (2)

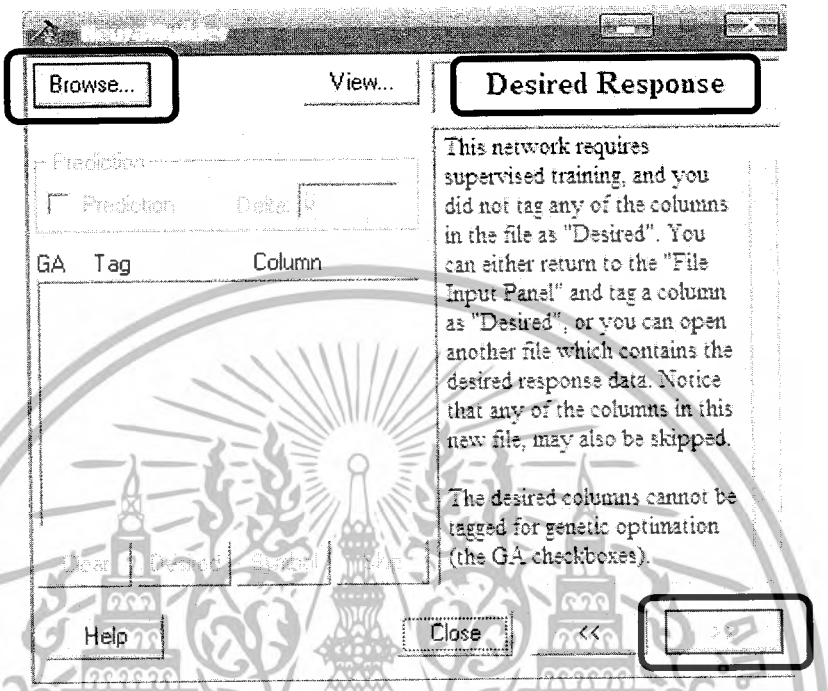
และเลือกไฟล์อินพุทที่รวมมาแล้ว จากนั้นกดปุ่ม"ต่อไป"



รูป จ.3.3 โปรแกรมนิวโร โซลูชัน (3)

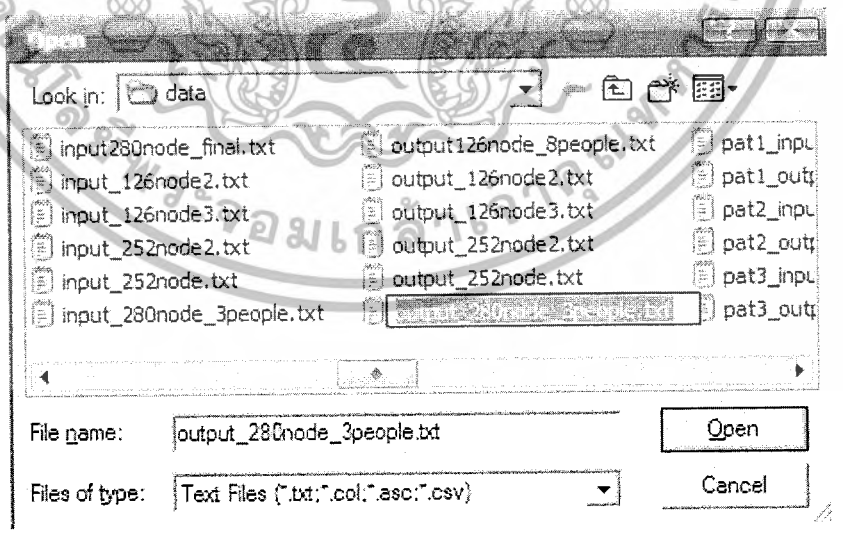
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) เมื่อปรากฏหน้าต่าง Desired Respondse ดังรูป กดปุ่ม Browse



รูป จ.3.4 โปรแกรมนิวโร โซลูชัน (4)

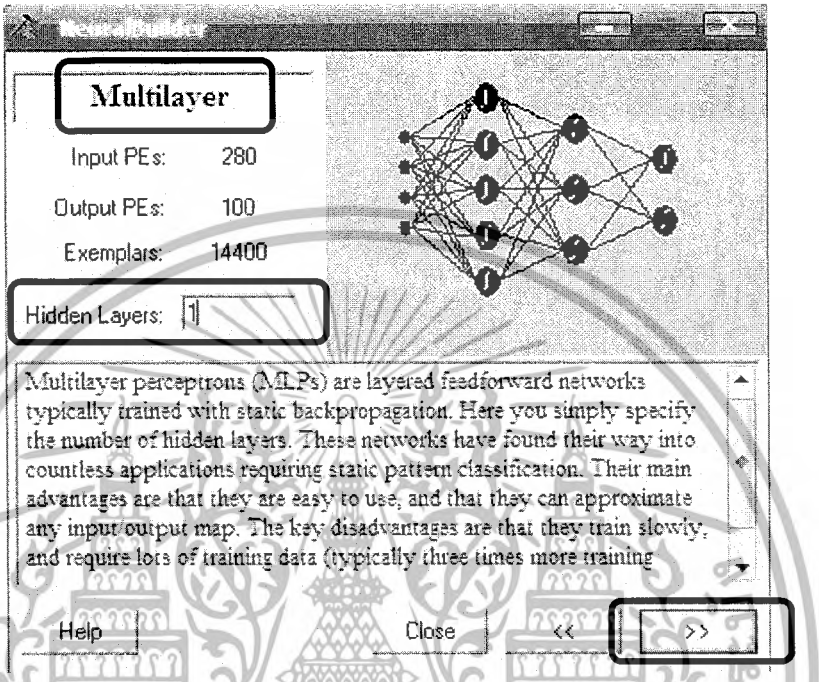
และเลือกไฟล์เอาท์พุทที่รวมมาแล้ว จากนั้นกดปุ่ม"ต่อไป"



รูป จ.3.5 โปรแกรมนิวโร โซลูชัน (5)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

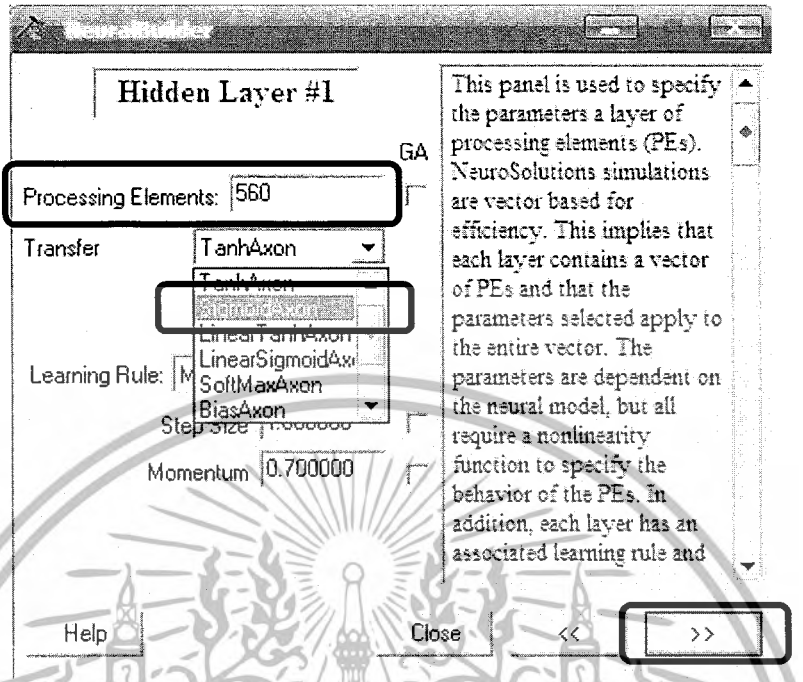
5) กดปุ่ม “ต่อไป” จนเจอหน้าต่าง Multilayer ในช่อง Hidden Layers เลือก 1 เลเยอร์ แล้วกดปุ่ม “ต่อไป”



รูป จ.3.6 โปรแกรมนิวโร โซลูชัน (6)

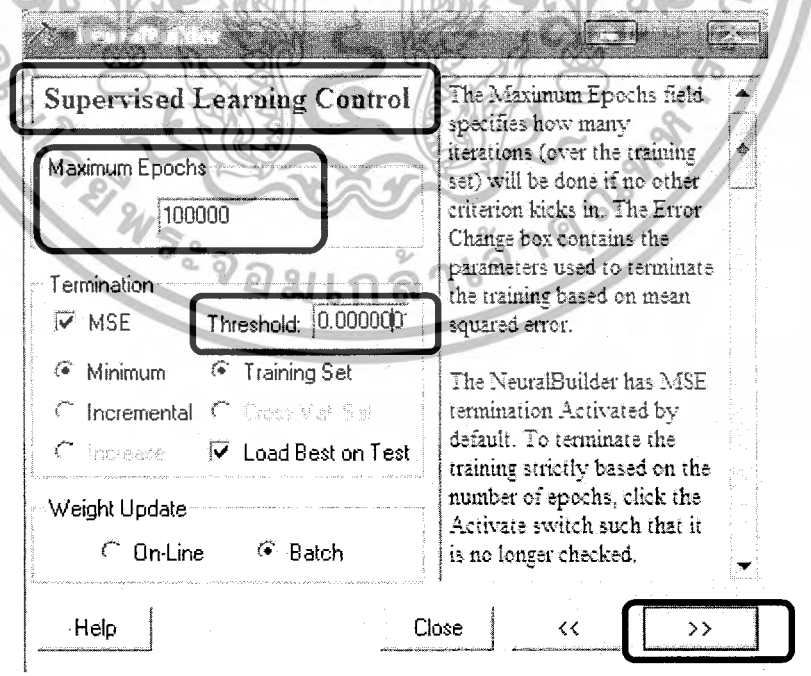
6) ในช่อง Processing Elements ใส่ค่าสองเท่าของโหนดอินพุท ในโครงงานนี้ คือ 560 ในช่อง Transfer เลือก SigmoidAxon แล้วกดปุ่ม “ต่อไป”

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป จ.3.7 โปรแกรมนิวโร โซลูชัน (7)

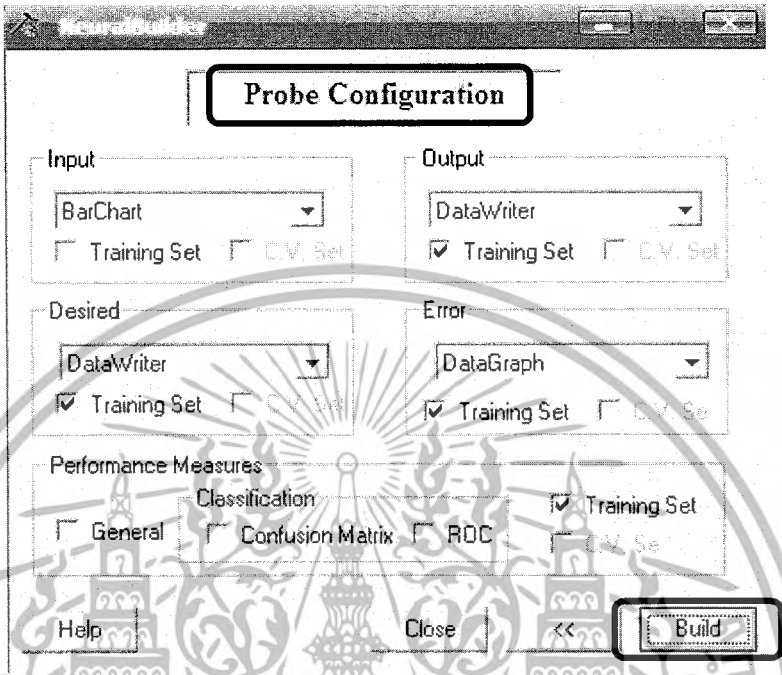
7) เมื่อเจอหน้า Supervised Learning Control ในช่อง Maximum Epochs ใส่ค่า 100,000 ส่วนค่า Threshold ตั้งค่าให้ต่ำมากๆประมาณ 0.000001 แล้วกดปุ่ม “ต่อไป”



รูป จ.3.8 โปรแกรมนิวโร โซลูชัน (8)

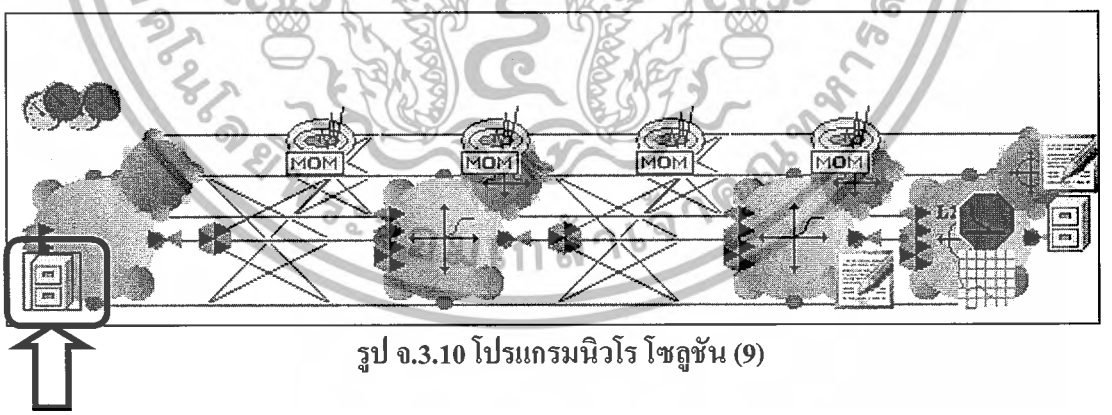
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8) เมื่อเจอหน้า Probe Configuration ให้กด Build เพื่อสร้างโครงข่ายประสาทเทียม



รูป จ.3.9 โปรแกรมนิวโร โซลูชัน (8)

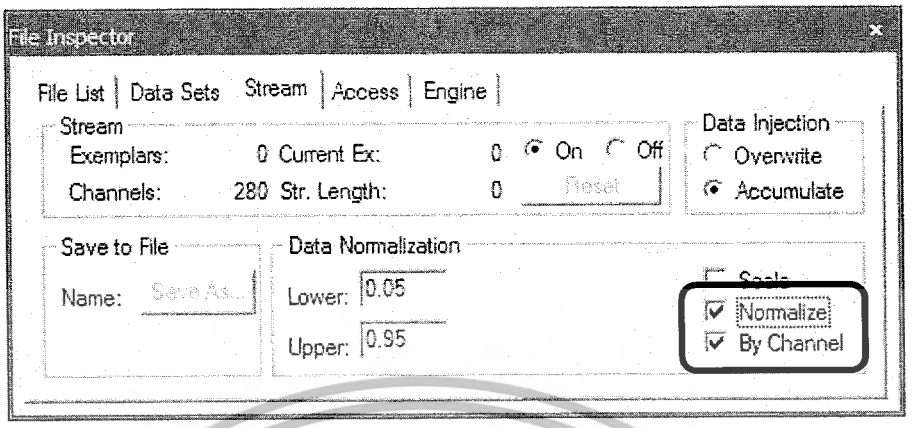
9) ดับเบิลคลิกเลือกที่รูปกล่องด้านซ้ายล่าง



รูป จ.3.10 โปรแกรมนิวโร โซลูชัน (9)

จากนั้น เมื่อปรากฏหน้าต่างข้างล่าง กดเลือกช่อง Normalize ออก ปิดหน้าต่างนี้และรอโครงข่ายปราสาทเทียมฝึกสอนจนเสร็จจึงนำค่าเวจน์ไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป จ.3.11 โปรแกรมนิวโร โซลูชัน (10)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้