



ระบบรู้จำเสียงสำหรับผู้พิการทางการได้ยินโดยใช้โครงข่ายประสาทเทียม  
SPEECH RECOGNITION SYSTEM FOR HEARING DISABILITY  
USING NEURAL NETWORK



คุณุตม์ ปิ่งศิริเจริญ

ระบบรู้จำเสียงสำหรับผู้พิการทางการได้ยินโดยใช้โครงข่ายประสาทเทียม



ปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมชีวการแพทย์  
คณะวิศวกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ  
ปีการศึกษา 2563  
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

SPEECH RECOGNITION SYSTEM FOR HEARING DISABILITY  
USING NEURAL NETWORK



KANUT PUENG SIRICHAROEN

A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of MASTER OF ENGINEERING  
(M.Eng. (Biomedical Engineering))  
Faculty of Engineering, Srinakharinwirot University

2020

Copyright of Srinakharinwirot University

ปริญญาานิพนธ์

เรื่อง

ระบบรู้จำเสียงสำหรับผู้พิการทางการได้ยินโดยใช้โครงข่ายประสาทเทียม

ของ

คุณุตม์ ปิ่งศิริเจริญ

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมชีวการแพทย์

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

คณบดีบัณฑิตวิทยาลัย

คณะกรรมการสอบปากเปล่าปริญญาานิพนธ์

ที่ปรึกษาหลัก

(ผู้ช่วยศาสตราจารย์ ดร.ดิเรก เสือสีนาค)

ประธาน

(รองศาสตราจารย์ ดร.ชูชาติ ปิณฑุจัน)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.สุชาดา ตันติสถิระพงษ์)

ชื่อเรื่อง	ระบบรู้จำเสียงสำหรับผู้พิการทางการได้ยินโดยใช้โครงข่ายประสาทเทียม
ผู้วิจัย	คุณุตม์ ปิ่งศิริเจริญ
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
ปีการศึกษา	2563
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร. ดิเรก เสือสีนาค

การฟัง ถือเป็นทักษะหนึ่งที่สำคัญในการเรียนรู้ภาษาและการพูดของมนุษย์ หากเกิดความบกพร่องในการติดต่อสื่อสารกับผู้อื่นอาจเกิดจากหลายสาเหตุ แต่สาเหตุสำคัญที่ทำให้เกิดความบกพร่องทางการพูดคือ การบกพร่องทางการฟังหรือการสูญเสียการได้ยิน ในงานวิจัยนี้จึงนำเสนอระบบรู้จำเสียงสำหรับฟื้นฟูการสื่อสารของคนที่มีภาวะหูหนวกทั้งหูสองข้าง ซึ่งมีคำศัพท์สำหรับฟื้นฟูการสื่อสารจำนวน 66 คำ สำหรับการหาลักษณะสำคัญของเสียงด้วยวิธีสัมประสิทธิ์เซปสตรัมของเมล (Mel-Frequency Cepstral Coefficient: MFCC) และสำหรับการจำแนกเสียงต่างๆ ด้วยวิธีการ โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN), โครงข่ายประสาทเทียม (Artificial Neural Network: ANN), โครงข่ายประสาท (Neural Network: NN) และการเรียนรู้เพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbor: KNN) จากผลการทดสอบด้วยชุดข้อมูลเสียงที่แบ่งเป็น 10 กลุ่มข้อมูล และ 20 กลุ่มข้อมูล วิธีการ ANN ได้ค่าความถูกต้องสูงที่สุดเมื่อเปรียบเทียบกับวิธีการอื่นๆ คือ 97.78% และ 86.11% ตามลำดับ สำหรับชุดข้อมูลเสียงที่แบ่งเป็น 30 กลุ่มข้อมูล, 40 กลุ่มข้อมูล, 50 กลุ่มข้อมูล, 60 กลุ่มข้อมูล และ 66 กลุ่มข้อมูล วิธีการ KNN ได้ค่าความถูกต้องสูงที่สุดเมื่อเปรียบเทียบกับวิธีการอื่นๆ คือ 89.63%, 81.67%, 82.67%, 85.37% และ 80.98% ตามลำดับ

คำสำคัญ : การสูญเสียการได้ยิน, ภาวะหูหนวก, สัมประสิทธิ์เซปสตรัมของเมล, โครงข่ายประสาทเทียม

Title	SPEECH RECOGNITION SYSTEM FOR HEARING DISABILITY USING NEURAL NETWORK
Author	KANUT PUENGSIRICHAROEN
Degree	MASTER OF ENGINEERING
Academic Year	2020
Thesis Advisor	Assistant Professor Dr. Direk Suseenak

Listening is one of the most important skills in learning human languages and in speech. However, some people have a disability in communicating with others for many reasons. One of the main causes of speech impairment is hearing impairment or hearing loss. This research proposes a speech recognition system for rehabilitation of communication in people with bilateral cochlear implant surgery. It includes the 66 words for the rehabilitation of communication. The Mel-Frequency Cepstral Coefficient (MFCC) was used for speech feature extraction. The Convolutional Neural Network (CNN), Artificial Neural Network (ANN), Neural Network (NN) and K-Nearest Neighbor (KNN) were used for speech classification. Based on the test results of the 10 classes and 20 classes datasets, the ANN algorithm achieved the highest accuracy compared to other algorithms, which were at 97.78% and 86.11%, respectively. For 30 classes, 40 classes, 50 classes, 60 classes and 66 classes datasets, the KNN algorithm achieved the highest accuracy compared to other algorithms, which were at 89.63%, 81.67%, 82.67%, 85.37% and 80.98%, respectively.

Keyword : Hearing loss, Cochlear implant, Mel-Frequency Cepstral Coefficient, Artificial neural networks

## กิตติกรรมประกาศ

ปริญญาานิพนธ์นี้สำเร็จลุล่วงได้ด้วยความสามารถจาก รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล คณบดีบัณฑิตวิทยาลัย

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ดิเรก เสือสีนาค และคณาจารย์ทุกท่าน ที่ได้ให้ความเมตตากรุณาเป็นที่ปรึกษาและให้ความช่วยเหลือชี้แนะแนวทางในสิ่งที่เป็นประโยชน์ต่อการศึกษและการทำปริญญาานิพนธ์นี้ด้วยความเอาใจใส่ตลอดมา

ขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ แพทย์หญิงจรินทร์ สิริรัฐวรรณ ดร.นงเยาว์ ศรีนางแย้ม และนางสาว ศิวดาดี ปิยะกัญญา ที่ได้ให้ความเมตตากรุณาเป็นที่ปรึกษาและให้ความช่วยเหลือชี้แนะแนวทางในการเก็บข้อมูลจากกลุ่มตัวอย่างทำให้การวิจัยสามารถดำเนินการจนสำเร็จลุล่วง

ขอขอบคุณพี่ๆ และเพื่อน ๆ สาขาวิชาวิศวกรรมชีวการแพทย์ รวมถึงบุคคลอีกหลายท่านที่ไม่ได้กล่าวนามไว้ ณ ที่นี้ที่ได้ให้ความช่วยเหลือและเป็นกำลังใจให้กับผู้วิจัยมาโดยตลอด

สุดท้ายนี้ ผู้วิจัยขอน้อมรำลึกถึงคุณของบิดามารดาและครูอาจารย์ ที่อบรมสั่งสอนให้ความรู้เป็นกำลังใจและให้การสนับสนุนผู้วิจัยด้วยดีตลอดมา

คุณุตม์ ปิงศิริเจริญ

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย .....	ง
บทคัดย่อภาษาอังกฤษ .....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ .....	ช
สารบัญตาราง.....	ญ
สารบัญรูปภาพ .....	ฎ
บทที่ 1 บทนำ.....	1
1.1 เหตุผลและความสำคัญของปัญหา .....	1
1.2 ความมุ่งหมายของงานวิจัย.....	2
1.3 ขอบเขตของการวิจัย .....	2
บทที่ 2 งานวิจัยและทฤษฎีที่เกี่ยวข้อง .....	4
2.1 งานวิจัยที่เกี่ยวข้อง .....	4
2.2 ความรู้พื้นฐานเกี่ยวกับการสูญเสียการได้ยิน.....	6
2.2.1 หลักการการได้ยิน.....	6
2.2.2 อาการของการสูญเสียการได้ยิน.....	6
2.2.3 สาเหตุของการสูญเสียการได้ยิน.....	7
2.2.4 ประเภทของการสูญเสียการได้ยิน.....	7
2.3 เสียงพื้นฐานที่ใช้ทดสอบการได้ยิน.....	8
2.3.1 ความหมายของ Ling Six Sounds .....	8
2.3.2 เสียงของ Ling Six Sounds .....	8
2.3.3 คลื่นเสียง Ling Six Sounds .....	10



3.5.2 Artificial Neural Network (ANN) .....	36
3.5.3 Neural Network (NN) .....	37
3.5.4 K-Nearest Neighbors (KNN) .....	37
3.6 การวัดประสิทธิภาพ (Performance test).....	38
3.6.1 เมทริกซ์ความสับสน (Confusion Matrix) .....	38
บทที่ 4 ผลการดำเนินงานวิจัย .....	41
4.1 ผลลัพธ์ของฐานข้อมูลเสียงพูด (Speech database) .....	41
4.2 ผลลัพธ์ของการจัดเตรียมข้อมูล (Preprocessing) .....	43
4.2.1 ผลลัพธ์ของการลบเสียงรบกวน และการปรับสภาพสัญญาณ (Noise reduction and Signal condition) .....	43
4.2.2 ผลลัพธ์การหาลักษณะสำคัญของเสียง (Feature extraction) .....	44
4.3 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ CNN.....	46
4.4 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ ANN .....	66
4.5 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ Neural Network .....	86
4.6 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ KNN .....	106
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ .....	126
5.1 สรุปผลการวิจัย.....	126
5.2 อภิปรายผลการวิจัย .....	128
5.3 ข้อเสนอแนะ .....	129
บรรณานุกรม .....	132
ประวัติผู้เขียน.....	137

## สารบัญตาราง

	หน้า
ตาราง 1 คำศัพท์ที่ใช้ในงานวิจัย.....	24
ตาราง 2 การแบ่งข้อมูล .....	33
ตาราง 3 CNN layers.....	34
ตาราง 4 Confusion Matrix .....	39
ตาราง 5 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล .....	46
ตาราง 6 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล .....	48
ตาราง 7 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล .....	50
ตาราง 8 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล .....	52
ตาราง 9 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล .....	55
ตาราง 10 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล .....	58
ตาราง 11 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล .....	61
ตาราง 12 ผลลัพธ์ Accuracy ของวิธีการ CNN กับฐานข้อมูลทั้งหมด .....	65
ตาราง 13 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล.....	66
ตาราง 14 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล.....	67
ตาราง 15 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล.....	70
ตาราง 16 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล.....	72
ตาราง 17 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล.....	75
ตาราง 18 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล.....	78
ตาราง 19 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล.....	81
ตาราง 20 ผลลัพธ์ Accuracy ของวิธีการ ANN กับฐานข้อมูลทั้งหมด.....	85
ตาราง 21 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล .....	86

ตาราง 22 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล .....	87
ตาราง 23 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล .....	90
ตาราง 24 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล .....	92
ตาราง 25 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล .....	95
ตาราง 26 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล .....	98
ตาราง 27 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล .....	101
ตาราง 28 ผลลัพธ์ Accuracy ของวิธีการ Neural Network กับฐานข้อมูลทั้งหมด .....	105
ตาราง 29 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล.....	106
ตาราง 30 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล.....	107
ตาราง 31 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล.....	110
ตาราง 32 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล.....	112
ตาราง 33 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล.....	115
ตาราง 34 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล.....	118
ตาราง 35 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล.....	121
ตาราง 36 ผลลัพธ์ Accuracy ของวิธีการ KNN กับฐานข้อมูลทั้งหมด.....	125
ตาราง 37 สรุปผล Accuracy จากการสอนและทดสอบโมเดล.....	126
ตาราง 38 ตารางเปรียบเทียบค่าความถูกต้อง (Accuracy) ระหว่างงานวิจัยนี้กับงานวิจัยที่เกี่ยวข้องกับระบบรู้จำเสียงภาษาไทย .....	128

## สารบัญรูปภาพ

	หน้า
ภาพประกอบ 1 ประเภทของการสูญเสียการได้ยิน.....	8
ภาพประกอบ 2 ช่วงความถี่ของ Ling Six Sounds .....	9
ภาพประกอบ 3 คลื่นเสียงของเสียง “ ah ” (17) .....	10
ภาพประกอบ 4 คลื่นเสียงของเสียง “ ee ” (17) .....	10
ภาพประกอบ 5 คลื่นเสียงของเสียง “ mm ” (17) .....	11
ภาพประกอบ 6 คลื่นเสียงของเสียง “ oo ” (17) .....	11
ภาพประกอบ 7 คลื่นเสียงของเสียง “ sh ” (17) .....	11
ภาพประกอบ 8 คลื่นเสียงของเสียง “ ss ” (17) .....	12
ภาพประกอบ 9 ขั้นตอนการสกัดเวกเตอร์ลักษณะเฉพาะของ MFCC .....	14
ภาพประกอบ 10 Mel Filter bank .....	15
ภาพประกอบ 11 สถาปัตยกรรม CNN.....	16
ภาพประกอบ 12 เมทริกซ์ภาพคูณกับเมทริกซ์ตัวกรองหรือ kernel .....	16
ภาพประกอบ 13 Max pooling .....	17
ภาพประกอบ 14 Flatten layer .....	17
ภาพประกอบ 15 หลัจาก pooling layer, flattened as FC layer .....	18
ภาพประกอบ 16 เซลล์ประสาท (Neuron) .....	18
ภาพประกอบ 17 ตัวอย่างโครงสร้างของ Neural Network.....	19
ภาพประกอบ 18 ตัวอย่างโครงสร้างของ Neural Network และ Deep Learning .....	20
ภาพประกอบ 19 ทิศทางการไหลของข้อมูลใน Feedforward ANN .....	20
ภาพประกอบ 20 ทิศทางการไหลของข้อมูลใน Feedback ANN .....	21
ภาพประกอบ 21 ลักษณะการทำงานของ KNN.....	22

ภาพประกอบ 22 การทำงานของระบบรู้จำเสียงพูด .....	23
ภาพประกอบ 23 ขั้นตอนการเก็บข้อมูลเสียงจากกลุ่มตัวอย่าง .....	28
ภาพประกอบ 24 ตัวอย่างรูปคลื่นของสัญญาณเสียง (ก) มีเสียงรบกวน (ข) ไม่มีเสียงรบกวน....	29
ภาพประกอบ 25 ตัวอย่างรูปคลื่นของเสียง “ ไก่ ” .....	29
ภาพประกอบ 26 ขั้นตอนการหาค่าลักษณะสำคัญของ MFCC.....	30
ภาพประกอบ 27 การแบ่งสัญญาณเสียง .....	31
ภาพประกอบ 28 ตัวอย่าง MFCC ในรูปแบบไฟล์ *.csv.....	31
ภาพประกอบ 29 ผลลัพธ์จากการเรียนรู้ของโมเดลโดยใช้ฐานข้อมูลภาพ 3D spectrogram.....	32
ภาพประกอบ 30 ผลลัพธ์จากการเรียนรู้ของโมเดลโดยใช้ฐานข้อมูลภาพ 2D spectrogram.....	32
ภาพประกอบ 31 ตัวอย่าง MFCC ของเสียง “ไก่” ในรูปแบบ 3D และ 2D spectrogram .....	32
ภาพประกอบ 32 CNN architecture.....	34
ภาพประกอบ 33 ReLU function.....	35
ภาพประกอบ 34 ANN architecture .....	36
ภาพประกอบ 35 Neural Network architecture .....	37
ภาพประกอบ 36 KNN architecture.....	38
ภาพประกอบ 37 ตัวอย่างรูปคลื่นของสัญญาณเสียงจากฐานข้อมูลเสียงพูด (ก) กลุ่มตัวอย่างคนที่ 1 (ข) กลุ่มตัวอย่างคนที่ 2 (ค) กลุ่มตัวอย่างคนที่ 3 (ง) กลุ่มตัวอย่างคนที่ 4.....	43
ภาพประกอบ 38 ตัวอย่างรูปคลื่นของสัญญาณเสียงที่ลบเสียงรบกวนและตัดข้อมูล (ก) เสียง “อา” (ข) เสียง “แอปเปิ้ล” (ค) เสียง “แซน” (ง) เสียง “กล้วย” (จ) เสียง “นก” (ฉ) เสียง “ดำ” .....	44
ภาพประกอบ 39 ตัวอย่างฐานข้อมูลของ MFCC ในรูปแบบไฟล์ *.csv .....	44
ภาพประกอบ 40 ตัวอย่างฐานข้อมูลของ MFCC ในรูปแบบไฟล์ *.png (ก) เสียง “อา” (ข) เสียง “แอปเปิ้ล” (ค) เสียง “แซน” (ง) เสียง “กล้วย” (จ) เสียง “นก” (ฉ) เสียง “ดำ” .....	46
ภาพประกอบ 41 Confusion Matrix ของ CNN 10 กลุ่มข้อมูล.....	47
ภาพประกอบ 42 Confusion Matrix ของ CNN 20 กลุ่มข้อมูล.....	49

ภาพประกอบ 43 Confusion Matrix ของ CNN 30 กลุ่มข้อมูล.....	51
ภาพประกอบ 44 Confusion Matrix ของ CNN 40 กลุ่มข้อมูล.....	54
ภาพประกอบ 45 Confusion Matrix ของ CNN 50 กลุ่มข้อมูล.....	57
ภาพประกอบ 46 Confusion Matrix ของ CNN 60 กลุ่มข้อมูล.....	60
ภาพประกอบ 47 Confusion Matrix ของ CNN 66 กลุ่มข้อมูล.....	64
ภาพประกอบ 48 Confusion Matrix ของ ANN 10 กลุ่มข้อมูล.....	67
ภาพประกอบ 49 Confusion Matrix ของ ANN 20 กลุ่มข้อมูล.....	69
ภาพประกอบ 50 Confusion Matrix ของ ANN 30 กลุ่มข้อมูล.....	71
ภาพประกอบ 51 Confusion Matrix ของ ANN 40 กลุ่มข้อมูล.....	74
ภาพประกอบ 52 Confusion Matrix ของ ANN 50 กลุ่มข้อมูล.....	77
ภาพประกอบ 53 Confusion Matrix ของ ANN 60 กลุ่มข้อมูล.....	80
ภาพประกอบ 54 Confusion Matrix ของ ANN 66 กลุ่มข้อมูล.....	84
ภาพประกอบ 55 Confusion Matrix ของ Neural Network 10 กลุ่มข้อมูล.....	87
ภาพประกอบ 56 Confusion Matrix ของ Neural Network 20 กลุ่มข้อมูล.....	89
ภาพประกอบ 57 Confusion Matrix ของ Neural Network 30 กลุ่มข้อมูล.....	91
ภาพประกอบ 58 Confusion Matrix ของ Neural Network 40 กลุ่มข้อมูล.....	94
ภาพประกอบ 59 Confusion Matrix ของ Neural Network 50 กลุ่มข้อมูล.....	97
ภาพประกอบ 60 Confusion Matrix ของ Neural Network 60 กลุ่มข้อมูล.....	100
ภาพประกอบ 61 Confusion Matrix ของ Neural Network 66 กลุ่มข้อมูล.....	104
ภาพประกอบ 62 Confusion Matrix ของ KNN 10 กลุ่มข้อมูล.....	107
ภาพประกอบ 63 Confusion Matrix ของ KNN 20 กลุ่มข้อมูล.....	109
ภาพประกอบ 64 Confusion Matrix ของ KNN 30 กลุ่มข้อมูล.....	111
ภาพประกอบ 65 Confusion Matrix ของ KNN 40 กลุ่มข้อมูล.....	114

ภาพประกอบ 66 Confusion Matrix ของ KNN 50 กลุ่มข้อมูล ..... 117

ภาพประกอบ 67 Confusion Matrix ของ KNN 60 กลุ่มข้อมูล ..... 120

ภาพประกอบ 68 Confusion Matrix ของ KNN 66 กลุ่มข้อมูล ..... 124



## บทที่ 1

### บทนำ

#### 1.1 เหตุผลและความสำคัญของปัญหา

จากข้อมูลขององค์การอนามัยโลก ในปี 2018 (1) พบว่ามีประชากรโลกที่มีปัญหาทางการได้ยินทั้งหมด 466 ล้านคนทั่วโลก คิดเป็นร้อยละ 6.1 ของประชากรโลก โดย 432 ล้านคนหรือร้อยละ 93 พบในผู้ใหญ่ และ 34 ล้านคนหรือร้อยละ 7 พบในเด็ก และยังพบว่า ในเด็กแรกเกิดทุก 1,000 คน จะพบภาวะสูญเสียการได้ยินจำนวน 1-3 คน (2, 3) จากสถิติในประเทศไทยโดยองค์การอนามัยโลก ปี 2007 (4) ได้รายงานว่ามีจำนวนประชากรที่มีปัญหา การได้ยินทั้งหมด 8,848,091 คน จากทั้งหมด 66,527,000 คน คิดเป็นร้อยละ 13.3 ของประชากร และจากข้อมูลของสำนักงานสถิติแห่งชาติ กระทรวง เทคโนโลยีสารสนเทศและการสื่อสารในปี 2017 (5) ได้ทำการสำรวจผู้ที่ได้รับการขึ้นทะเบียนเป็นผู้พิการทางการได้ยิน จำแนกตามลักษณะความบกพร่อง พบว่ามีผู้ที่มีปัญหาการได้ยินทั้งหมดจำนวน 107,697 คน

โดยประเภทของการสูญเสียการได้ยินสามารถแบ่งออกได้ 2 ประเภท คือ 1. ผู้ที่มีปัญหาสูญเสียการได้ยินเกิดขึ้นโดยที่ยังไม่เคยมีภาษาพูดมาก่อน (Pre-lingual Deafness) หมายถึงผู้ที่มี การสูญเสียการได้ยินที่เป็นมาตั้งแต่เกิดหรือเกิดก่อนระยะที่จะเรียนรู้ภาษา คืออายุน้อยกว่า 5 ปี โดยในกลุ่มนี้เมื่อมีปัญหาไม่ได้ยินเสียงในช่วงที่จำเป็นในการเรียนรู้ภาษาก็จะส่งผลกระทบต่อ พัฒนาการการใช้ภาษา ทำให้เกิด ความพิการทางการพูดตามมา เรียกว่าหูหนวกและเป็นใบ้ โดยผู้ป่วยกลุ่มนี้ถ้าได้รับการรักษาฝึกระยะ 5 ปี พบว่าจะมีพัฒนาการทาง ภาษา ได้ใกล้เคียงคนปกติ (6) แต่ถ้าผู้ป่วยที่สูญเสียการได้ยินในกลุ่มนี้ไม่ได้รับการรักษา แก้ไขจน เลยช่วงที่สมองจะมีการพัฒนาด้านภาษาแล้วหรืออายุมากกว่า 5 ปี ถึงแม้จะได้รับการผ่าตัดฝึกระยะ หูเทียมในภายหลังก็จะไม่สามารถพูดสื่อสาร ได้เหมือนคนปกติ เนื่องจากเลยระยะเวลาที่ สมอง จะพัฒนาด้านการเรียนรู้ภาษาไปแล้ว ถึงแม้ว่าการ ผ่าตัดจะทำให้ผู้ป่วยได้ยินเสียงใน สิ่งแวดล้อมได้ก็ตาม 2. กลุ่มผู้ที่มีปัญหาการได้ยินที่เกิดหลังจากที่มี ภาษาพูดมาก่อนแล้ว (Post-lingual Deafness) โดยในกลุ่มนี้พบในเด็กหรือผู้ใหญ่ที่การสูญเสียการ ได้ยินนั้นเกิดขึ้นภายหลัง จากที่มีภาษาพูดมาแล้ว ผู้ป่วยกลุ่มนี้เมื่อได้รับการผ่าตัดฝึกระยะ หูเทียมแล้ว จะสามารถ กลับมาได้ยิน พูดสื่อสารได้เหมือนปกติ (6)

ดังนั้น ผู้ที่ได้รับการผ่าตัดฝึกระยะ หูเทียม หลังการผ่าตัดจะมีปัญหาด้านการฟังและการพูดเกิดขึ้น ถ้าได้รับการฝึกฝนทักษะการฟังและการสื่อสารอย่างสม่ำเสมอและถูกต้องจากนัก แก้ไขการพูด จะทำให้สมองส่วนที่ทำหน้าที่รับและแปลงสัญญาณเสียงได้เกิดการปรับตัว หลังจาก

นั้นผู้ป่วยควรทำการฝึกฝนด้วยตนเองที่บ้าน เพื่อให้ได้ผลลัพธ์ที่ดีขึ้นก่อนเข้ารับการประเมินที่โรงพยาบาลในครั้งถัดไป แต่ปัญหาที่นักบำบัดการพูดของโรงพยาบาลพบคือ ผู้ป่วยส่วนใหญ่ไม่ได้ฝึกทักษะเพิ่มเติมด้วยตนเอง ทำให้พัฒนาการเป็นไปได้ช้า นักบำบัดการพูดจึงต้องการตัวช่วยเสริมที่ให้ผู้ป่วยใช้ฝึกทักษะการพูดด้วยตนเองที่บ้านหรือสถานที่อื่นๆได้

ในงานวิจัยนี้จึงนำเสนอระบบรู้จำเสียงพูดสำหรับฟื้นฟูการสื่อสารของคนไข้ผ่าตัดประสาทหูเทียมทั้งสองข้างหรือระดับการสูญเสียการได้ยินมากกว่า 40 เดซิเบล โดยวิธีการหาลักษณะสำคัญของเสียงด้วยวิธี สัมประสิทธิ์เซปสตรัมของเมล (Mel-Frequency Cepstral Coefficient: MFCC) และใช้ฐานข้อมูลเสียง 7 แบบ ได้แก่ 10 กลุ่มคำ, 20 กลุ่มคำ, 30 กลุ่มคำ, 40 กลุ่มคำ, 50 กลุ่มคำ, 60 กลุ่มคำ และ 66 กลุ่มคำ โดยนำฐานข้อมูลดังกล่าวไปจำแนกคำด้วยวิธีการ 4 ชนิด ได้แก่ โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN) โดยโครงสร้างข้างในของ CNN ในงานวิจัยนี้มีต้นแบบมาจาก AlexNet (7), โครงข่ายประสาทเทียม (Artificial Neural Network: ANN), โครงข่ายประสาท (Neural Network: NN) และการเรียนรู้เพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbor: KNN) ในการสร้างโมเดลของระบบรู้จำเสียงพูดขึ้นมา เพื่อเปรียบเทียบค่าความถูกต้องของแต่ละวิธีการและฐานข้อมูล

## 1.2 ความมุ่งหมายของงานวิจัย

1.2.1 เพื่อพัฒนาระบบรู้จำเสียงสำหรับฟื้นฟูการสื่อสารของคนไทยที่ผ่าตัดประสาทหูเทียม

1.2.2 เพื่อเปรียบเทียบประสิทธิภาพของวิธีการจำแนกชนิดต่างๆ

## 1.3 ขอบเขตของการวิจัย

1.3.1 งานวิจัยนี้ใช้งานเฉพาะกับคำศัพท์ที่กำหนดไว้ได้แก่ Ling 6 Sound และคำศัพท์ที่ใช้ในชีวิตประจำวัน 60 คำ ซึ่งได้ปรึกษากับนักแก้ไขการพูดในการเลือกคำศัพท์

1.3.2 งานวิจัยนี้เก็บข้อมูลเสียงจากคนไทยที่มีอายุ 5 ถึง 50 ปี หรือผู้สูงอายุที่มีอายุ 60 ถึง 80 ปี และมีการผ่าตัดประสาทหูเทียมทั้งสองข้าง (การสูญเสียการได้ยินมากกว่า 40 dB)

1.3.3 งานวิจัยนี้ใช้วิธีการหาลักษณะสำคัญของเสียงด้วยวิธีการ สัมประสิทธิ์เซปสตรัมของเมล (Mel-Frequency Cepstral Coefficient: MFCC) และใช้วิธีการ โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN), โครงข่ายประสาทเทียม (Artificial Neural Network: ANN), โครงข่ายประสาท (Neural Network: NN) และการเรียนรู้เพื่อนบ้านที่ใกล้ที่สุด

(K-Nearest Neighbor: KNN) สำหรับการแยกแยะเสียงของหกเสียงของ Ling (Ling Six Sounds) และคำศัพท์ในชีวิตประจำวัน 60 คำ

1.3.4 งานวิจัยนี้ได้รับการรับรองจริยธรรมการวิจัยในมนุษย์ เลขที่ SWUEC-177/2562E และเก็บข้อมูลที่หน่วย ผู้ป่วยนอกแผนกโสต ศอ นาสิก ราลิ่งซ์ และ คลินิกฝึกพูด ณ โรงพยาบาล ศูนย์การแพทย์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี คณะแพทยศาสตร์ มศว



## บทที่ 2 งานวิจัยและทฤษฎีที่เกี่ยวข้อง

### 2.1 งานวิจัยที่เกี่ยวข้อง

ในปี พ.ศ. 2558 Srijiranon et al. (8) ได้ทำการเปรียบเทียบระบบ Neuro-Fuzzy กับ การ แจกแจงผสมแบบเกาส์เซียน (Gaussian Mixture Models: GMM), ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM), การเรียนรู้ต้นไม้ตัดสินใจ (Decision-tree) และ การจำแนก ของเบย์ (Byes) เพื่อตรวจสอบประสิทธิภาพในการจดจำเสียงของมนุษย์โดยใช้ Neuro-Fuzzy ใน งานวิจัยนี้ใช้ การประมาณพารามิเตอร์แบบอิงการรับฟังของมนุษย์ (Perceptual Linear Predictive: PLP) สำหรับการหาลักษณะสำคัญของเสียงและ ฐานข้อมูลมาจากการบันทึกคำศัพท์ภาษาไทย แก้วคำในหลายสภาพแวดล้อม ผลการวิจัยพบว่าความแม่นยำของ Neuro-Fuzzy นั้นสูงกว่าวิธีการ ยอดนิยมอื่น ๆ แต่ GMM และ SVM ก็ให้ความแม่นยำที่ค่อนข้างสูงเช่นกัน

ในปี พ.ศ. 2559 Kartiki Gupta และ Divya Gupta (9) ได้ทำการเปรียบเทียบการหา ลักษณะสำคัญของเสียงด้วยเทคนิคสัมประสิทธิ์การประมาณพารามิเตอร์เชิงเส้น (Linear Predictive Coding: LPC) , การกรองสเปกตรัมสัมพัทธ์ (Relative Spectral Filtering: RASTA) และ สัมประสิทธิ์เซปสตรัมของเมล (Mel-Frequency Cepstral Coefficient: MFCC) ในระบบรู้จำเสียง โดยสรุปข้อดีและข้อจำกัดของแต่ละวิธีการหาลักษณะสำคัญ และเปรียบเทียบผลการทดลองจาก งานวิจัยต่างๆ (10-12) ซึ่งผลลัพธ์สำหรับภาษาอังกฤษคือ MFCC feature ได้ค่าความถูกต้องที่ 99.9 เปอร์เซ็นต์, RASTA feature ได้ค่าความถูกต้องที่ 94.27 เปอร์เซ็นต์ และ LPC feature ได้ค่า ความถูกต้องที่ 91.4 เปอร์เซ็นต์ U. G. Patil et al. (13) , Muhammad Atif Imtiaz และ Gulistan Raja (14) นำเสนอระบบรู้จำเสียงสำหรับภาษา Hindi และ ภาษาอังกฤษ โดยใช้ MFCC feature และการเรียนรู้เพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbor: KNN) ในการแยกแยะเสียง โดย ฐานข้อมูลเก็บมาจากผู้หญิง 2 คน ผู้ชาย 4 คน รวมเป็น 44 ข้อมูล ตัวอย่างคำศัพท์ที่ใช้สำหรับเก็บ ข้อมูล ได้แก่ **आपना , जल , घेहू , दाल , हंडुतान** และ **सपना** เป็นต้น (13), คนพูด 5 คน เก็บจำนวน 200 ข้อมูล คำศัพท์ที่ใช้สำหรับเก็บข้อมูลได้แก่ Dark, Wash, Water, Year, Don't, Carry, Greasy, Like, Oily และ That (14) Yanmin Qian et al. (15) นำเสนอระบบรู้จำเสียงพูด โดยใช้ CNN และใส่ค่าขนาดของตัวกรอง, ชั้น pooling และ input feature maps ที่แตกต่างกัน แล้วนำมาเปรียบเทียบกับหน่วยความจำระยะสั้นระยะยาว (Long Short-Term Memory: LSTM) เพื่อสร้างโมเดลที่ทนทานต่อสิ่งรบกวนมากขึ้น

ในปี พ.ศ. 2560 Adiwijaya et al. (16) ได้นำเสนอการเปรียบเทียบ MFCC และ LPC feature โดยใช้กับ KNN ในการจำแนกการออกเสียงตัวอักษรภาษา Hijaiyah ผลที่ได้คือ ค่าความถูกต้องที่ใช้ LPC ได้ 92 เปอร์เซ็นต์ และ MFCC ได้ 87 เปอร์เซ็นต์ P. Sangsai และ P. Detyong (17) ได้ทำการพัฒนาแอปพลิเคชันบนระบบปฏิบัติการแอนดรอยด์ สำหรับช่วยฝึกฝนทักษะการฟังและการสื่อสารเบื้องต้นของผู้ผ่าตัดฝังประสาทหูเทียมได้ด้วยตนเอง ซึ่งแอปพลิเคชันนี้ใช้ระบบรู้จำเสียงในการจำแนกเสียง Ling Six Sounds โดยใช้เทคนิค MFCC สำหรับการหาลักษณะสำคัญของเสียงและ ใช้เครื่องมือ CMUSphinx สำหรับการจำแนกเสียง Ling Six Sounds ผลการทดสอบของระบบคือ เมื่อทดสอบในบริเวณที่มีเสียงรบกวน 57.3 เดซิเบล ได้ค่าความถูกต้องเฉลี่ยของทั้งระบบ 81.67 เปอร์เซ็นต์ และเมื่อทดสอบในบริเวณที่มีเสียงรบกวน 63.6 เดซิเบล ได้ค่าความถูกต้องเฉลี่ยของทั้งระบบ 55 เปอร์เซ็นต์

ในปี พ.ศ. 2561 Ericks Rachmat Swedia et al. (18) ได้นำเสนอระบบรู้จำเสียงภาษาอินโดนีเซียตัวเลข 0 ถึง 9 โดยใช้ LSTM และเปรียบเทียบค่าความถูกต้องของฐานข้อมูลที่ใช้วิธีการหาลักษณะสำคัญของเสียงที่แตกต่างกันได้แก่ LPC และ MFCC ใช้ข้อมูลฝึกจำนวน 7990 และข้อมูลทดสอบจำนวน 790 ข้อมูล P P Mini et al. (19) ได้ทำการศึกษาเปรียบเทียบ MFCC , การแปลงจริงที่แมปตามลำดับ (Sequency based Mapped Real Transform: SMRT) และ MFCC + SMRT features ในระบบรู้จำเสียงพูดโดยใช้ SVM ในการจำแนกเสียงและใช้ฐานข้อมูลของ The KARA ONE Database: Phonological Categories in imagined and articulated speech Zheli Liu et al. (20) นำเสนอการใช้ Convolutional Neural Network (CNN) ในการเพิ่มประสิทธิภาพระบบรู้จำเสียงที่ใช้ Gaussian mixture model (GMM) เป็นตัวจำแนกใน CNN โดย image input ขนาด 256x256 px และปรับ parameter เป็น In - 96Conv11 - 96Pool3 - 256Conv5 - 256Pool3 - 384Conv3 - 384Conv3 - 256Conv3 - 256Pool3 - 4096Fc - 1000Fc - Out. S. Tantisatirapong et al. (21) เสนอการเปรียบเทียบการหาลักษณะสำคัญของเสียงโดยใช้ MFCC, สเปกโตรแกรม (Spectrogram: SPT), พลังงานความหนาแน่นสเปกตรัม (Energy Spectral Density: ESD) และกำลังความหนาแน่นสเปกตรัม (Power Spectral Density: PSD) สำหรับระบบจดจำเสียงพูดภาษาไทยโดยมีสำเนียงจากภาคตะวันออกเฉียงเหนือ, ภาคใต้ และภาคกลาง ฐานข้อมูลมาจากการบันทึกตัวเลขศูนย์ถึงเก้าในภาษาไทยจากผู้พูดหญิง 30 คนและผู้พูดชาย 30 คน ผลการศึกษาพบว่าความแม่นยำของคุณลักษณะ MFCC สูงกว่า ESD, SPT และ PSD

ในปี พ.ศ. 2562 Ali Bou Nassif et al. (22) ได้ทำการศึกษาเรื่องการใช้ Deep Neural Networks กับระบบรู้จำเสียง โดยหาข้อมูลจากงานวิจัยตั้งแต่ปี 2006 ถึง 2018 รวมทั้งหมด 174

งานวิจัย ซึ่งอธิบายเกี่ยวกับ GMM base on Hidden Markov Model (HMM), Support Vector Machine (SVM) และ CNN สรุปได้ว่าการใช้ Deep learning มีประสิทธิภาพในระบบรู้จำเสียงมากกว่าการใช้ Machine learning

ในปี พ.ศ. 2564 K. Puengsiricharoen et al. (23) ได้ทำการศึกษาเรื่อง Ling Six Sounds นั้น มีความแตกต่างของแต่ละเสียงเพียงพหรือไม่ สำหรับการจำแนกเสียงของระบบรู้จำเสียง โดยใช้ MFCC สำหรับการหาลักษณะสำคัญของเสียง และใช้ไดนามิกไทม์วอร์ปิง (Dynamic Time Warping: DTW) สำหรับตรวจสอบความคล้ายคลึงกันของแต่ละเสียงใน Ling Six Sounds ผลการศึกษาพบว่าระยะทาง DTW ของคำที่เหมือนกันมีค่าน้อยกว่า 3000 และระยะทาง DTW ของคำที่ต่างกันมีค่ามากกว่า 4000 แสดงให้เห็นว่า Ling Six Sounds สามารถใช้ในการจำแนกเสียงของระบบรู้จำเสียงได้เนื่องจากมีความแตกต่างกันของแต่ละเสียงที่เพียงพอ

## 2.2 ความรู้พื้นฐานเกี่ยวกับการสูญเสียการได้ยิน

การสูญเสียการได้ยิน (hearing loss) หมายถึง ความบกพร่องในการรับฟังเสียงที่ต่างไปจากความสามารถตามปกติเมื่อเทียบกับมาตรฐานที่กำหนดไว้

### 2.2.1 หลักการการได้ยิน

การได้ยินของมนุษย์แบ่งเป็น 3 ส่วน คือ

2.1.1.1 การนำเสียง (Inductive function) ทำหน้าที่โดยหูชั้นนอก หูชั้นกลาง จะเริ่มตั้งแต่ใบหูเข้าไปจนถึงหูชั้นกลางบริเวณนอกต่อ oval window ออกมา

2.1.1.2 การรับฟังเสียง (sensorineural function) ทำหน้าที่โดยหูชั้นใน ได้แก่ ประสาทที่อยู่ภายในอวัยวะรูปก้นหอย (cochlear) ซึ่งอยู่ด้านในต่อ oval window เข้าไปและเส้นประสาทรับฟังเสียง (auditory nerve)

2.1.1.3 การแปลความหมาย (Inductive function) เป็นหน้าที่ของสมองส่วนกลางซึ่งจะรับและแปลความหมายของเสียงที่ได้ยิน

### 2.2.2 อาการของการสูญเสียการได้ยิน

การสูญเสียการได้ยินนั้นแตกต่างกันไปขึ้นอยู่กับชนิดของการสูญเสียการได้ยิน สาเหตุของการสูญเสียการได้ยินและระดับการสูญเสียโดยทั่วไปแล้ว คนที่มีปัญหาการได้ยินอาจพบปัญหาใดๆ ต่อไปนี้หรือทั้งหมด เช่น มีความเข้าใจน้อยในการสนทนาในชีวิตประจำวัน มีความสามารถในการได้ยินแต่ไม่เข้าใจ ต้องเพิ่มเสียงทีวีหรือวิทยุ ขอให้ผู้อื่นพูดบ่อยๆ หลีกเสียงสถานการณ์การเข้าร่วมสังคมที่เคยสนุก มีความยุ่งยากในการสื่อสารในสถานการณ์ที่มีเสียงดัง

เช่น ร้านอาหาร ริมถนน การสังสรรค์กับเพื่อนหรือในครอบครัวที่มีชีวิตชีวา ในรถยนต์ บนรถประจำทาง หรือในการประชุม มีอาการหูอื้อ หรือมีเสียงก้องและ/หรือเสียงพึมพำในหู

### 2.2.3 สาเหตุของการสูญเสียการได้ยิน

สาเหตุของการสูญเสียการได้ยินเป็นเรื่องสำคัญที่ต้องเข้าใจ เนื่องจากเป็นปัจจัยสำคัญในการพิจารณาการรักษาที่เหมาะสม การสูญเสียการได้ยินมีหลายสาเหตุ และบางสาเหตุทำให้เกิดการสูญเสียการได้ยินเฉพาะบางประเภท ตัวอย่างเช่น การอุดตันของขี้หู ทำให้เกิดการสูญเสียการได้ยินแบบชั่วคราว แต่ถ้าเกิดขึ้นบ่อยสามารถทำให้เกิดการสูญเสียการได้ยินแบบถาวรได้ การสูญเสียการได้ยินอาจเกิดจากสิ่งต่อไปนี้ อายุที่เพิ่มขึ้น (การสูญเสียการได้ยินที่เกี่ยวข้องกับอายุ เรียกว่า presbycusis) ยาบางชนิดซึ่งบางครั้งเรียกว่ายา "ototoxic" การบาดเจ็บหรือความเสียหายที่ศีรษะ ปัจจัยทางพันธุกรรม การได้รับเสียงดังมากเกินไป โรคที่ทำให้มีการสูญเสียการได้ยินที่พบได้บ่อย

### 2.2.4 ประเภทของการสูญเสียการได้ยิน

การสูญเสียการได้ยินแบ่งออกเป็น 3 ประเภทหลัก

#### 2.2.4.1 การสูญเสียการได้ยินชนิดประสาทรับฟังเสียงบกพร่อง

การสูญเสียการได้ยินชนิดประสาทรับฟังเสียงบกพร่อง (Sensorineural hearing loss: SNHL) เป็นประเภทที่พบบ่อยที่สุดของการสูญเสียการได้ยิน เกิดจากความผิดปกติที่หูชั้นใน (cochlea) หรือประสาทรับฟังเสียง (auditory nerve) ทำให้มีความลำบากในการรับฟังเสียง โดยเฉพาะเสียงสนทนา คือ ได้ยินเสียงแต่ฟังไม่รู้เรื่อง (24)

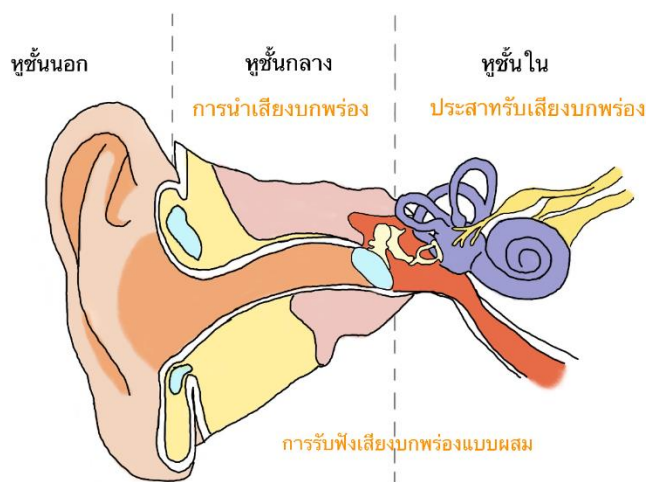
#### 2.2.4.2 การสูญเสียการได้ยินชนิดการนำเสียงบกพร่อง

การนำเสียงบกพร่อง (conductive hearing loss) เป็นภาวะการรับฟังเสียงบกพร่อง ซึ่งเป็นผลจากโรคที่ทำให้มีความผิดปกติที่หูชั้นนอกและหูชั้นกลางนอก oval window ออกมา เป็นผลให้มีความผิดปกติของกลไกการส่งผ่านคลื่นเสียงไปสู่หูชั้นใน โรคเหล่านี้สามารถแก้ไขให้หายได้ด้วยการรักษาทางยาหรือการผ่าตัด (24)

#### 2.2.4.3 สูญเสียการได้ยินชนิดการรับฟังเสียงแบบผสม

การรับฟังเสียงบกพร่องแบบผสม (Mixed hearing loss) เป็นภาวะที่เกิดจากความผิดปกติในการนำเสียงร่วมกับประสาทรับฟังเสียงบกพร่อง พบในโรคที่มีความผิดปกติที่หูชั้นนอกและ/หรือหูชั้นกลาง ร่วมกับความผิดปกติของหูชั้นใน เช่น โรคหูน้ำหนวกเรื้อรังซึ่งลูกกลมเข้าไปในหูชั้นใน โรคในหูชั้นกลางของผู้สูงอายุซึ่งมีปัญหาประสาทรับฟังเสียงเสื่อมด้วย โรคที่มีกระดูกจับแน่นที่แผ่นขาของกระดูกโกลนและพยาธิสภาพในถึงชั้นในร่วมด้วย (cochlear otosclerosis) โรค

เหล่านี้การรักษาโดยผ่าตัดอาจทำให้การได้ยินดีขึ้นได้ระดับหนึ่ง คือ แก้ไขความผิดปกติเกี่ยวกับการนำเสียงได้แต่ไม่สามารถแก้ภาวะประสาทรับฟังเสียงบกพร่องได้ (24)



ภาพประกอบ 1 ประเภทของการสูญเสียการได้ยิน

## 2.3 เสียงพื้นฐานที่ใช้ทดสอบการได้ยิน

ในปี 2012 Susan Scollie และคณะ (25) ได้ทำการศึกษาเกี่ยวกับหกเสียงของ Ling (Ling Six Sounds) ซึ่งเป็นแบบทดสอบเสียง 6 เสียงของ Daniel Ling เป็นเสียงพื้นฐาน 6 เสียง คือ /อา/ /อู/ /อี/ /อิม/ /ซ.../ /ส.../ ซึ่งเป็นเสียงที่ครอบคลุมเสียงในภาษาพูดที่เด็กจะได้ยินซึ่งจะครอบคลุมช่วงความถี่จากความถี่ต่ำ (low frequency) ไปความถี่สูง (high frequency)

### 2.3.1 ความหมายของ Ling Six Sounds

Ling Six Sounds (26) เป็นเครื่องมือที่ง่ายต่อการใช้งาน ซึ่งจะประเมินว่าเด็ก ๆ จะได้ยินเสียงด้วยการผ่าตัดฝังประสาทหูเทียมและ/หรือเครื่องช่วยฟังได้ดีเพียงใด การทดสอบ คือ การตรวจสอบการรับฟังพฤติกรรมที่ใช้เสียงง่ายๆ เพื่อตรวจหาการเปลี่ยนแปลงคุณภาพของการได้ยินของเด็ก การทดสอบควรจะทำเสร็จสิ้นในช่วงเริ่มต้นของแต่ละวันในโรงเรียน หรือ ช่วงการบำบัด และเมื่อใดก็ตามที่คุณสงสัยว่าเด็กไม่ได้ยินเหมือนที่พวกเขาได้ยินปกติ

### 2.3.2 เสียงของ Ling Six Sounds

#### 2.3.2.1 เสียง “ oo ”

เสียง “ oo ” นี้ใช้เพื่อวัดว่าสามารถได้ยินเสียงที่ต่ำมาก ๆ รวมทั้งสระเสียงต่ำทุกเสียง เสียงนี้อาจจะช่วยในเรื่องการรับรู้อารมณ์ของเพลงด้วย

#### 2.3.2.2 เสียง “ mm ”

เสียง “ mm ” ใช้เพื่อวัดว่าสามารถได้ยินเสียงความถี่ต่ำ เช่น สระในทุกคำที่เราได้ยิน

### 2.3.2.3 เสียง “ ah ”

เสียง “ ah ” ใช้เพื่อวัดว่าสามารถได้ยินเสียงความถี่ระดับกลาง ผู้ที่ไม่สามารถรับรู้เสียงนี้ได้อย่างมั่นใจ พวกเขาอาจจะได้ยินคำในบทสนทนาที่ไม่แน่นอน

### 2.3.2.4 เสียง “ sh ”

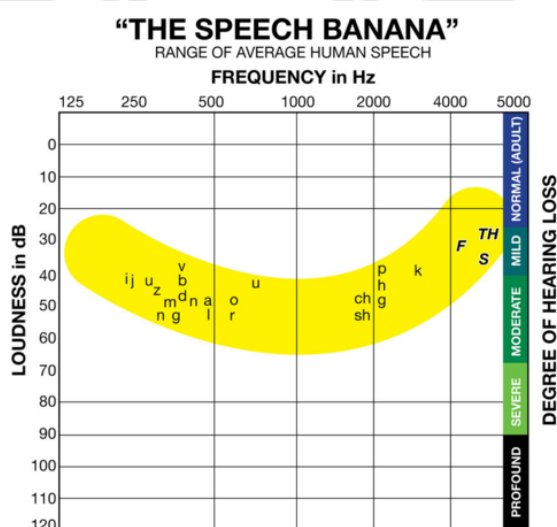
เสียง “ sh ” เป็นเสียงที่มีความถี่สูง ผู้ที่สูญเสียการได้ยินแบบรุนแรง อาจไม่สามารถได้ยินเสียงนี้ ถ้าไม่ได้ใส่เครื่องช่วยฟัง

### 2.3.2.5 เสียง “ ss ”

เสียง “ ss ” เป็นเสียงที่มีความถี่สูงมาก ๆ ผู้ที่สูญเสียการได้ยินแบบรุนแรง อาจไม่สามารถได้ยินเสียงนี้ ถ้าไม่ได้ใส่เครื่องช่วยฟัง เมื่อเราฟังการสนทนา เสียงคำพูดส่วนใหญ่มักอยู่ในช่วงเสียงสูง ดังนั้น ถ้าคนที่ไม่สามารถรับรู้เสียงนี้ได้ จะยากที่จะเข้าใจภาษาพูดและการสนทนากับผู้อื่น

### 2.3.2.6 เสียง “ ee ”

เสียง “ ee ” เป็นได้ทั้งเสียงสูงและเสียงต่ำของช่วงความถี่เสียงพูด ผู้ที่พูดตามเสียงนี้ได้จะสามารถบอกเราได้หลายอย่างว่าพวกเขาสามารถได้ยินเสียงอะไรบ้าง ถ้าพวกเขาออกเสียงตามได้อย่างถูกต้องแสดงว่าพวกเขาสามารถรับรู้ได้ทั้งความถี่สูงและต่ำได้ดี ถ้าพวกเขาพูดว่า “ss” เมื่อเลียนแบบเสียง แสดงว่าพวกเขามีปัญหาการได้ยินความถี่เสียงต่ำ ถ้าพวกเขาพูดว่า “oo” เมื่อเลียนแบบเสียง แสดงว่าพวกเขามีปัญหาการได้ยินความถี่เสียงสูง (26)



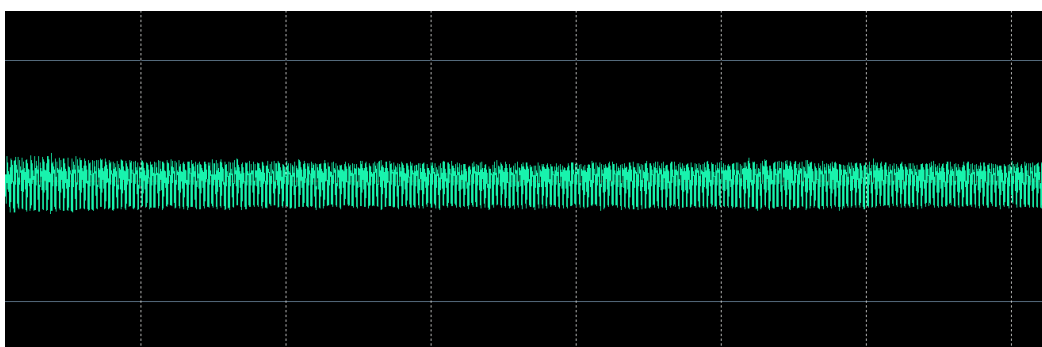
ภาพประกอบ 2 ช่วงความถี่ของ Ling Six Sounds

ที่มา: [https://commons.wikimedia.org/wiki/File:Speech\\_Banana\\_no\\_icons.png](https://commons.wikimedia.org/wiki/File:Speech_Banana_no_icons.png)

### 2.3.3 คลื่นเสียง Ling Six Sounds

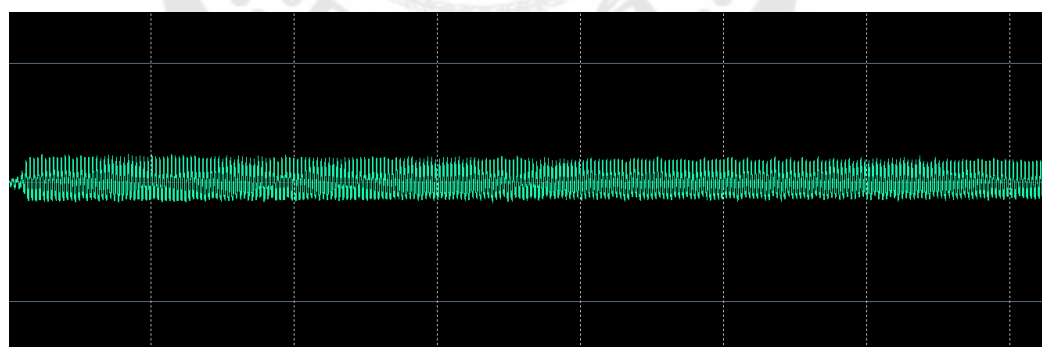
เสียง Ling Six Sounds ประกอบด้วยเสียงทั้งหมด 6 เสียง ได้แก่ เสียง “ah”, เสียง “ee”, เสียง “mm”, เสียง “oo”, เสียง “sh” และ เสียง “ss” โดยมีรูปคลื่นเสียงดังต่อไปนี้

#### 2.3.3.1 เสียง “ ah ”



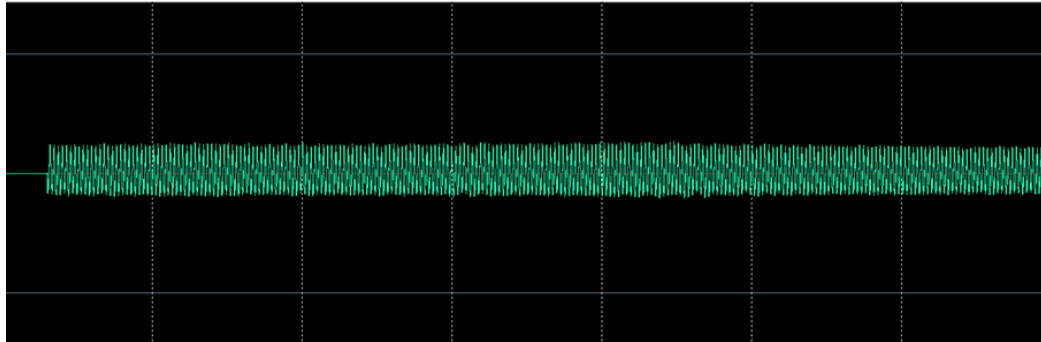
ภาพประกอบ 3 คลื่นเสียงของเสียง “ ah ” (17)

#### 2.3.3.2 เสียง “ ee ”



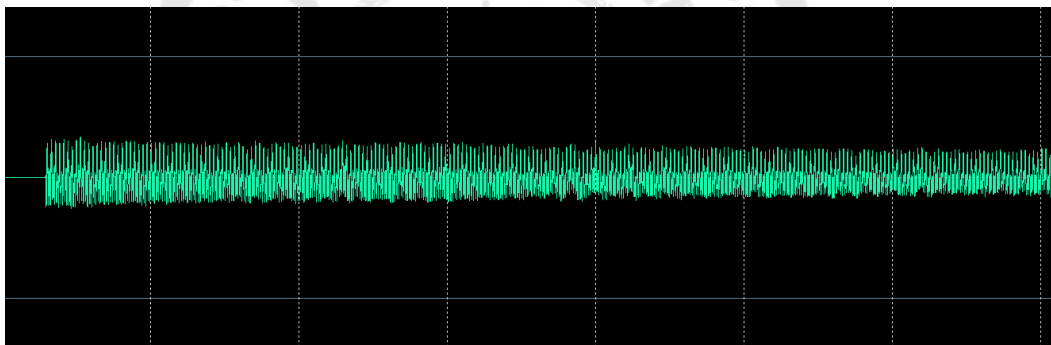
ภาพประกอบ 4 คลื่นเสียงของเสียง “ ee ” (17)

## 2.3.3.3 เสียง “ mm ”



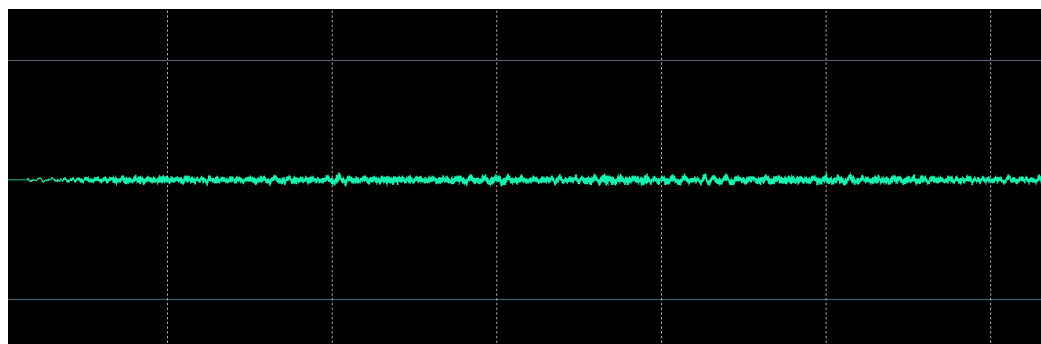
ภาพประกอบ 5 คลื่นเสียงของเสียง “ mm ” (17)

## 2.3.3.4 เสียง “ oo ”



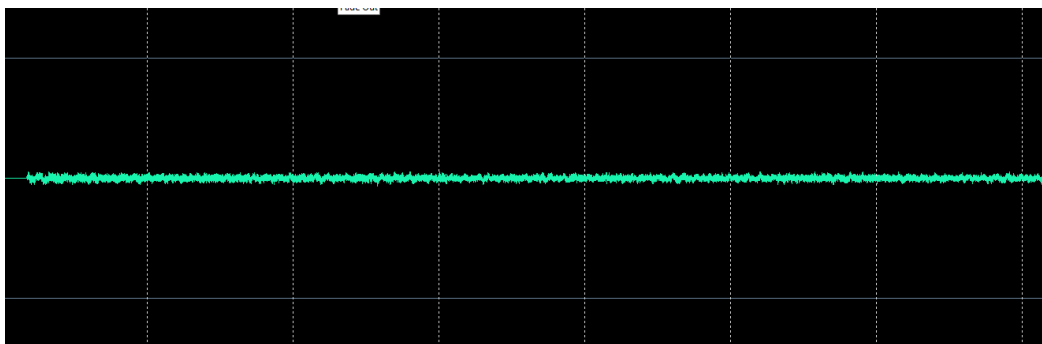
ภาพประกอบ 6 คลื่นเสียงของเสียง “ oo ” (17)

## 2.3.3.5 เสียง “ sh ”



ภาพประกอบ 7 คลื่นเสียงของเสียง “ sh ” (17)

### 2.3.3.6 เสียง “ ss ”



ภาพประกอบ 8 คลื่นเสียงของเสียง “ ss ” (17)

### 2.3.4 การทดสอบด้วย Ling Six Sounds

โดยปกติแล้วจะพูด 1 เสียงจากเสียงทั้ง 6 เสียงแบบสุ่ม ซึ่งสามารถตรวจสอบการได้ยินได้ 2 อย่างคือ อย่างแรกจะเป็นการตรวจสอบการได้ยินของแต่ละเสียง เนื่องจากแต่ละเสียงที่ได้ยินจะบอกถึงความสามารถในการได้ยินเสียงพูด อย่างที่สองจะเป็นการตรวจสอบการได้ยินของเครื่องช่วยฟังหรือตรวจสอบว่าความถูกต้องของกระบวนการทำงานของเสียง (27)

#### 2.3.4.1 วิธีการทดสอบ

- 1) ต้องทดสอบในสภาพแวดล้อมที่เงียบ
- 2) ระยะระหว่างผู้พูดและผู้ฟังประมาณ 50 เซนติเมตร ถึง 1 เมตร เหมือนกับการนั่งสนทนาทั่วไปตามความดังระดับเสียงพูด
- 3) เสียงที่พูดควรจะเหมือนกับเสียงพูดปกติไม่ยาวเกินไป
- 4) ถ้าผู้ฟังใส่เครื่องช่วยฟังหรือประสาทหูเทียมทั้ง 2 ข้าง ให้ทดสอบแต่ละข้างอย่างละครั้ง และทั้งสองข้างพร้อมกัน 1 ครั้ง
- 5) การทดสอบสำหรับเด็ก ผู้พูดต้องไม่ให้เด็กเห็นปากเวลาพูด ให้เด็กมองไปที่อื่นหรือหาอะไรมาฟังปากเวลาพูด ถ้าเด็กที่นั่งอยู่ในห้องทดสอบอาจทำให้ตัวสั่น สามารถหลีกเลี่ยงโดยการให้เด็กนั่งที่อื่นหรือนั่งตักแล้วให้คนอื่นพูด สำหรับผู้ใหญ่ สามารถให้พวกเขาหลับตา หรือมองไปทางอื่น

## 2.4 การแปลงฟูเรียร์อย่างรวดเร็ว (Fast Fourier Transform: FFT)

Fast Fourier Transform (FFT) (28) คือ วิธีการที่ช่วยให้สามารถคำนวณ discrete Fourier transform (DFT) ในเวลา  $O(n \log n)$  แนวคิดพื้นฐานของ FFT คือการใช้การแบ่งและการ

พหุนาม  $A(x)$  โดยแบ่งเวกเตอร์สัมประสิทธิ์ของพหุนามออกเป็นสองเวกเตอร์และคำนวณ DFT แบบวนซ้ำ สำหรับแต่ละเวกเตอร์และรวมผลลัพธ์เพื่อคำนวณ DFT ของพหุนามที่สมบูรณ์

ดังนั้นให้พหุนาม  $A(x)$  ที่มีดีกรี  $n-1$  โดยที่  $n$  คือยกกำลัง 2 และ  $n > 1$ :

$$A(x) = a_0 x^0 + a_1 x^1 + \dots + a_{n-1} x^{n-1} \quad (1)$$

โดยแบ่งออกเป็นสองพหุนามที่เล็กกว่าตัวที่มีค่าสัมประสิทธิ์ของตำแหน่งคู่เท่านั้นและอีกตัวที่มีค่าสัมประสิทธิ์ของตำแหน่งคี่:

$$A_0(x) = a_0 x^0 + a_2 x^1 + \dots + a_{n-2} x^{\frac{n}{2}-1} \quad (2)$$

$$A_1(x) = a_1 x^0 + a_3 x^1 + \dots + a_{n-1} x^{\frac{n}{2}-1} \quad (3)$$

สมการในรูปแบบง่ายคือ:

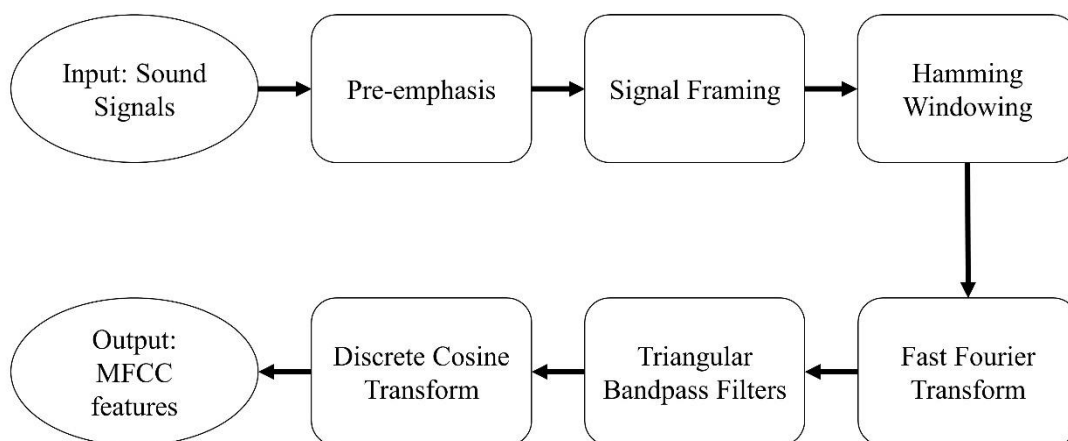
$$A(x) = A_0(x^2) + xA_1(x^2) \quad (4)$$

พหุนาม  $A_0$  และ  $A_1$  มีค่าสัมประสิทธิ์เพียงครึ่งเดียวของพหุนาม  $A$  หากเราสามารถคำนวณ DFT(A) ในเวลาเชิงเส้นโดยใช้ DFT(A<sub>0</sub>) และ DFT(A<sub>1</sub>) เราจะได้  $T_{DFT}(n) = 2T_{DFT}\left(\frac{n}{2}\right) + O(n)$  สำหรับความซับซ้อนของเวลาซึ่งส่งผลให้  $T_{DFT}(n) = O(n \log n)$  โดย master theorem

## 2.5 สัมประสิทธิ์เซปสตรัมของเมล (Mel-Frequency Cepstral Coefficient: MFCC)

MFCC คือค่าสัมประสิทธิ์ Cepstral ความถี่ของเมล (29) MFCC คำนึงถึงการรับรู้ของมนุษย์ในเรื่องความถี่ในความถี่ที่เหมาะสมโดยการแปลงความถี่ทั่วไปเป็นค่าของ Mel (Mel Scale) จึงเหมาะสำหรับงานการรู้จำเสียงได้ดี (เนื่องจากเหมาะสำหรับการทำความเข้าใจมนุษย์และความถี่ที่มนุษย์พูดหรือเปล่งเสียง).

เซปสตรัล (Cepstral) คือ การแปลงโคไซน์ไม่ต่อเนื่อง (Discrete Cosine Transform: DCT) ของลอการิทึมจากสเปกตรัมของสัญญาณแบบสั้น ค่าสัมประสิทธิ์ cepstral บน Mel Scale เป็นเทคนิคที่ได้รับการปรับปรุงของ cepstral โดยการปรับมาตราส่วนสเปกตรัมในระดับที่เหมาะสมกับการได้ยินของมนุษย์โดยพิจารณาจากลักษณะของสัญญาณเสียง สัญญาณเสียงความถี่ต่ำมีความสำคัญมากกว่าช่วงความถี่สูง ดังนั้น มาตรฐานสเปกตรัม (Spectrum scale) จึงได้รับการออกแบบมาเพื่อจับรายละเอียดของสัญญาณความถี่ต่ำการออกแบบนี้จึงเรียกว่า Mel Scale ขั้นตอนในการคำนวณค่าสัมประสิทธิ์ของเซลล์ใน Mel Scale มีดังต่อไปนี้:



ภาพประกอบ 9 ขั้นตอนการสกัดเวกเตอร์ลักษณะเฉพาะของ MFCC

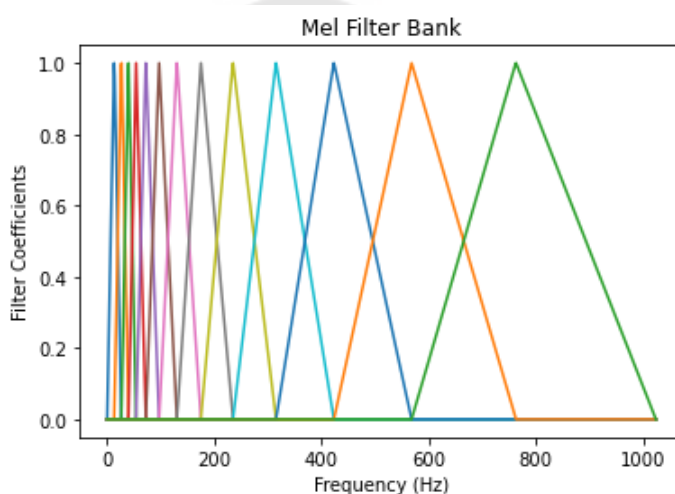
### 2.5.1 เมลสเปกตรัม (Mel spectrum)

Mel spectrum (29) คำนวณโดยการส่งสัญญาณที่แปลงฟูเรียร์ผ่านชุดตัวกรองแบนด์พาส (Band-pass) ที่เรียกว่าชุดตัวกรองความถี่เมล (Mel-filter bank) และเมล (Mel) เป็นหน่วยวัดตามความถี่ที่หูของมนุษย์ได้รับ ไม่สอดคล้องเชิงเส้นตรงกับความถี่ทางกายภาพของโทนเสียงเนื่องจากระบบการได้ยินของมนุษย์ไม่ได้รับระดับเสียงเชิงเส้น Mel scale ประมาณระยะห่างความถี่เชิงเส้นที่ต่ำกว่า 1000 Hz และระยะห่างลอการิทึมที่สูงกว่า 1000 Hz การวัดเมลจากความถี่ทางกายภาพสามารถแสดงเป็น

$$f_{Mel} = 2595 \log_{10} \left( 1 + \frac{f}{700} \right) \quad (5)$$

เมื่อ  $f$  หมายถึงความถี่ทางกายภาพเป็น Hz และ  $f_{Mel}$  หมายถึงความถี่ที่รับรู้

ชุดตัวกรองความถี่ (Filter banks) สามารถใช้ได้ทั้งในโดเมนเวลาและโดเมนความถี่ สำหรับการคำนวณ MFCC โดยทั่วไป Filter banks จะใช้ในโดเมนความถี่ โดยปกติความถี่กลางของตัวกรองจะมีระยะห่างเท่า ๆ กันบนแกนความถี่ อย่างไรก็ตามเพื่อเลียนแบบการรับรู้ของหูของมนุษย์แกนที่บิดเบี้ยวจะถูกใช้ตามการแสดงออกของฟังก์ชันไม่เชิงเส้นในสมการข้างต้น ตัวกรองฟิลเตอร์ที่ใช้กันมากที่สุดคือรูปสามเหลี่ยม Filter banks รูปสามเหลี่ยมที่มีการแปรปรวนความถี่ Mel แสดงในภาพประกอบ 10



ภาพประกอบ 10 Mel Filter bank

Mel spectrum ของสเปกตรัมขนาด  $X(k)$  คำนวณโดยการคูณสเปกตรัมขนาดโดยตัวกรองน้ำหนัก Mel รูปสามเหลี่ยมแต่ละตัว โดยแสดงเป็นสมการดังนี้

$$s(m) = \sum_{k=0}^{N-1} [|X(k)|^2 H_m(k)]; \quad 0 \leq m \leq M - 1 \quad (6)$$

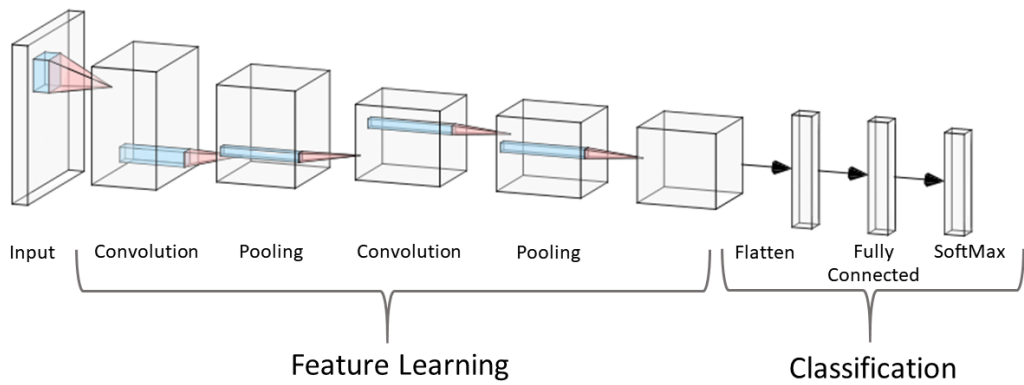
เมื่อ  $M$  คือจำนวนตัวกรองน้ำหนัก Mel รูปสามเหลี่ยมทั้งหมด,  $H_m(k)$  คือน้ำหนักที่กำหนดให้กับพลังงานสเปกตรัม  $k$  ที่มีส่วนกับแถบเอ๊าต์พุต  $m$

## 2.6 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN)

โครงข่ายประสาทแบบคอนโวลูชัน (CNN) (30) เป็นวิธีการการเรียนรู้เชิงลึกที่สามารถจำแนกข้อมูลด้วยภาพอินพุต, กำหนดความสำคัญ (น้ำหนักและ biases ที่เรียนรู้ได้) ให้กับวัตถุ

ต่างๆในภาพและสามารถแยกความแตกต่างจากวัตถุอื่น ๆ ได้ สำหรับ CNN การประมวลผลล่วงหน้าที่จะจำเป็นจะต่ำกว่าวิธีการการจำแนกประเภทอื่น ๆ มากเนื่องจาก CNN มีความสามารถในการเรียนรู้ตัวกรอง / คุณสมบัติเหล่านี้จากการฝึกอบรมข้อมูลอินพุต

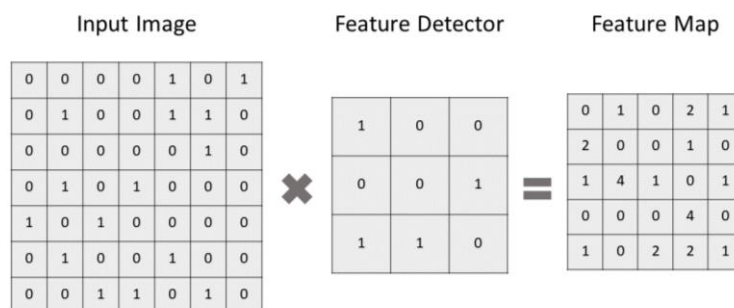
ในการฝึกและทดสอบโมเดล CNN ภาพอินพุตแต่ละภาพจะส่งผ่านชั้นคอนโวลูชัน (Convolution) ชุดหนึ่งด้วยฟิลเตอร์ (Kernels), ชั้น Pooling, ชั้น fully connected (FC) และใช้ฟังก์ชัน Softmax เพื่อจำแนกวัตถุที่มีค่าความน่าจะเป็นระหว่าง 0 ถึง 1



ภาพประกอบ 11 สถาปัตยกรรม CNN

### 2.6.1 ชั้นคอนโวลูชัน (Convolution Layer)

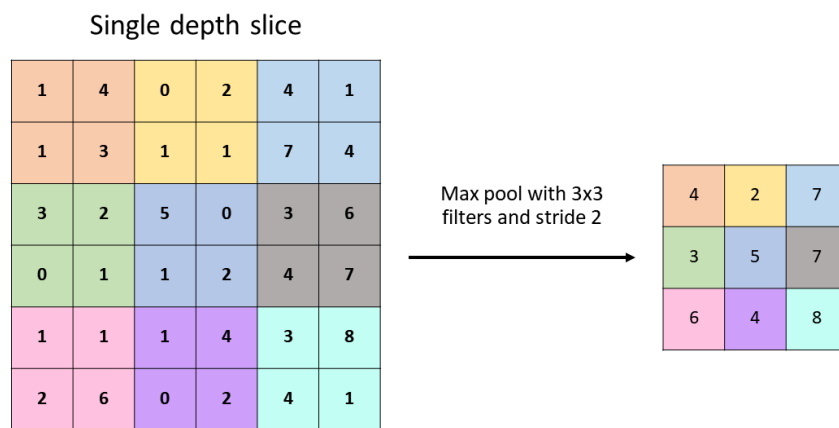
คอนโวลูชันเป็นชั้นแรกที่ดึงลักษณะสำคัญจากภาพอินพุต (30) คอนโวลูชันรักษาความสัมพันธ์ระหว่างพิกเซลโดยเรียนรู้ลักษณะสำคัญของภาพโดยใช้ข้อมูลอินพุตสี่เหลี่ยมเล็ก ๆ เป็นการดำเนินการทางคณิตศาสตร์ที่ใช้อินพุตสองอินพุตคือเมทริกซ์รูปภาพและตัวกรองหรือ kernel ผลลัพธ์ของ convolution ของเมทริกซ์ภาพคูณกับเมทริกซ์ตัวกรองเรียกว่า “ Feature Map”



ภาพประกอบ 12 เมทริกซ์ภาพคูณกับเมทริกซ์ตัวกรองหรือ kernel

## 2.6.2 ชั้นพูลลิ่ง (Pooling Layer)

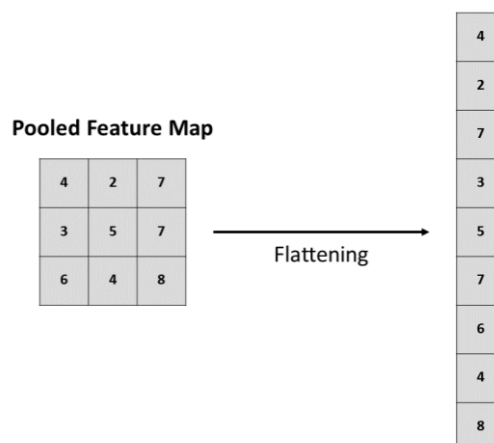
ในส่วน pooling layer จะลดจำนวนพารามิเตอร์เมื่อรูปภาพมีขนาดใหญ่เกินไป Spatial pooling เรียกอีกอย่างว่า subsampling หรือ down sampling ซึ่งจะลดขนาดของแต่ละ map แต่ยังคงรักษาข้อมูลที่สำคัญไว้ Spatial pooling สามารถแบ่งได้ 3 ประเภท ได้แก่ Max Pooling, Average Pooling และ Sum Pooling (30)



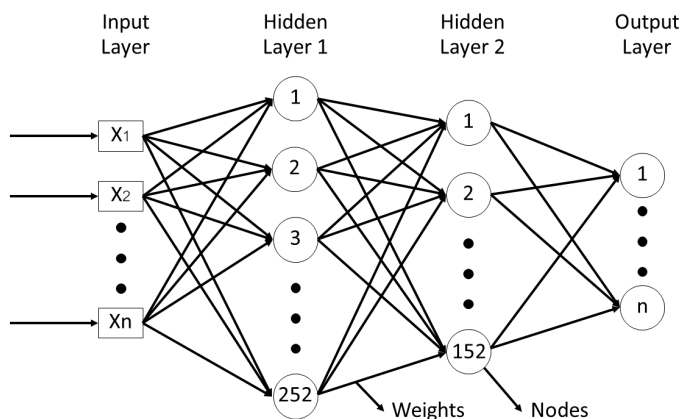
ภาพประกอบ 13 Max pooling

## 2.6.3 ชั้นเชื่อมต่ออย่างสมบูรณ์ (Fully Connected Layer: FC)

Fully Connected layer (FC) คือชั้นที่ทำให้เมทริกซ์ Pooling Feature Map แบนราบลงในรูปแบบเวกเตอร์และวางไว้ใน fully connected layer เหมือนกับเครือข่ายประสาทเทียม(30)



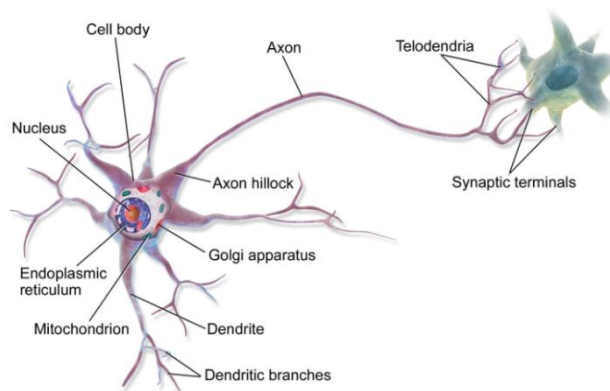
ภาพประกอบ 14 Flatten layer



ภาพประกอบ 15 หลักจาก pooling layer, flattened as FC layer

## 2.7 โครงข่ายประสาท (Neural Network: NN)

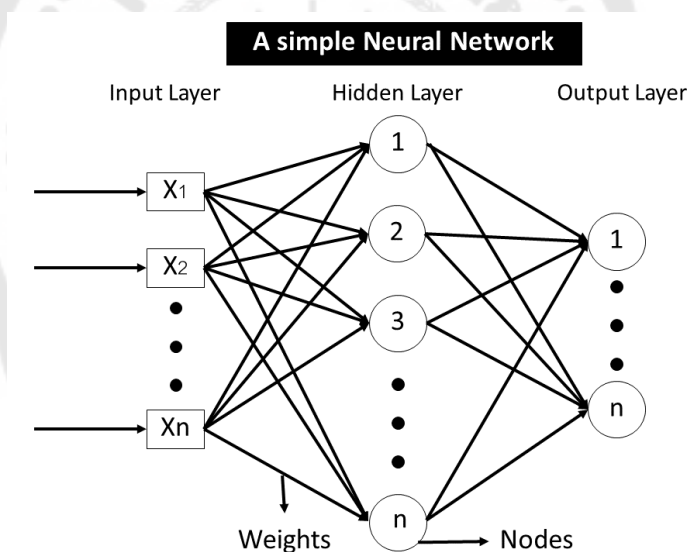
Neural Network (NN) (31) เป็นวิธีการที่จำลองให้เหมือนกับโครงข่ายประสาทในสมองของมนุษย์ ในการเรียนรู้และประมวลผลสิ่งต่างๆ ที่มนุษย์ได้รับรู้จากประสาทสัมผัสเช่น การมองเห็น การได้ยิน การได้กลิ่น การรับรส และการสัมผัส ซึ่งการเรียนรู้จดจำของสมองจะอยู่ในรูปแบบของตัวเลข เพราะฉะนั้นเราจึงต้องแปลงข้อมูลต่างๆ ในโลกเป็นตัวเลข เช่น ภาพ, เสียง และข้อความ เป็นต้น โดยโครงข่ายในสมองของมนุษย์จะมีเซลล์ประสาท (Neuron) มีหน้าที่รับข้อมูลมาจาก Dendrite และส่งข้อมูลไปยังเซลล์ประสาทตัวอื่นผ่านจุดประสานประสาท (Synapses) ตามภาพประกอบ 16 การเชื่อมต่อกันของเซลล์ประสาทหลายๆ ตัวนั้น จะทำให้เกิดเป็นเครือข่ายที่ทำงานร่วมกันขึ้น



ภาพประกอบ 16 เซลล์ประสาท (Neuron)

ที่มา : <https://commons.wikimedia.org/w/index.php?curid=28761830>

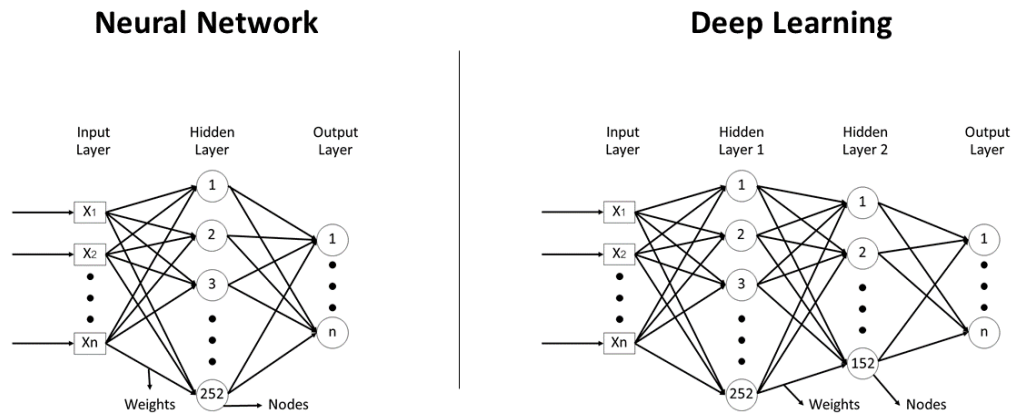
โครงสร้างของ Neural Network (32, 33) จะประกอบด้วย 3 ส่วน นั่นก็คือ 1. Input layer มี 1 ชั้น 2. Hidden layer มี 1 ชั้น หรือ มากกว่า 1 ชั้น (เรียกว่า Deep learning) 3. Output layer มี 1 ชั้น ซึ่งในแต่ละ layer จะมี Node เป็นหน่วยประมวลผลย่อย มีหน้าที่คือ ประมวลผลต่างๆ ใน Neural Network โดย Node เป็นการจำลองลักษณะการทำงานการส่งข้อมูลของเซลล์ประสาทในสมองมนุษย์ ซึ่งในแต่ละ Node จะมีจุดเชื่อมต่อกับจุดประสานประสาท (Synapses) ในสมอง ทำให้สามารถส่งข้อมูลไประหว่าง Node อื่นๆ ที่เชื่อมต่อกันได้ โดยกำหนดค่าน้ำหนัก (weight) กำกับอยู่ที่เส้นเชื่อมทุกเส้น ซึ่ง Node จะรวมชุดของค่าสัมประสิทธิ์ (coefficients) หรือ ค่าน้ำหนัก (weights) ของ Input จากข้อมูลที่ใส่เข้าไป เพื่อกำหนดความสำคัญให้กับ Input ที่เกี่ยวข้องกับงานที่วิธีการพยายามเรียนรู้ หลังจากนั้นค่า Input weights จะถูกรวมเข้าด้วยกันแล้ว ถูกส่งผ่าน Activation function เพื่อพิจารณาว่าสัญญาณนั้นควรจะก้าวหน้าไปมากน้อยเพียงใด ผ่านเครือข่ายเพื่อส่งผลต่อผลลัพธ์ที่สูงที่สุด



ภาพประกอบ 17 ตัวอย่างโครงสร้างของ Neural Network

## 2.8 โครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

Artificial Neural Network (ANN) (34, 35) เป็นวิธีการแขนงหนึ่งของการเรียนรู้แบบเชิงลึก (Deep learning) ซึ่งโครงสร้างของวิธีการมีความเหมือนกับ Neural Network แต่มีความแตกต่างที่ Hidden layer โดยใน Deep learning หรือ Deep Neural Network จะมี Hidden layer มากกว่า 1 ชั้น เพราะฉะนั้นจะมีจำนวน Node ที่ข้อมูลต้องผ่านในกระบวนการเรียนรู้ของวิธีการมากกว่า Neural Network ทำให้ได้ผลลัพธ์ที่มีประสิทธิภาพที่มากกว่า

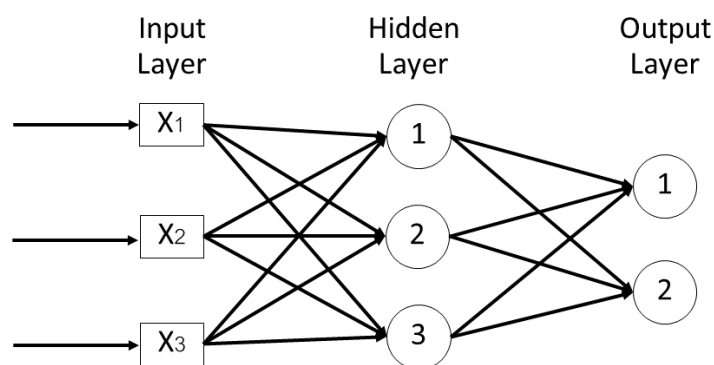


ภาพประกอบ 18 ตัวอย่างโครงสร้างของ Neural Network และ Deep Learning

ประเภทของ ANN หลักๆ สามารถแบ่งออกเป็น 2 ประเภท ได้แก่

### 2.8.1 ข่ายงานประสาทแบบป้อนไปหน้า (Feedforward ANN)

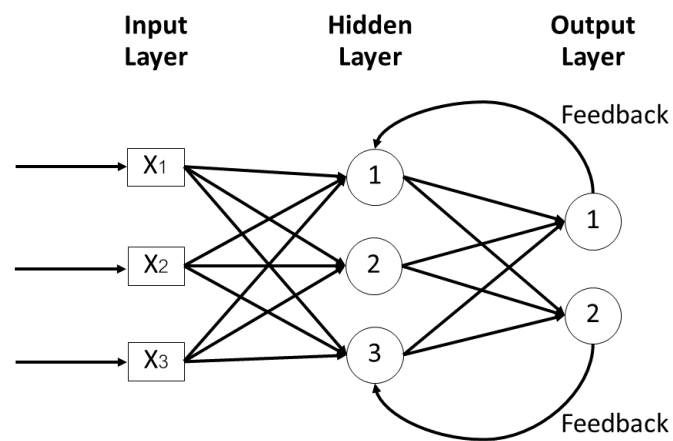
ใน Feedforward ANN (34, 35) การไหลของข้อมูลจะเกิดขึ้นในทิศทางเดียวเท่านั้น นั่นคือการไหลของข้อมูลมาจาก Input layer ไปยัง Hidden layer และสุดท้ายไปยัง Output layer ซึ่งจะไม่มีกระบวนการวน feedback loop โดยส่วนใหญ่ Feedforward ANN จะใช้กับงาน Supervised learning เช่น การจำแนกประเภท การจดจำรูปภาพ เป็นต้น



ภาพประกอบ 19 ทิศทางการไหลของข้อมูลใน Feedforward ANN

### 2.8.2 ข่ายงานประสาทแบบส่งผ่านย้อนกลับ (Feedback ANN)

ใน Feedback ANN (34, 35) จะมีการวน feedback loops ทำให้การไหลของข้อมูลสามารถเดินทางได้สองทิศทาง "สถานะ" ใน Feedback ANN จะเปลี่ยนแปลงไปเรื่อย ๆ จนกว่าจะถึงจุดสมดุล และจะคงอยู่ที่จุดสมดุลจนกว่า Input จะเปลี่ยนไปและต้องหาจุดสมดุลใหม่



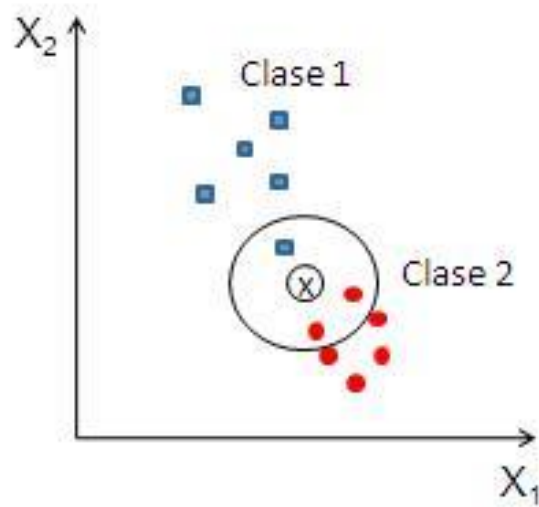
ภาพประกอบ 20 ทิศทางการไหลของข้อมูลใน Feedback ANN

### 2.9 การเรียนรู้เพื่อนบ้านที่ใกล้ที่สุด (K-Nearest Neighbor: KNN)

K-Nearest Neighbor (KNN) (36) เป็นวิธีการที่อยู่ในกลุ่ม Supervised learning ซึ่ง KNN ใช้ในการแบ่งกลุ่มข้อมูล โดยจะตัดสินใจว่าข้อมูลมีความคล้ายคลึงหรือใกล้เคียงกับกลุ่มข้อมูลใด โดยการตรวจสอบข้อมูลบางจำนวน (K) และจะตรวจสอบด้วยวิธีการ เพื่อนบ้านที่ใกล้ที่สุด (Nearest Neighbor) ซึ่งจะคำนวณระยะทางของข้อมูลว่าอยู่ใกล้กับกลุ่มข้อมูลไหนมากที่สุด โดยส่วนใหญ่จะใช้สมการหาระยะทางแบบ Euclidean ในการคำนวณสำหรับข้อมูลสำหรับฝึก และข้อมูลสำหรับทดสอบ สมการหาระยะทางแบบ Euclidean สามารถแสดงดังต่อไปนี้

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (7)$$

เมื่อ  $d(x, y)$  คือ ระยะสเกลาร์จากสองเวกเตอร์ของข้อมูล  $x$  และ  $y$ ,  $n$  คือ จำนวนของมิติข้อมูล

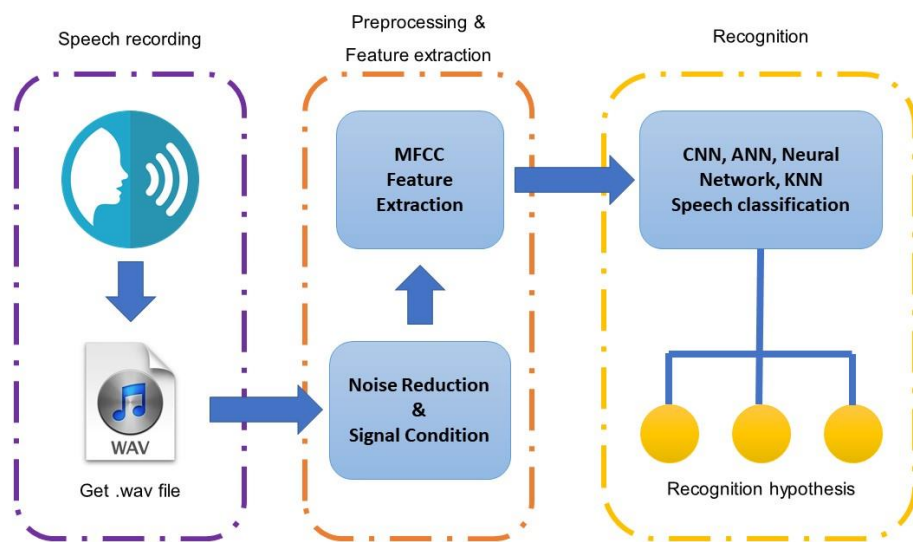


ภาพประกอบ 21 ลักษณะการทำงานของ KNN

ที่มา : [https://commons.wikimedia.org/wiki/File:KNN\\_detec.JPG](https://commons.wikimedia.org/wiki/File:KNN_detec.JPG)

### บทที่ 3 วิธีดำเนินการวิจัย

งานวิจัยนี้เป็นการวิจัยและพัฒนาระบบรู้จำเสียงสำหรับผู้พิการทางการได้ยิน โดยระบบนี้ประกอบไปด้วย 1) การเก็บข้อมูลเสียงจากผู้ป่วยที่ผ่าตัดประสาทหูเทียมจากหูทั้งสองข้าง 2) การจัดเตรียมข้อมูล (Preprocessing) และการหาลักษณะสำคัญของเสียง (Feature extraction) 3) การเรียนรู้ให้ระบบสามารถจดจำและแบ่งแยกคำศัพท์ได้ (Speech Recognition) โดยนำข้อมูลเสียงทั้งหมดเข้าวิธีการเรียนรู้แบบต่างๆ โดยการทำงานของระบบรู้จำเสียงสามารถแสดงได้ดังภาพประกอบ 22



ภาพประกอบ 22 การทำงานของระบบรู้จำเสียงพูด

มีขั้นตอนการดำเนินการระเบียบวิธีวิจัยดังนี้

1. ฐานข้อมูลเสียงพูด (Speech Database)
2. การบันทึกข้อมูลเสียง (Speech recording)
3. การจัดเตรียมข้อมูล (Preprocessing)
4. การจัดแบ่งข้อมูล (Data Segmentation)
5. การพัฒนาระบบรู้จำเสียง
6. การวัดประสิทธิภาพ (Performance test)

### 3.1 ฐานข้อมูลเสียงพูด (Speech database)

ฐานข้อมูลเสียงพูดนี้ได้มาจากการบันทึกเสียง Ling 6 sound และคำศัพท์ในชีวิตประจำวันจำนวน 60 คำ โดยเก็บข้อมูลจากผู้ชาย 10 คน และผู้หญิง 10 คน ที่มีการผ่าตัดประสาทหูเทียมทั้งสองข้าง โดยได้รับการรับรองจริยธรรมการวิจัยในมนุษย์ เลขที่ SWUEC-177/2562E และคำปรึกษาจากนักฝึกการพูดของศูนย์การแพทย์สมเด็จพระรัตนราชสุตาภยยามบรมราชกุมารีในด้านของคำศัพท์ในชีวิตประจำวัน และทำการบันทึกเสียง 3 ครั้งต่อหนึ่งคำ คำศัพท์ที่ใช้ในการบันทึกเสียงได้แก่

ตาราง 1 คำศัพท์ที่ใช้ในงานวิจัย

คำศัพท์	การออกเสียง
<b>หมวด Ling 6 Sounds</b>	
อา	อา
อี	อี
อู	อู
อิม	อิม
ช	ชู่
ช	ชู่
<b>หมวดสัตว์</b>	
เสื่อ	เสื่อ
ลิง	ลิง
สิงโต	สิง - โต
ไก่	ไก่
งู	งู
ช้าง	ช้าง
นก	นก
หมา	หมา

ตาราง 1 (ต่อ)

คำศัพท์	การออกเสียง
เปิด	เปิด
จระเข้	จอ - ระ - เค้
<b>หมวดสิ่งของ</b>	
พัดลม	พัด - ลม
กางเกง	กาง - เกง
เสื้อ	เซื่อ
ดินสอ	ดิน - สอ
โต๊ะ	โต๊ะ
หมวก	หมวก
แปรงสีฟัน	แปรง - สี - ฟัน
หวี	หฺวี
ปิ่น	ปิ่น
แก้ว	แก้ว
<b>หมวดคำกริยา</b>	
เดิน	เดิน
เด็ดตัว	เด็ด - ตัว
วิ่ง	วิ่ง
เขียน	เขียน
ดูทีวี	ดู - ที - วี
รดน้ำต้นไม้	รด - น้ำ - ต้น - ม้าย
ยีน	ยีน
กินข้าว	กิน - ข้าว
เตะฟุตบอล	เตะ - ฟุ่ด - บ็อล

ตาราง 1 (ต่อ)

คำศัพท์	การออกเสียง
อาบน้ำ	อาบ - น้ำ
<b>หมวดอวัยวะ</b>	
หู	หู
ตา	ตา
ปาก	ปาก
จมูก	จะ - หมูก
ลิ้น	ลิ้น
เท้า	เท้า
แขน	แขน
มือ	มือ
ข้อศอก	คอ - สอก
ฟัน	ฟัน
<b>หมวดอาหาร</b>	
กล้วย	กล้วย
ไข่	ไข่
แตงโม	แตง - โม
สับปะรด	สับ - ปะ - รด
ส้ม	ส้ม
มะม่วง	มะ - ม่วง
แอปเปิ้ล	แอป - เปิ้ล
มังคุด	มัง - คุด
ชมพู	ชม - พู่
มะละกอ	มะ - ละ - กอ

ตาราง 1 (ต่อ)

คำศัพท์	การออกเสียง
<b>หมวดสี่</b>	
แดง	แดง
ขาว	ขาว
ดำ	ดำ
ฟ้า	ฟ้า
เหลือง	เหลือง
เขียว	เขียว
ส้ม	ส้ม
ม่วง	ม่วง
น้ำเงิน	น้ำ - เงิน
น้ำตาล	น้ำ - ตาน

### 3.2 การบันทึกข้อมูลเสียง (Speech recording)

ข้อมูลเสียงที่บันทึกมาจากผู้ป่วยที่ผ่าตัดประสาทหูเทียมทั้งสองข้างจากหน่วย ผู้ป่วยนอกแผนกโสต ศอ นาสิก ราลิ่งซ์ และ คลินิกฝึกพูด ณ โรงพยาบาลศูนย์การแพทย์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี โดยข้อมูลเสียงที่บันทึกอยู่ในรูปแบบไฟล์ \*.wav โดยใช้ sampling rate ที่ 44.1 kHz และมีความละเอียดที่ 16 bit รูปแบบข้อมูลในการบันทึกจะเป็นแบบ 1 คน ต่อ 1 ไฟล์เสียง เนื่องจากสถานที่ในการเก็บข้อมูลและเวลาของกลุ่มตัวอย่าง จึงไม่สามารถเก็บข้อมูลในรูปแบบ 1 คำ ต่อ 1 ไฟล์เสียงได้



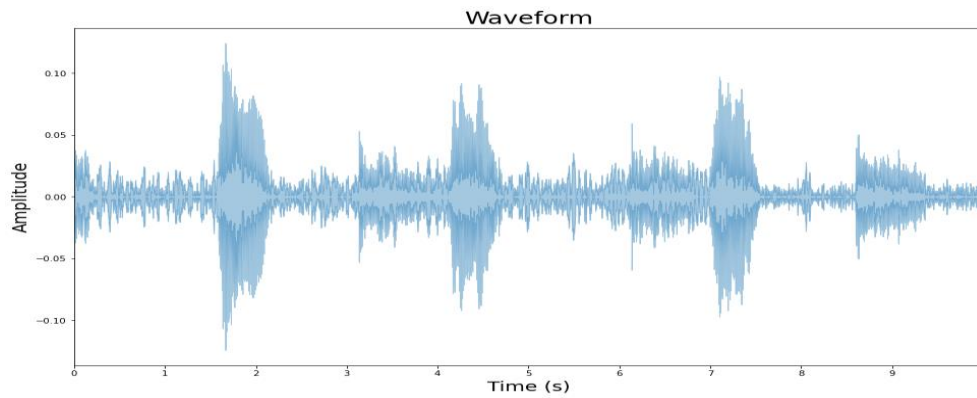
ภาพประกอบ 23 ขั้นตอนการเก็บข้อมูลเสียงจากกลุ่มตัวอย่าง

### 3.3 การจัดเตรียมข้อมูล (Preprocessing)

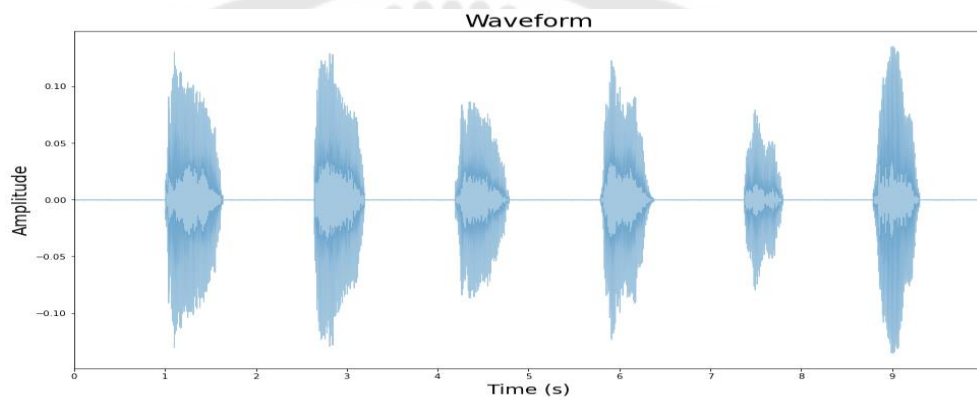
ในขั้นตอนนี้จะเป็นการเตรียมข้อมูลให้เหมาะสมสำหรับการพัฒนาระบบรู้จำเสียง โดยมีขั้นตอนดังต่อไปนี้

#### 3.3.1 การลบเสียงรบกวน และการปรับสภาพสัญญาณ (Noise reduction and Signal condition)

เนื่องจากขณะทำการบันทึกเสียงได้มีเสียงรบกวน (Noise) แทรกอยู่ ซึ่งสามารถมีผลต่อประสิทธิภาพในขั้นตอนการหาลักษณะสำคัญของเสียง (Feature extraction) จึงจำเป็นต้องลบเสียงรบกวนออก โดยในงานวิจัยนี้ใช้โปรแกรม Audacity ในการลดเสียงรบกวน วิธีการลดเสียงรบกวนใช้การวิเคราะห์แบบการแปลงฟูเรียร์อย่างรวดเร็ว (Fast Fourier Transform) (37) โดยค้นหาสเปกตรัมของโทนเสียงบริสุทธิ์ที่ประกอบเป็นเสียงพื้นหลังในส่วนเสียงเงียบที่เลือก ซึ่งเรียกว่าสเปกตรัมความถี่ (Frequency spectrum) ของเสียง ซึ่งจะสร้างลายนิ้วมือของเสียงรบกวนพื้นหลังคงที่ในไฟล์เสียง เมื่อทำการลดเสียงรบกวนจากเสียงโดยรวมระบบจะค้นหาสเปกตรัมความถี่ของแต่ละช่วงสั้นๆ ของเสียง โทนเสียงบริสุทธิ์ใดๆ ที่ไม่ตรงกับระดับเฉลี่ยในลายนิ้วมือจะถูกลดระดับเสียงลง และเนื่องจากข้อมูลที่บันทึกได้ทำการบันทึกในรูปแบบ 1 คน ต่อ 1 ไฟล์เสียง ดังนั้นจึงมีการตัดข้อมูลออกเป็นคำศัพท์ต่างๆ โดยทุกคำศัพท์จะกำหนดเวลาให้เท่ากันคือ 1 วินาที

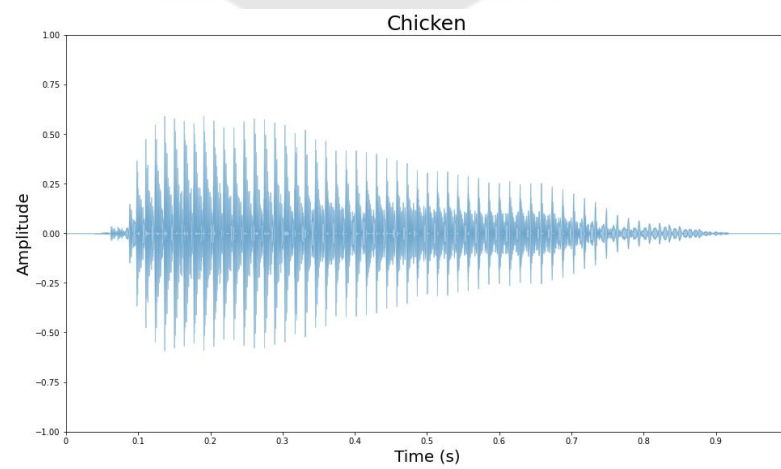


(ก)



(ข)

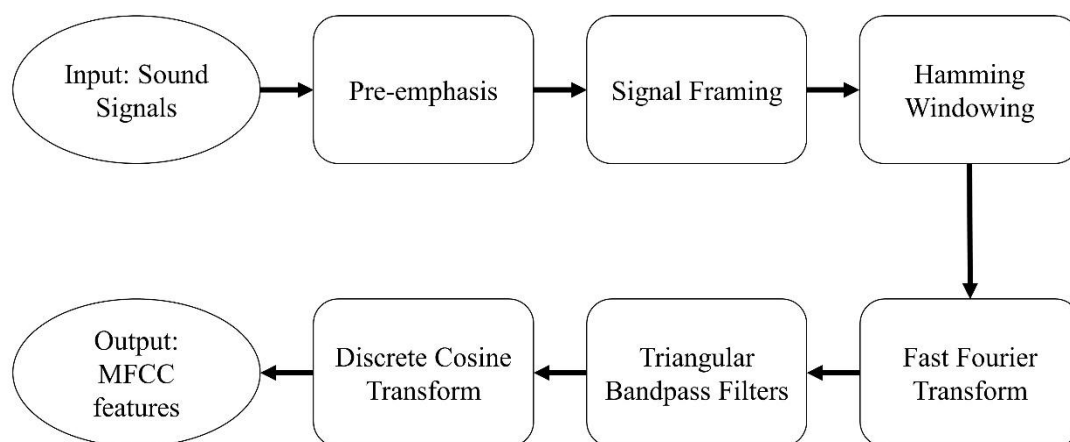
ภาพประกอบ 24 ตัวอย่างรูปคลื่นของสัญญาณเสียง (ก) มีเสียงรบกวน (ข) ไม่มีเสียงรบกวน



ภาพประกอบ 25 ตัวอย่างรูปคลื่นของเสียง "ไก่"

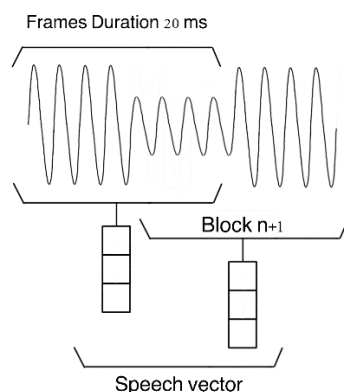
### 3.3.2 การหาลักษณะสำคัญของเสียง (Feature extraction)

ในขั้นตอนนี้จะเป็นการหาลักษณะสำคัญของเสียงพูดนั้นๆ เพื่อหาค่าพารามิเตอร์ต่างๆ ที่ใช้เป็นตัวแทนของเสียงพูดนั้นๆ และสามารถลดจำนวนข้อมูลจากการแปลงข้อมูลจำนวนมากเป็นชุดข้อมูลที่มีจำนวนน้อยลง โดยยังคงคุณสมบัติสำคัญของข้อมูลเดิมไว้ สำหรับในงานวิจัยนี้จะใช้วิธีการหาลักษณะสำคัญของเสียงแบบ Mel-Frequency Cepstral Coefficient (MFCC) โดยสามารถแสดงภาพรวมการคำนวณหาค่าลักษณะสำคัญดังภาพประกอบ 26



ภาพประกอบ 26 ขั้นตอนการหาค่าลักษณะสำคัญของ MFCC

จากภาพประกอบ 26 ในขั้นตอนแรกจะเป็นการนำสัญญาณเสียงเข้ามา แล้วไปขั้นตอน Pre-emphasis นั่นก็คือสัญญาณเสียงจะถูกส่งผ่านตัวกรองที่เน้นที่ความถี่ที่สูง โดยจะเพิ่มพลังงานให้กับสัญญาณที่มีความถี่สูง จากนั้นสัญญาณเสียงจะถูกแบ่งออกเป็นช่วงสั้นๆ 20 มิลลิวินาที หรือที่เรียกว่าเฟรม (Frames) และมีการใช้ Frames Overlap กันขนาด 10 มิลลิวินาที ดังภาพประกอบ 27 หลังจากนั้น Frames แต่ละ Frames จะถูกรวมด้วย Hamming Window function เพื่อรักษาสัญญาณให้มีความต่อเนื่องกัน ในขั้นตอนต่อไปจะเป็นการใช้เทคนิค Fast Fourier Transform (FFT) เพื่อแปลง time domain เป็น frequency domain โดยผลลัพธ์จะเป็นแบบ spectrum หลังจากนั้นสัญญาณจะผ่านชุดตัวกรอง Triangular band pass filters แล้วผ่าน Discrete Cosine Transform (DCT) จะทำการแปลง frequency domain เป็น quefrequency domain และใช้ Mel-scale cepstral coefficients จำนวน 13 ซึ่งจะได้คุณลักษณะ MFCC ออกมา



ภาพประกอบ 27 การแบ่งสัญญาณเสียง

โดยผลลัพธ์ของ MFCC จะทำการบันทึกออกเป็น 2 แบบ คือ

1) ทำการบันทึกในรูปแบบไฟล์ \*.csv โดยจะแบ่งเป็นคอลัมน์ต่างๆ ได้แก่ filename, chroma\_stft, rmisspectral\_centroid, spectral\_bandwidth, rolloff, zero\_crossing\_rate, mfcc 1-13 และ label สำหรับเป็นฐานข้อมูลของวิธีการ ANN, Neural Network และ KNN

filename	chroma	rmisspectr	spectral	ba	zero_cros	mfcc1	mfcc2	mfcc3	mfcc4	mfcc5	mfcc6	mfcc7	mfcc8	mfcc9	mfcc10	mfcc11	mfcc12	mfcc13	label
bird19.wav	0.3218	0.029012	1249.238	1083.13	1815.52	0.07483	-513.94	141.55	12.73	-52.933	-33.082	9.4309	-5.37	-38.18	-45.362	-14.11	-15.88	-4.3814	bird
bird17.wav	0.2861	0.033247	1131.718	1203.47	2151.27	0.04318	-510.05	133.911	5.1943	-50.305	-27.46	6.6664	-23.72	-38.96	-31.223	-14.75	-14.05	-16.335	bird
bird18.wav	0.2858	0.01764	1270.964	1109.2	2135.38	0.0744	-566.66	133.283	15.058	-51.607	-33.673	2.036	-14.9	-35.34	-41.671	-16.49	-3.773	-6.897	bird
bird25.wav	0.3468	0.036058	1007.841	1313.89	1473.62	0.08205	-481.65	140.039	38.441	16.304	-17.638	-30.73	-14.61	-6.041	-6.0855	-9.739	-8.85	-1.5802	bird
bird12.wav	0.4243	0.0345	1214.613	1341.46	2229.18	0.13099	-496.7	148.089	35.174	19.993	-13.215	-21.99	-13	-5.792	-4.5109	-15.23	-5.446	-4.5988	bird
bird5.wav	0.3836	0.030485	1565.243	1187.95	2162.21	0.1747	-516.44	146.549	45.257	21.289	-10.867	-26.22	-9.722	-4.24	-4.2207	-12.03	-11.23	-0.1396	bird
bird37.wav	0.3665	0.089842	2105.935	1495.79	2924.6	0.25411	-388.91	122.07	-3.3948	7.5854	-14.788	-24.44	10.738	-4.728	-9.3576	2.8961	-15.44	4.333	bird
bird1.wav	0.3903	0.077748	2194.402	1554.65	3210.79	0.25401	-404.34	119.028	-2.0118	1.3193	-10.472	-16.77	10.073	-7.757	-7.4579	2.5963	-13.27	4.9571	bird
bird26.wav	0.3458	0.074393	1822.171	1311.79	2655.62	0.20882	-407.39	118.456	-0.2631	4.6169	-7.5959	-17.27	8.3091	-5.224	-3.6052	3.8537	-8.401	5.8032	bird
bird3.wav	0.3848	0.07635	1803.508	1626.22	2706.01	0.21356	-445.31	93.4556	32.52	15.93	3.0249	-7.145	-8.231	8.0537	-6.124	-5.385	-5.576	7.8034	bird

ภาพประกอบ 28 ตัวอย่าง MFCC ในรูปแบบไฟล์ \*.csv

2) ทำการบันทึกในรูปแบบไฟล์ \*.png ซึ่งจะเป็นภาพ spectrogram (3D) แล้วทำการแปลง 3D spectrogram เป็น 2D spectrogram เพื่อลดจำนวนข้อมูลในการคำนวณของระบบ เนื่องจากเราสนใจเฉพาะความแตกต่างของขนาดแมกนิจูด (Magnitude) และ ความถี่ (Frequency) ของสัญญาณ โดยภาพ 3D spectrogram จะมี 3 แกน ได้แก่ x, y และ z ส่วนภาพ 2D spectrogram จะตัดแกน z ออก เพราะฉะนั้นในภาพ 2D spectrogram จะมี 2 แกน ได้แก่ x และ y การทำวิธีนี้จะช่วยให้ระบบสามารถวิเคราะห์ข้อมูลได้เร็วยิ่งขึ้น โดยผลลัพธ์นี้จะใช้เป็นฐานข้อมูลสำหรับวิธีการ CNN

```

Epoch 98/100
6/6 [=====] - 4s 665ms/step - loss: 0.0054 - accuracy: 0.9992 - val_loss: 1.1137 - val_accuracy: 0.7969
Epoch 99/100
6/6 [=====] - 4s 658ms/step - loss: 0.0040 - accuracy: 1.0000 - val_loss: 1.0817 - val_accuracy: 0.7969
Epoch 100/100
6/6 [=====] - 4s 625ms/step - loss: 0.0062 - accuracy: 1.0000 - val_loss: 1.0991 - val_accuracy: 0.7812
Processing time: 10.031755419572194

```

ภาพประกอบ 29 ผลลัพธ์จากการเรียนรู้ของโมเดลโดยใช้ฐานข้อมูลภาพ 3D spectrogram

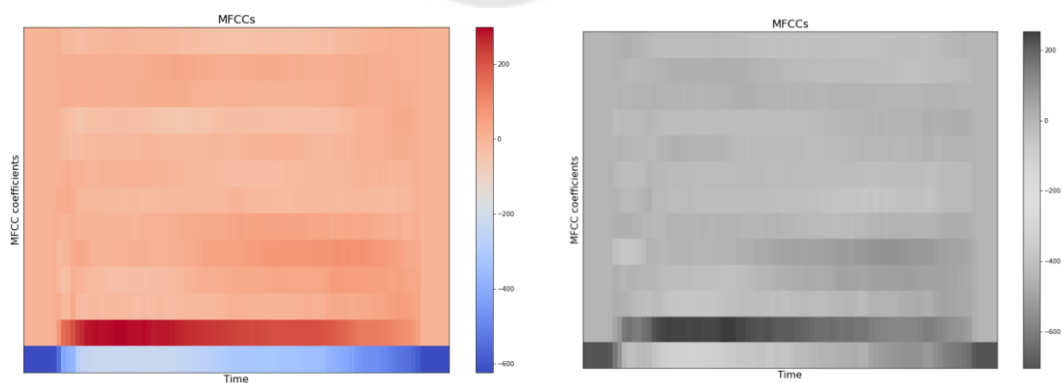
```

Epoch 98/100
6/6 [=====] - 4s 658ms/step - loss: 0.0216 - accuracy: 0.9916 - val_loss: 0.8695 - val_accuracy: 0.8281
Epoch 99/100
6/6 [=====] - 4s 725ms/step - loss: 0.0084 - accuracy: 0.9972 - val_loss: 0.8848 - val_accuracy: 0.8438
Epoch 100/100
6/6 [=====] - 4s 668ms/step - loss: 0.0280 - accuracy: 0.9944 - val_loss: 1.1184 - val_accuracy: 0.7969
Processing time: 7.408873788515726

```

ภาพประกอบ 30 ผลลัพธ์จากการเรียนรู้ของโมเดลโดยใช้ฐานข้อมูลภาพ 2D spectrogram

จากภาพประกอบ 29 และ 30 แสดงถึงผลลัพธ์จากการเรียนรู้ของโมเดลโดยใช้วิธีการ CNN ในการจำแนกเสียง จากการสอบทวน (Validation) ของโมเดลที่ใช้ฐานข้อมูลภาพ 3D spectrogram ได้ค่าความถูกต้องที่ 78.12 เปอร์เซ็นต์ และใช้เวลาการเรียนรู้ของโมเดลที่ 10.03 นาที จากการสอบทวน (Validation) ของโมเดลที่ใช้ฐานข้อมูลภาพ 2D spectrogram ได้ค่าความถูกต้องที่ 79.69 เปอร์เซ็นต์ และใช้เวลาการเรียนรู้ของโมเดลที่ 7.41 นาที ซึ่งผลลัพธ์จากการเรียนรู้ของโมเดลโดยใช้ฐานข้อมูลทั้ง 2 แบบ จะเห็นได้ว่าฐานข้อมูลทั้ง 2 แบบ ได้ค่าความถูกต้องของการสอบทวนที่ใกล้เคียงกัน แต่ระยะเวลาที่ใช้ในการเรียนรู้ของโมเดลของฐานข้อมูลภาพ 2D spectrogram ใช้ระยะเวลาการเรียนรู้ที่น้อยกว่าฐานข้อมูลภาพ 3D spectrogram เนื่องจากฐานข้อมูลภาพ 2D spectrogram มีจำนวนข้อมูลที่น้อยกว่าฐานข้อมูล 3D spectrogram



ภาพประกอบ 31 ตัวอย่าง MFCC ของเสียง “ไก่” ในรูปแบบ 3D และ 2D spectrogram

### 3.4 การจัดแบ่งข้อมูล (Data Segmentation)

การแบ่งข้อมูลสำหรับการสอนโมเดลแบ่งออกเป็น 3 ส่วน คือ ฝึกสอน (Train) : การสอบ  
ทวน (Validation) : การทดสอบ (Test) เป็นอัตราส่วน 70 : 15 : 15 ตามลำดับ โดยแบ่งข้อมูลดัง  
ตาราง

ตาราง 2 การแบ่งข้อมูล

ประเภท	หมวด	Train	Validation	Test	รวม
10 กลุ่มข้อมูล	สัตว์	420	90	90	600
20 กลุ่มข้อมูล	สัตว์, อาหาร	840	180	180	1200
30 กลุ่มข้อมูล	สัตว์, อาหาร, อวัยวะ	1260	270	270	1800
40 กลุ่มข้อมูล	สัตว์, อาหาร, อวัยวะ, สิ่งของ	1680	360	360	2400
50 กลุ่มข้อมูล	สัตว์, อาหาร, อวัยวะ, สิ่งของ, คำกริยา	2100	450	450	3000
60 กลุ่มข้อมูล	สัตว์, อาหาร, อวัยวะ, สิ่งของ, คำกริยา, สี	2520	540	540	3600
66 กลุ่มข้อมูล	สัตว์, อาหาร, อวัยวะ, สิ่งของ, คำกริยา, สี, Ling Six Sounds	2772	594	594	3960

### 3.5 การพัฒนาระบบรู้จำเสียง

ในขั้นตอนจะเป็นการจำแนกเสียงพูด (Speech Classification) โดยการใช้วิธีการ CNN, ANN, Neural Network และ KNN โดยได้ผลลัพธ์ (output) เป็นการรู้จำเสียงของระบบ ซึ่งเป็นสมมติฐานจากการรู้จำ (Recognition hypothesis)

#### 3.5.1 Convolutional Neural Network (CNN)

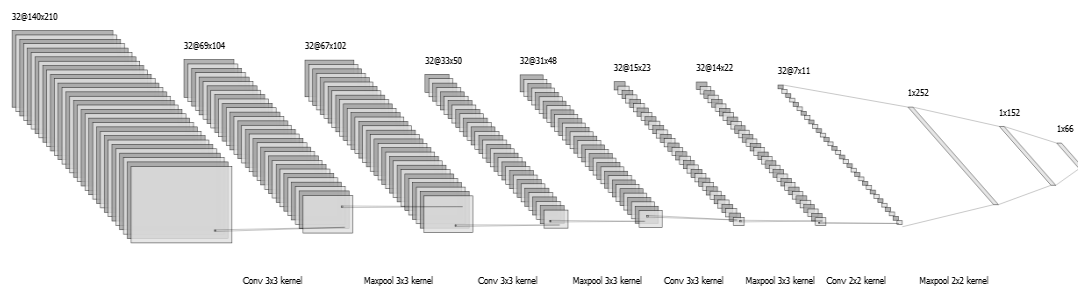
เป็นวิธีการการเรียนรู้เชิงลึกที่สามารถจำแนกข้อมูลด้วยภาพอินพุต, กำหนด  
ความสำคัญ (น้ำหนักและ biases ที่เรียนรู้ได้) ให้กับวัตถุต่างๆในภาพและสามารถแยกความ  
แตกต่างจากวัตถุอื่น ๆ ได้ สำหรับ CNN การ preprocessing ที่จำเป็นจะต่ำกว่าวิธีการการจำแนก

ประเภทอื่น ๆ เนื่องจาก CNN มีความสามารถในการเรียนรู้คุณสมบัติเหล่านี้จากการฝึกอบรวมข้อมูล Input

ในการฝึกและทดสอบโมเดล CNN จะใช้ภาพอินพุตขนาด 210 x 140 pixels โดยแต่ละภาพจะส่งผ่าน Convolution layers ชุดหนึ่งด้วยฟิลเตอร์ (Kernels), Pooling โดยในงานวิจัยนี้ จะใช้ Pooling ประเภท Max-Pooling, Fully Connected layers (FC) และใช้ฟังก์ชัน Softmax เพื่อจำแนกวัตถุที่มีค่าความน่าจะเป็นระหว่าง 0 ถึง 1 ซึ่งในงานวิจัยนี้มีการปรับโมเดลดังตาราง 3 และสามารถแสดงเป็นแผนผังดังภาพประกอบ 29

ตาราง 3 CNN layers

Layer	Kernel	Output Size
Conv 1	3 x 3	140 x 210 x 32
Max-Pooling 1	3 x 3	69 x 104 x 32
Conv 2	3 x 3	67 x 102 x 32
Max-Pooling 2	3 x 3	33 x 50 x 32
Conv 3	3 x 3	31 x 48 x 32
Max-Pooling 3	3 x 3	15 x 23 x 32
Conv 4	2 x 2	14 x 22 x 32
Max-Pooling 4	2 x 2	7 x 11 x 32
FC 1	-	252
FC 2	-	152
FC 3	-	Number Output Classes

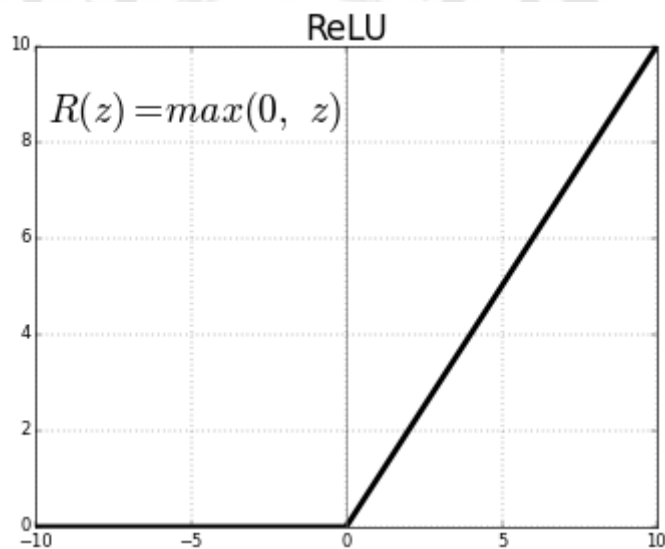


ภาพประกอบ 32 CNN architecture

Activation function (38) คือ Node ที่วางไว้ที่ส่วนท้ายหรือระหว่าง Neural Networks มีหน้าที่ช่วยในการตัดสินใจว่าข้อมูลใน Node ควรจะผ่านไปยัง Node ต่อไปหรือไม่ โดยในงานวิจัยนี้จะใช้ Activation function ประเภท ReLU (Rectified Linear Unit) function ซึ่งจะอยู่ใน Convolution layers และ FC layers ยกเว้น FC layer สุดท้าย จะใช้ Softmax function

ReLU function (38) เป็น Activation function ที่ใช้กันอย่างแพร่หลายมากที่สุด ในโครงข่ายประสาทเทียมในปัจจุบัน ReLU function จะแปลง Input ( $x$ ) เชิงลบทั้งหมดเป็นศูนย์และ Node จะไม่ทำงาน ทำให้การคำนวณมีประสิทธิภาพมากขึ้นเนื่องจากการเปิดใช้งาน Node เพียงไม่กี่ตัวต่อครั้ง สามารถแสดงเป็นสมการดังนี้

$$f(x) = \max(0, x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases} \quad (8)$$



ภาพประกอบ 33 ReLU function

ที่มา : [https://miro.medium.com/max/724/1\\*njuH4XVXf-l9pR\\_RorUOrA.png](https://miro.medium.com/max/724/1*njuH4XVXf-l9pR_RorUOrA.png)

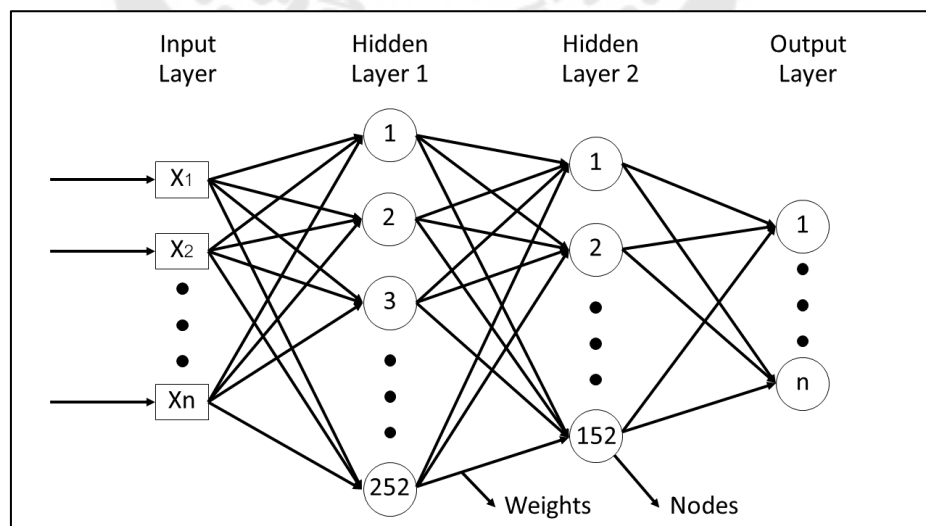
Softmax function (39) เป็นฟังก์ชันที่เปลี่ยนเวกเตอร์ของ Logit จำนวนจริงให้เป็นเวกเตอร์ของค่าจริง K ที่รวมเป็น 1 โดยค่า Input อาจจะเป็นค่าบวก ค่าลบ ศูนย์ หรือมากกว่าหนึ่งค่า แต่ Softmax function จะเปลี่ยนค่าเหล่านี้เป็นค่าระหว่าง 0 และ 1 เพื่อให้สามารถหาค่าความน่าจะเป็นว่าเป็นเท่าไร โดยสมการ Softmax function สามารถแสดงได้ดังนี้

$$\sigma(\vec{x})_i = \frac{e^{x_i}}{\sum_{j=1}^K e^{x_j}} \quad (9)$$

เมื่อ  $\vec{x}$  คือ ค่าเวกเตอร์ Input,  $x_i$  คือ ค่าทั้งหมดที่เป็นองค์ประกอบของเวกเตอร์ Input,  $K$  คือ จำนวนของกลุ่มข้อมูล

### 3.5.2 Artificial Neural Network (ANN)

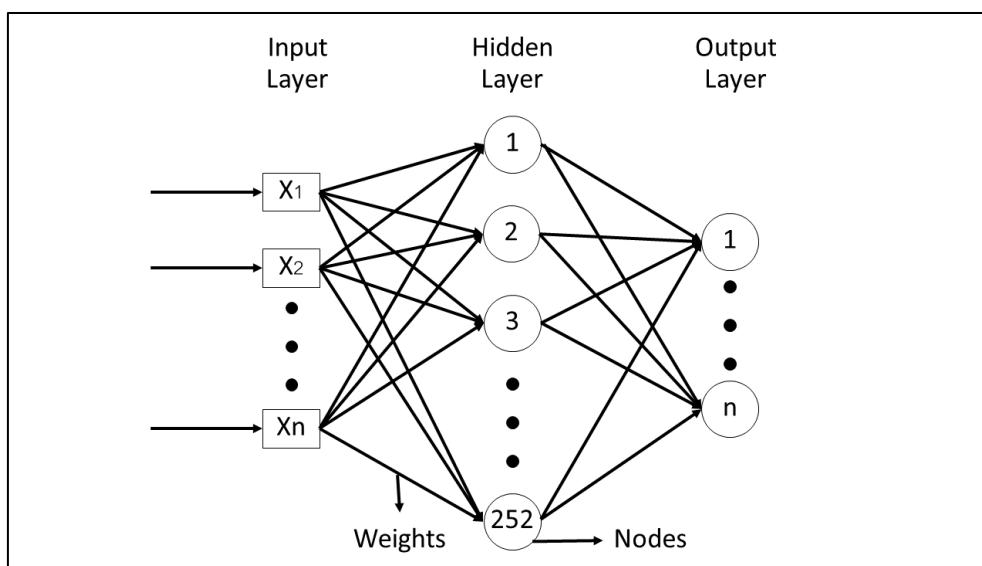
Artificial Neural Network (ANN) เป็นวิธีการการเรียนรู้เชิงลึกที่ถูกพัฒนาจากการทำงานของสมองมนุษย์เป็นต้นแบบ โดยสมองมนุษย์ประกอบไปด้วยหน่วยประมวลผลเรียกว่า นิวรอน ( เซลล์ประสาท หรือ neuron) ซึ่งมีจำนวนมากและเชื่อมต่อกันอย่างซับซ้อน ซึ่งใน ANN ประกอบไปด้วย 1 ชั้นอินพุต (Input layer) , 1 ชั้นเอาต์พุต (Output layer) โดยจะมีจำนวน nodes ตามจำนวนของ Output กลุ่มข้อมูล และ ชั้นซ่อน (Hidden layers) ซึ่งสามารถมีได้มากกว่า 1 ชั้น ซึ่งในงานวิจัยนี้ใช้ ANN ประเภท Feedforward ANN โดยมี Hidden layers จำนวน 2 ชั้น ชั้นที่ 1 มีจำนวน 252 nodes และ ชั้นที่ 2 มีจำนวน 152 nodes ซึ่งใน Hidden layers จะใช้ Activation function ประเภท ReLU function และใน Output layer จะใช้ Softmax function โดยสามารถแสดงแผนผังโครงสร้าง ANN ดังภาพประกอบ 31



ภาพประกอบ 34 ANN architecture

### 3.5.3 Neural Network (NN)

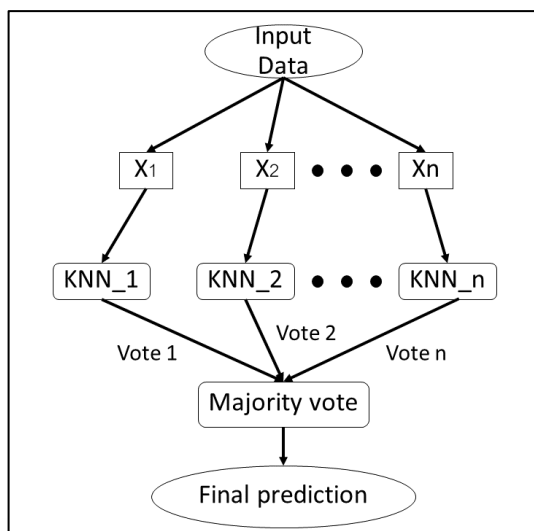
งานวิจัยนี้เลือกใช้ Multi-layer Perceptron classifier ซึ่งเป็นโครงข่ายประสาทเทียมที่มีหลายชั้น แต่จะประกอบไปด้วย 1 ชั้นอินพุต (Input layer) , 1 ชั้นซ่อน (Hidden layer) และ 1 ชั้นเอาต์พุต (Output layer) ซึ่งจะมีปรับและเลือกใช้ให้เหมาะสมกับข้อมูล ซึ่งงานวิจัยนี้ปรับชั้น Hidden layer โดยมีจำนวน 252 nodes และใช้ Activation function ประเภท ReLU function โดยสามารถแสดงแผนผังโครงสร้าง Neural Network ดังภาพประกอบ 32



ภาพประกอบ 35 Neural Network architecture

### 3.5.4 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) เป็นวิธีการที่ใช้ในการแบ่งกลุ่มข้อมูล โดยจะคำนวณระยะทางของข้อมูลว่าอยู่ใกล้กับกลุ่มข้อมูลไหนมากที่สุด โดยการจะตรวจสอบข้อมูลบางจำนวน (K) ซึ่งในงานวิจัยนี้ใช้ค่า K หรือจำนวนข้อมูลที่ใช้ในการตรวจสอบข้อมูลเท่ากับ 1 และใช้สมการหาระยะทางแบบ Euclidean ในการคำนวณหาระยะทางของข้อมูลฝึกและข้อมูลทดสอบ โดยสามารถแสดงแผนผังโครงสร้าง KNN ดังภาพประกอบ 33



ภาพประกอบ 36 KNN architecture

### 3.6 การวัดประสิทธิภาพ (Performance test)

ในงานวิจัยนี้จะใช้เมทริกซ์ความสับสน (Confusion Matrix) ในการวัดความถูกต้องของระบบ (Accuracy), ค่าความแม่นยำ (Precision), ความถูกต้องของแต่ละกลุ่มข้อมูล (Sensitivity) และ F1-score ของระบบรู้จำเสียงพูด

#### 3.6.1 เมทริกซ์ความสับสน (Confusion Matrix)

Confusion matrix (40) คือ ตารางใช้ประเมินประสิทธิภาพของโมเดลในการจำแนกกลุ่มข้อมูล ซึ่งเป็นเมทริกซ์ขนาด  $N \times N$  ซึ่ง  $N$  คือ จำนวนคลาสที่จำแนก โดยเมทริกซ์จะเปรียบเทียบระหว่างผลเฉลยกับผลที่ได้จากการทำนายของโมเดล ตารางนี้จะช่วยให้เห็นองค์ประกอบโดยรวมของการจัดหมวดหมู่ของงานวิจัยว่ามีประสิทธิภาพเพียงพอหรือไม่ และเกิดข้อผิดพลาดประเภทใด ดังตาราง 4

ตาราง 4 Confusion Matrix

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	True Positive (TP)	False Positive (FP)
	Negative (0)	False Negative (FN)	True Negative (TN)

เมื่อ True Positive (TP) คือ สิ่งที่ทำแบบจำลองทำนาย (Output) ว่าเป็น 1 และข้อมูลจริง (Status) บอกว่าเป็น 1

True Negative (TN) คือ สิ่งที่ทำแบบจำลองทำนาย (Output) ว่าเป็น 0 และข้อมูลจริง (Status) บอกว่าเป็น 0

False Positive (FP) คือ สิ่งที่ทำแบบจำลองทำนาย (Output) ว่าเป็น 1 และข้อมูลจริง (Status) บอกว่าเป็น 0

False Negative (FN) คือ สิ่งที่ทำแบบจำลองทำนาย (Output) ว่าเป็น 0 และข้อมูลจริง (Status) บอกว่าเป็น 1

โดยตาราง Confusion matrix สามารถใช้ประเมินประสิทธิภาพได้หลายส่วน ได้แก่

1) ความถูกต้องของระบบ (Accuracy) คือ จำนวนครั้งที่ทำนายถูกหารด้วยจำนวนครั้งที่ทำนายทั้งหมด หมายความว่าทำนายถูกต้องและแม่นยำแค่ไหน แบบรวม ๆ โดยคำนวณจากสมการดังนี้

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (10)$$

2) ค่าความแม่นยำ (Precision) คือ จำนวนครั้งที่ทำนายว่าเป็น 1 แล้วถูก หารด้วยจำนวนครั้งที่ทำนายว่าเป็น 1 ทั้งหมด หมายความว่าระบบมีความแม่นยำแค่ไหน โดยคำนวณจากสมการดังนี้

$$\mathbf{Precision} = \frac{TP}{(TP+FP)} \quad (11)$$

3) ความถูกต้องของแต่ละกลุ่มข้อมูล (Sensitivity) คือ จำนวนครั้งที่ทำนายว่าเป็น 1 แล้วถูก หารด้วยจำนวนเป็น 1 ทั้งหมดในข้อมูล หมายความว่าความถูกต้องครอบคลุมแค่ไหน โดยคำนวณจากสมการดังนี้

$$\mathbf{Sensitivity} = \frac{TP}{TP+FN} \quad (12)$$

4) F1-score คือ ค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ Sensitivity โดยคำนวณจากสมการดังนี้

$$\mathbf{F1} = 2 \times \left( \frac{\mathbf{precision} \times \mathbf{sensitivity}}{\mathbf{precision} + \mathbf{sensitivity}} \right) \quad (13)$$

## บทที่ 4

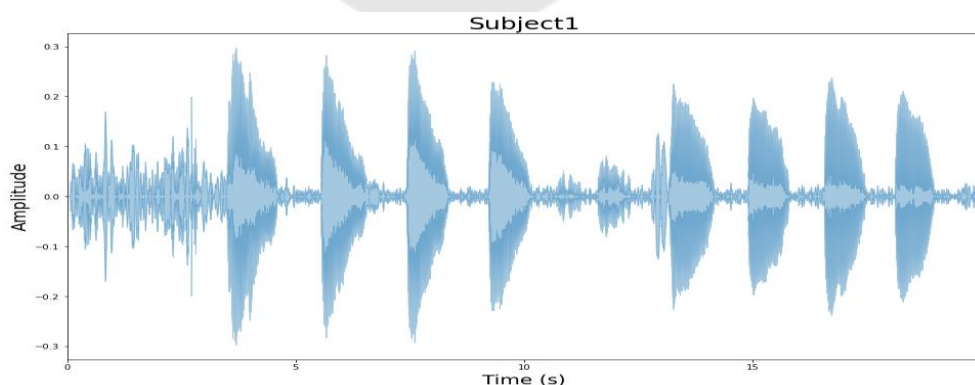
### ผลการดำเนินงานวิจัย

การวิจัยเพื่อให้ได้มาซึ่งการพัฒนาระบบรู้จำเสียงพูด เพื่อฟื้นฟูการสื่อสารสำหรับผู้ที่มีภาวะหูตึง ผู้วิจัยได้ดำเนินการวิจัยโดยการศึกษาตามขอบเขตและขั้นตอนต่างๆ จนกระทั่งประเมินประสิทธิภาพของระบบที่สร้างขึ้นและให้เป็นไปตามวัตถุประสงค์ที่ได้กำหนดไว้ได้ดังนี้

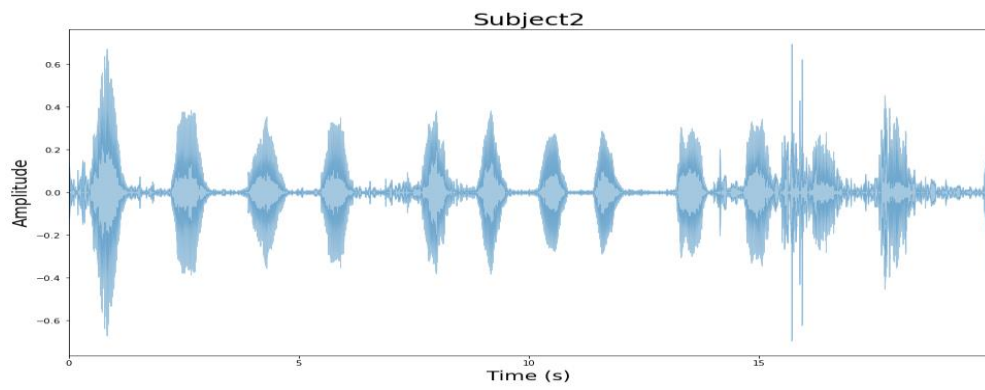
1. ผลลัพธ์ของฐานข้อมูลเสียงพูด (Speech database)
2. ผลลัพธ์ของการจัดเตรียมข้อมูล (Preprocessing)
3. ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ CNN
4. ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ ANN
5. ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ Neural Network
6. ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ KNN

#### 4.1 ผลลัพธ์ของฐานข้อมูลเสียงพูด (Speech database)

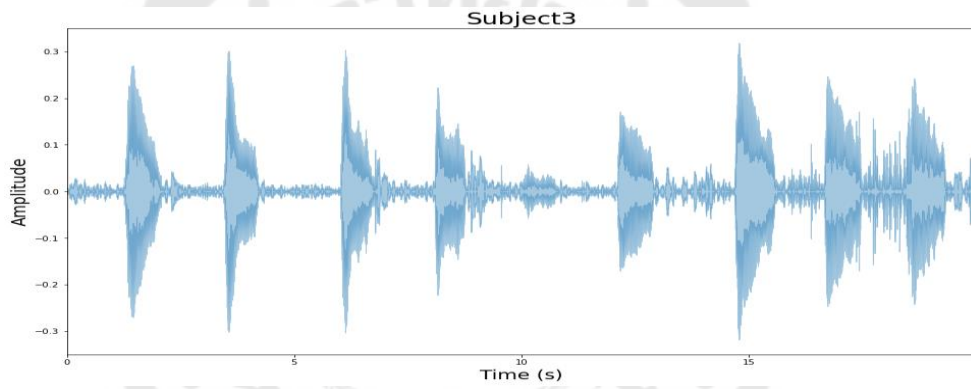
จากการบันทึกข้อมูลเสียงที่เก็บจากผู้ป่วยชายจำนวน 10 คน และหญิงจำนวน 10 คน ที่ผ่าตัดประสาทหูเทียมจากหูทั้งสองข้างที่หน่วย ผู้ป่วยนอกแผนกโสต ศอ นาสิก ราลิงซ์ และ คลินิก ผ่าตัด ณ โรงพยาบาลศูนย์การแพทย์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารี ได้ผลลัพธ์ดังนี้



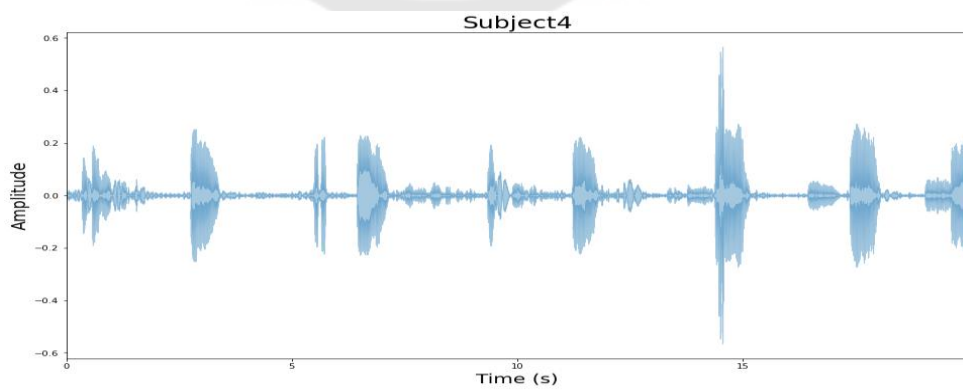
(ก)



(๑)



(๒)



(๓)

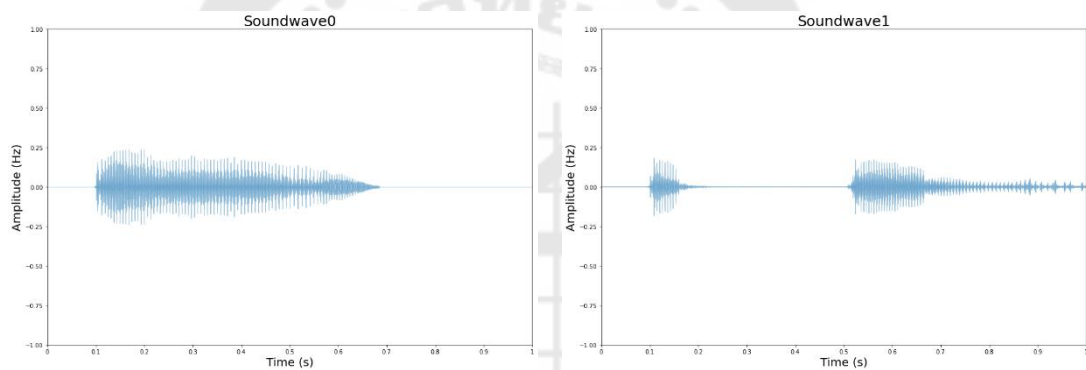
ภาพประกอบ 37 ตัวอย่างรูปคลื่นของสัญญาณเสียงจากฐานข้อมูลเสียงพูด (ก) กลุ่มตัวอย่างคนที่ 1 (ข) กลุ่มตัวอย่างคนที่ 2 (ค) กลุ่มตัวอย่างคนที่ 3 (ง) กลุ่มตัวอย่างคนที่ 4

#### 4.2 ผลลัพธ์ของการจัดเตรียมข้อมูล (Preprocessing)

การเตรียมข้อมูลให้เหมาะสมสำหรับการพัฒนาระบบรู้จำเสียง ได้ผลลัพธ์ดังต่อไปนี้

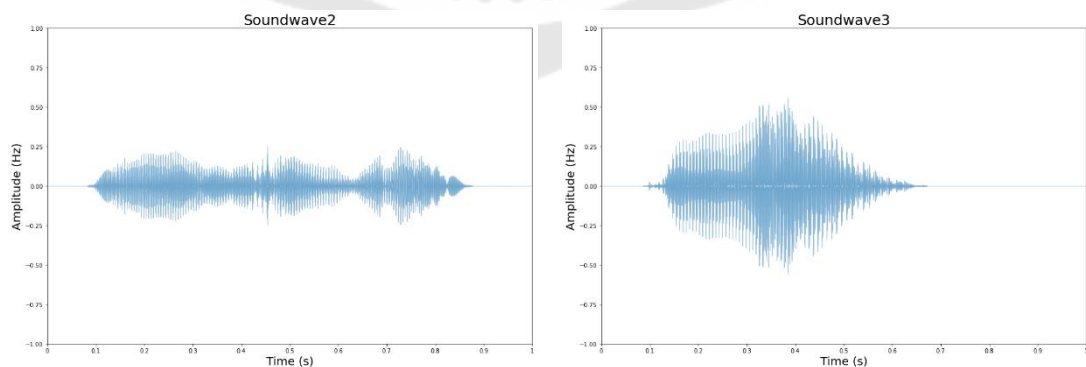
##### 4.2.1 ผลลัพธ์ของการลบเสียงรบกวน และการปรับสภาพสัญญาณ (Noise reduction and Signal condition)

จากการลบเสียงรบกวนออก โดยใช้เทคนิค Fast Fourier Transform (FFT) ในการลบเสียงรบกวน และการตัดข้อมูลออกเสียงเป็นคำศัพท์ต่างๆ โดยทุกคำศัพท์จะกำหนดเวลาให้เท่ากันคือ 1 วินาที ได้ผลลัพธ์ดังต่อไปนี้



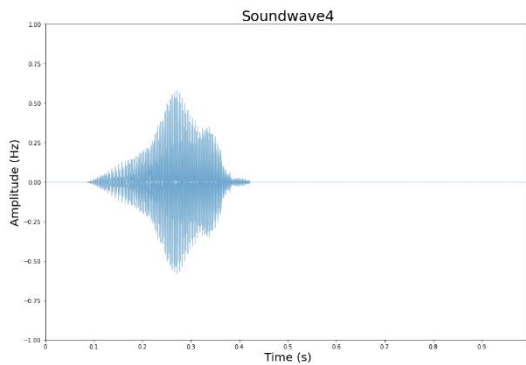
(ก)

(ข)

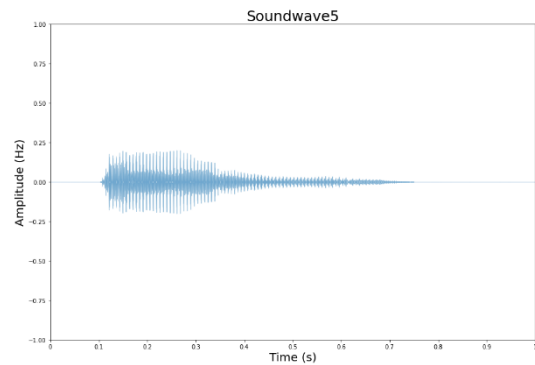


(ค)

(ง)



(จ)



(ข)

ภาพประกอบ 38 ตัวอย่างรูปคลื่นของสัญญาณเสียงที่ลบเสียงรบกวนและตัดข้อมูล (ก) เสียง “อา”  
 (ข) เสียง “แอปเปิ้ล” (ค) เสียง “แซน” (ง) เสียง “กล้วย” (จ) เสียง “นก” (ฉ) เสียง “ดำ”

### 4.2.2 ผลลัพธ์การหาลักษณะสำคัญของเสียง (Feature extraction)

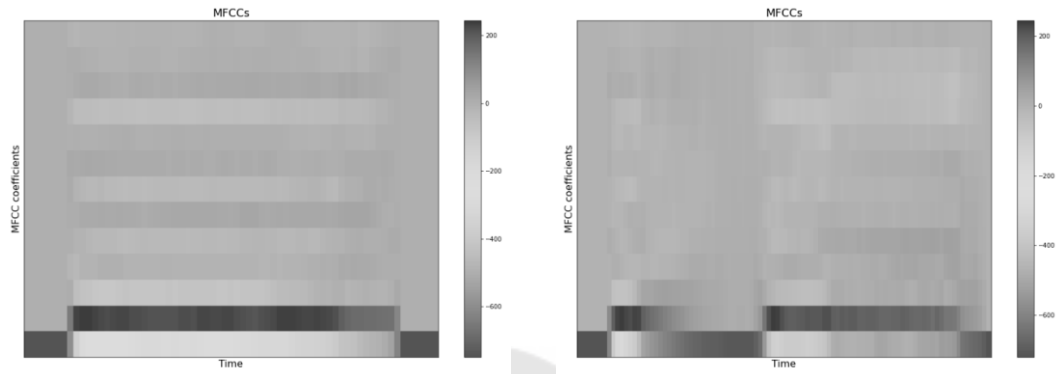
การหาลักษณะสำคัญของเสียงด้วยวิธีการ MFCC ได้ผลลัพธ์ 2 แบบ คือ

1) บันทึกข้อมูลในรูปแบบไฟล์ \*.csv โดยได้ผลลัพธ์ดังต่อไปนี้

chroma_stfmspectral	spectral	bar	rolloff	zero_cross	mfcc1	mfcc2	mfcc3	mfcc4	mfcc5	mfcc6	mfcc7	mfcc8	mfcc9	mfcc10	mfcc11	mfcc12	mfcc13	label
-1.17955	-0.902699	-0.200647	-1.77536	-0.44512	0.538194	-1.18018	0.604779	0.60982	-1.95938	-0.92271	1.27369	0.638547	-0.40011	-0.15412	0.172023	2.34333	1.11797	3
0.737214	0.151079	0.0745141	-0.00847	0.383469	-0.25097	1.04119	0.349955	-0.94352	-0.30044	-0.57147	-0.52807	1.37403	0.51486	-2.34318	0.852606	0.832091	-0.20156	41
0.752711	-0.442325	0.200202	1.11306	0.484796	0.455345	-0.34358	-0.65676	0.321738	1.40537	0.775901	-0.51337	-0.44691	-0.45003	-1.15178	-0.80308	-1.58294	0.387425	12
-0.531499	-0.629567	1.08721	0.182833	1.10134	0.894104	0.908579	-0.28459	-3.32231	0.244374	-0.94187	1.12386	0.905735	0.539597	0.617526	-0.41636	-0.0896	2.22592	0
-0.64938	-0.763963	-0.974785	-0.6241	-1.30754	-0.73657	-1.54746	-0.93418	0.669618	1.08336	1.75973	1.422	2.23385	2.50285	1.02447	1.25297	1.75488	-1.22584	31
-0.012201	0.236058	-0.921637	-1.15982	-0.73445	-1.31423	0.176327	-0.2955	-0.11135	-0.95237	-0.61855	-0.18159	-0.85283	-1.20643	0.014772	-0.89117	-0.27282	0.376388	29
2.33953	-0.693694	-0.111188	-1.22438	-0.24127	-0.05677	0.068844	1.77496	-1.31567	-1.38123	0.25727	-0.00111	0.620662	-0.14193	0.354995	0.828234	-0.32435	0.131584	48
1.92313	0.894196	-0.890421	-1.5509	-1.09491	-0.57833	-0.21886	-0.0409	0.425804	0.55558	1.20489	0.196166	-0.64563	1.33466	1.51087	-0.40955	0.799769	-0.55315	50
-0.928843	0.0455355	-0.0522104	1.91619	0.49527	-0.70628	0.361895	-0.90196	1.18286	1.41961	-0.0453	-0.91382	-0.40976	0.674118	-1.58229	-0.14665	-1.79063	-1.63222	32
1.75222	0.217344	0.0902185	-0.73985	-0.00772	0.076165	0.910976	1.02006	-1.59616	-1.15424	-0.91349	1.0935	0.858344	0.262501	1.24641	-0.25653	1.81635	0.42089	33
0.837499	-0.0636303	0.37892	0.729231	0.379394	0.437247	0.274088	-0.04396	0.190868	-0.59119	-0.72776	0.159835	1.47313	0.356695	0.483029	0.00729	-0.69869	0.615437	30
-1.08456	-0.651212	-0.90099	0.05082	-0.5734	-1.5255	-1.5685	-0.6877	1.19548	1.4373	0.80402	0.228514	0.082462	0.222816	0.347293	-0.31244	-1.65833	-0.19775	55
1.20048	0.383759	-0.578002	0.690424	0.133063	-0.6952	0.360143	-0.08944	0.510467	2.8138	1.58931	-0.49662	1.07377	-0.34128	0.548675	2.71057	-1.19642	1.45071	15
-0.087815	-0.623581	-0.504846	-0.63998	-0.47512	-0.61769	0.137676	0.750464	-0.76169	-0.13174	-0.01166	-0.5206	-0.28493	-1.24689	-0.30632	0.258717	0.594074	-0.30183	58
-0.516187	2.21849	-0.715416	-0.98942	-0.89942	-0.62322	0.765879	-0.14112	1.0744	-0.2022	-0.77962	-0.92363	-0.77478	-0.67801	-1.07625	-0.5924	0.363888	-1.67814	44
-1.07637	-0.876584	0.114176	0.442237	0.087075	-0.25777	-0.33044	0.183678	-0.72166	-2.07969	0.069871	1.01458	-1.42735	-0.28453	1.88075	-1.1787	-0.39551	2.56822	54
0.692383	0.587523	0.583716	0.551882	0.795289	0.409314	0.733506	-1.04341	-0.28865	0.315881	-0.20837	-0.49758	0.8828	1.65933	-0.63821	0.909198	0.533018	-0.66651	37
-0.549672	-0.198655	0.676039	-0.29837	0.582098	0.970532	-0.13196	-0.78572	-1.18965	-0.66287	-0.47765	1.27947	1.05859	-1.60239	-0.16731	-0.46123	-0.32734	-1.58038	37
0.709846	1.03888	-0.209105	0.569055	-0.10147	-0.46718	1.48056	0.260637	-0.23911	-0.53898	-1.1686	1.46077	0.186991	-0.76698	0.214302	-1.31485	0.22859	1.27704	7
1.00012	1.02318	-0.0395953	0.535113	0.152289	-0.10641	1.22844	0.354088	-1.14088	-0.11904	0.181901	-0.37146	1.80454	-0.16795	-0.00331	0.618657	-0.13139	0.655203	40
-0.629984	-0.807858	0.487378	-0.63384	0.365532	0.736351	-0.94356	0.344375	1.23331	-1.47792	-0.91505	1.46517	1.25438	0.189613	-1.17708	0.343429	1.00112	-0.08973	40
-1.11186	0.025628	-0.421409	-0.18054	-0.49659	-0.22585	-0.2483	-0.18652	-0.05122	0.271693	0.5654	0.236497	0.071451	-1.32171	-0.20278	0.506451	-0.72453	0.983708	60
-0.759973	-1.05014	-0.685095	-1.98318	-0.83851	-0.11629	-2.3449	0.675522	2.19823	0.81893	0.62428	0.070598	-1.94776	-2.89914	-0.84767	1.04391	1.99324	-0.65875	35
1.01088	-0.945682	1.72157	1.04009	1.55865	1.48103	1.17547	-1.76197	-2.48249	0.736837	1.1266	0.712213	-1.45334	-0.22402	0.046684	1.20125	-0.22316	0.735164	46
-1.69916	4.67759	-0.765691	-0.60026	-1.0376	-0.64356	1.39923	-0.23199	0.247002	0.498425	0.043915	0.376317	-0.98935	-1.19558	-0.13622	-1.62992	-1.3366	0.194588	59
-0.371828	0.562421	-0.968845	-0.64879	-0.87071	-0.86715	0.071541	0.347556	0.192034	1.04859	1.4158	0.667357	1.26803	0.767091	-1.57297	0.280112	1.58833	-0.46112	25
1.28984	-0.64491	1.0098	1.59958	1.01839	0.730729	0.001719	-0.61948	0.107357	-0.57774	0.046256	-0.21856	-1.1512	1.18975	-0.55112	-0.81831	-1.37313	-0.01281	41
-0.41229	-0.804785	-0.450534	0.348234	-0.72407	-0.65833	-0.82136	0.332605	0.619879	-0.68258	0.202917	0.685261	-0.97068	0.523671	-1.02719	-1.43816	-0.70572	-0.67871	56
2.24158	1.1015	1.03055	-0.67754	0.848923	1.38111	1.63961	0.285134	-3.3086	-1.20326	-0.12402	-0.35134	0.774024	-1.68789	-0.10018	-0.1083	-1.08652	0.613238	33

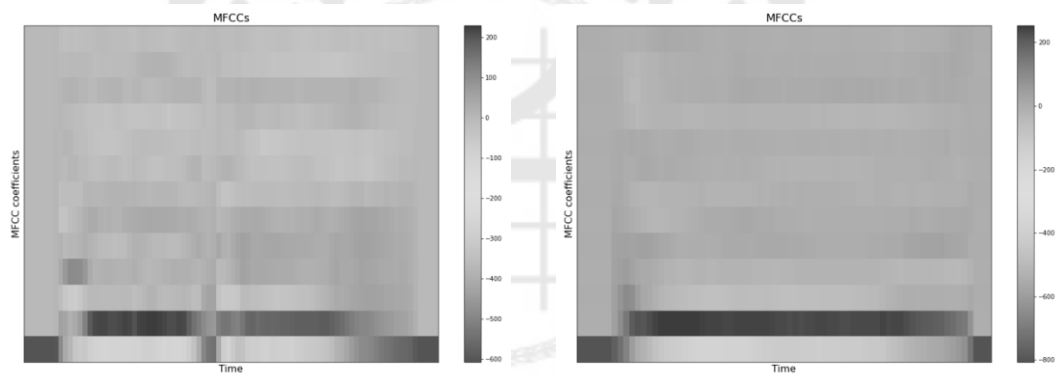
ภาพประกอบ 39 ตัวอย่างฐานข้อมูลของ MFCC ในรูปแบบไฟล์ \*.csv

2) บันทึกข้อมูลในรูปแบบไฟล์ \*.png โดยได้ผลลัพธ์ดังต่อไปนี้



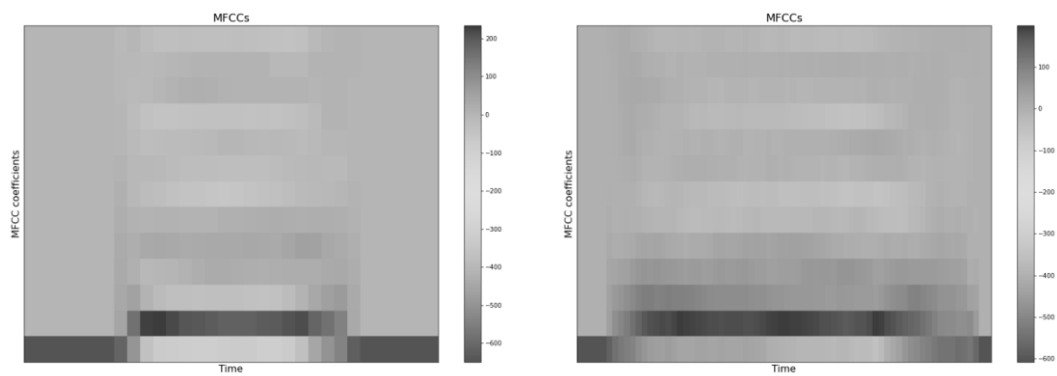
(ก)

(ข)



(ค)

(ง)



(จ)

(ฉ)

ภาพประกอบ 40 ตัวอย่างฐานข้อมูลของ MFCC ในรูปแบบไฟล์ \*.png (ก) เสียง “อา” (ข) เสียง “แอปเปิ้ล” (ค) เสียง “แซน” (ง) เสียง “กล้วย” (จ) เสียง “นก” (ฉ) เสียง “ดำ”

#### 4.3 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ CNN

Loss function (41) คือ การคำนวณปริมาณที่ model ควรพยายามลดให้เหลือน้อยที่สุดในระหว่างการ train ข้อมูล โดยในงานวิจัยนี้ได้ใช้ Loss function แบบ Categorical crossentropy ซึ่งเป็นฟังก์ชันที่ใช้ในงานการจำแนกประเภทที่มีหลาย กลุ่มข้อมูล โดยจะคำนวณจากค่าเฉลี่ยของ Cross-Entropy ที่เกิดจากการแจกแจงความน่าจะเป็นของ กลุ่มข้อมูล ต่างๆที่ model ทำนายได้ (Predicted) และค่าจริง (Actual) โดยใช้สมการดังสมการ (14) ดังนี้

$$\text{Loss} = - \sum_{i=1}^{\text{output size}} y_i \cdot \log \hat{y}_i \quad (14)$$

เมื่อ  $\hat{y}_i$  คือ ค่าสเกลาร์ลำดับที่  $i$  ใน output ของ model,  $y_i$  คือ ค่าที่สอดคล้องกับ target และ *output size* คือ จำนวนค่าสเกลาร์ ใน output ของ model

Optimization Algorithm (42) เป็นการปรับปรุงค่า error และ loss เพื่อให้ได้ค่าที่ดีที่สุด โดยจะทำการปรับค่าน้ำหนัก (Weight) และค่า Bias ที่เชื่อมกับ Node โดยในงานวิจัยนี้ได้ใช้ Optimizer แบบ Adam ซึ่งเป็น Optimizer ที่สามารถแก้ปัญหา decaying ของ gradients ในแต่ละ step ที่ผ่านมาได้และสามารถปรับค่า learning rates สำหรับพารามิเตอร์ในแต่ละครั้งได้

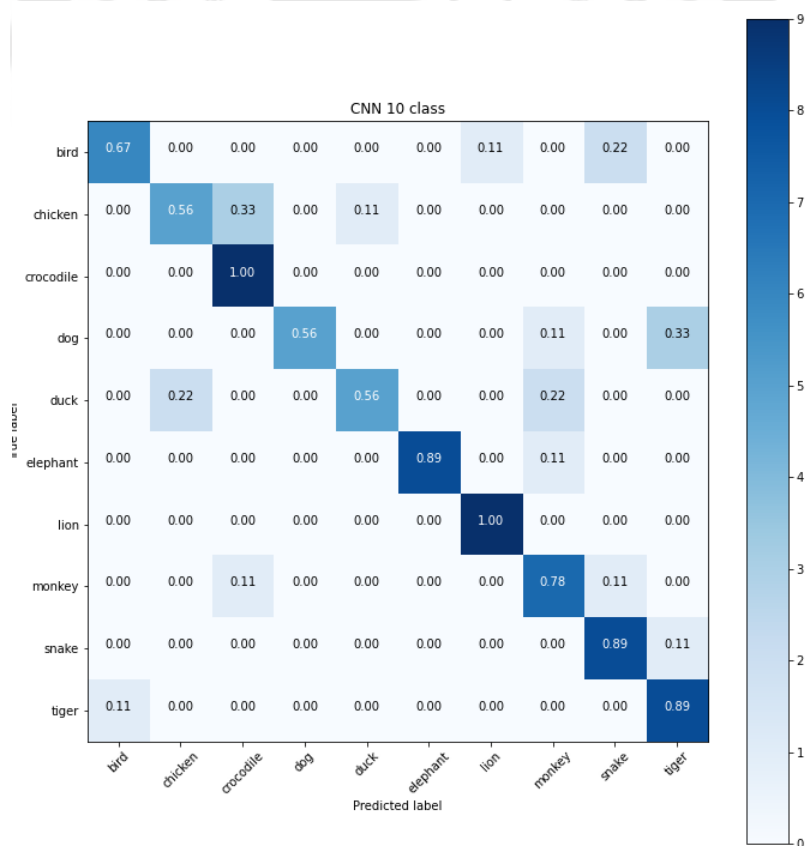
ในขั้นตอนการ train model ของ CNN ได้ใช้ epochs จำนวน 100 และนำ model ที่ได้ไปทดสอบกับ test data ในการประเมินประสิทธิภาพของ model จะใช้ Confusion Matrix ในการหาค่า Accuracy, Sensitivity, F1-score และ Precision ของ model ซึ่งได้ผลลัพธ์ดังตารางต่อไปนี้

ตาราง 5 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.86	0.67	0.75
ไก่	0.71	0.56	0.63
จระเข้	0.69	1.00	0.82
หมา	1.00	0.56	0.71

ตาราง 5 (ต่อ)

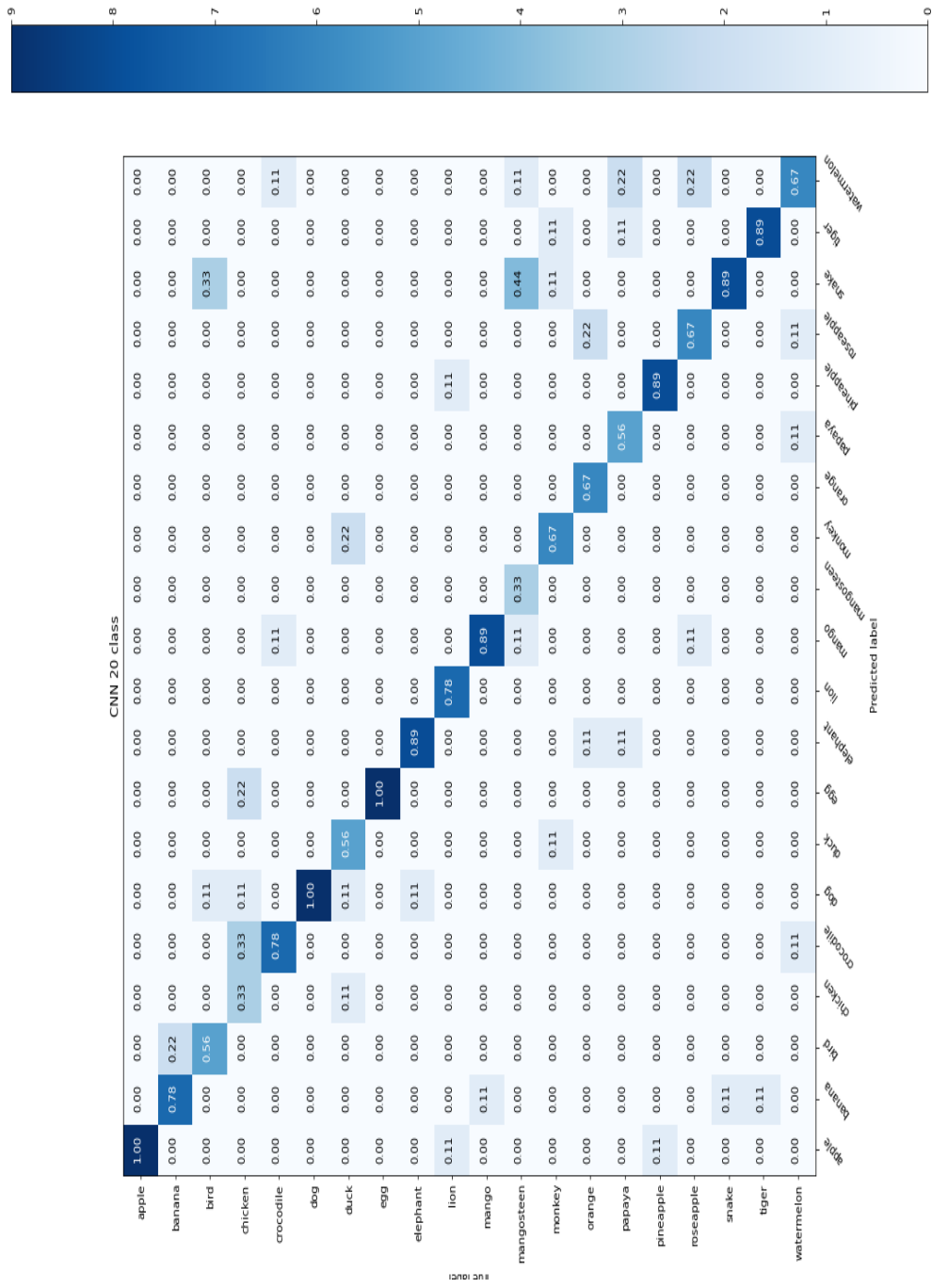
ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
เปิด	0.83	0.56	0.67
ข้าง	1.00	0.89	0.94
สิงโต	0.90	1.00	0.95
ลิง	0.64	0.78	0.70
งู	0.73	0.89	0.80
เสือดอ	0.67	0.89	0.76
Accuracy			0.78
Macro avg	0.80	0.78	0.77
Weighted avg	0.80	0.78	0.77



ภาพประกอบ 41 Confusion Matrix ของ CNN 10 กลุ่มข้อมูล

ตาราง 6 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล

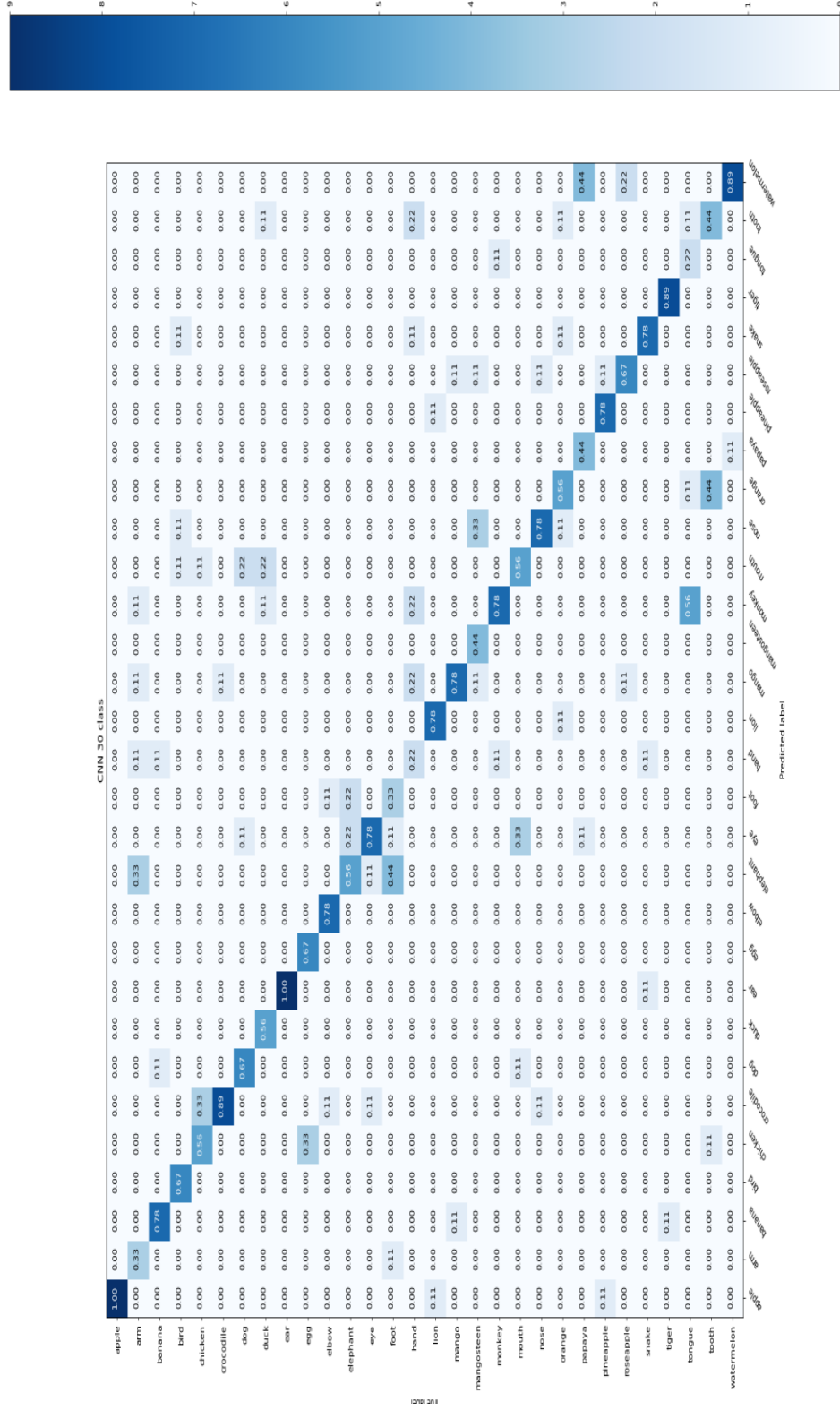
ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
แอปเปิ้ล	0.82	1.00	0.90
กล้วย	0.70	0.78	0.74
นก	0.71	0.56	0.63
ไก่	0.75	0.33	0.46
จระเข้	0.64	0.78	0.70
หมา	0.69	1.00	0.82
เบ็ด	0.83	0.56	0.67
ไข่	0.82	1.00	0.90
ข้าง	0.80	0.89	0.84
สิงโต	1.00	0.78	0.88
มะม่วง	0.73	0.89	0.80
มังคุด	1.00	0.33	0.50
ลิง	0.75	0.67	0.71
ส้ม	1.00	0.67	0.80
มะละกอ	0.83	0.56	0.67
สับปะรด	0.89	0.89	0.89
ชมพู	0.67	0.67	0.67
งู	0.50	0.89	0.64
เสือ	0.80	0.89	0.84
แตงโม	0.50	0.67	0.57
Accuracy			0.74
Macro avg	0.77	0.74	0.73
Weighted avg	0.77	0.74	0.73



ภาพประกอบ 42 Confusion Matrix ของ CNN 20 กลุ่มข้อมูล

ตาราง 7 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
แอปเปิ้ล	0.82	1.00	0.90
แขน	0.75	0.33	0.46
กล้วย	0.78	0.78	0.78
นก	1.00	0.67	0.80
ไก่	0.56	0.56	0.56
จระเข้	0.57	0.89	0.70
หมา	0.75	0.67	0.71
เปิด	1.00	0.56	0.71
หู	0.90	1.00	0.95
ไข่	1.00	0.67	0.80
ข้อศอก	1.00	0.78	0.88
ข้าง	0.38	0.56	0.45
ตา	0.47	0.78	0.58
เท้า	0.50	0.33	0.40
มือ	0.33	0.22	0.27
สิงโต	0.88	0.78	0.82
มะม่วง	0.54	0.78	0.64
มังคุด	1.00	0.44	0.62
ลิง	0.44	0.78	0.56
ปาก	0.45	0.56	0.50
จมูก	0.58	0.78	0.67
ส้ม	0.50	0.56	0.53
มะละกอ	0.80	0.44	0.57
สับปะรด	0.88	0.78	0.82
ชมพู่	0.60	0.67	0.63
งู	0.70	0.78	0.74
เสื่อ	1.00	0.89	0.94
ลิ้น	0.67	0.22	0.33
ฟัน	0.44	0.44	0.44
แดงโม	0.57	0.89	0.70
Accuracy			0.65
Macro avg	0.70	0.65	0.65
Weighted avg	0.70	0.65	0.65



ภาพประกอบที่ 43 Confusion Matrix ของ CNN 30 กลุ่มข้อมูล

ตาราง 8 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
แอปเปิ้ล	0.73	0.89	0.80
แขน	0.67	0.44	0.53
กล้วย	0.70	0.78	0.74
นก	1.00	0.56	0.71
ไก่	0.36	0.56	0.43
หวี	0.69	1.00	0.82
จระเข้	0.55	0.67	0.60
หมา	0.43	0.67	0.52
เปิด	1.00	0.44	0.62
หู	0.58	0.78	0.67
ไข่	0.50	0.22	0.31
ข้อศอก	0.60	0.67	0.63
ข้าง	0.55	0.67	0.60
ตา	0.36	0.44	0.40
พัดลม	0.56	0.56	0.56
เท้า	0.70	0.78	0.74
แก้ว	0.67	0.44	0.53
ปืน	1.00	0.67	0.80
มือ	0.00	0.00	0.00
หมวก	0.60	0.33	0.43
สิงโต	1.00	0.89	0.94
มะม่วง	0.89	0.89	0.89
มังคุด	1.00	0.67	0.80
ลิง	0.38	0.56	0.45
ปาก	0.57	0.44	0.50
จมูก	0.75	0.67	0.71
ส้ม	0.50	0.89	0.64
กางเกง	0.43	0.67	0.52
มะละกอ	1.00	0.67	0.80
ดินสอ	0.62	0.56	0.59

สับปะรด	0.75	0.67	0.71
---------	------	------	------

ตาราง 8 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
ชมพู่	0.71	0.56	0.63
เสื่อ	0.50	0.44	0.47
งู	0.38	0.67	0.48
โต๊ะ	1.00	0.22	0.36
เสื่อ	0.40	0.67	0.50
ลิ้น	0.67	0.22	0.33
พิน	0.43	0.33	0.38
แปรงสีฟัน	0.67	0.89	0.76
แตงโม	0.33	0.44	0.38
Accuracy			0.59
Macro avg	0.63	0.59	0.58
Weighted avg	0.63	0.59	0.58



ตาราง 9 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
แอปเปิ้ล	1.00	0.67	0.80
แขน	0.80	0.44	0.57
กล้วย	0.83	0.56	0.67
นก	1.00	0.33	0.50
ไก่	0.57	0.44	0.50
หวี	0.64	1.00	0.78
จระเข้	0.38	0.67	0.48
หมา	0.50	0.44	0.47
เปิด	1.00	0.67	0.80
หู	0.62	0.56	0.59
กิน	0.44	0.44	0.44
ไข่	0.58	0.78	0.67
ข้อศอก	0.75	0.67	0.71
ข้าง	0.29	0.44	0.35
ตา	0.42	0.56	0.48
พัดลม	0.62	0.56	0.59
เท้า	0.50	0.89	0.64
เตะฟุตบอล	0.57	0.44	0.50
แก้ว	0.73	0.89	0.80
ปิ่น	0.67	0.67	0.67
มือ	0.00	0.00	0.00
หมวก	0.43	0.33	0.38
สิงโต	0.86	0.67	0.75
มะม่วง	0.70	0.78	0.74
มังคุด	0.46	0.67	0.55
ลิง	0.36	0.56	0.43
ปาก	0.29	0.44	0.35
จมูก	0.58	0.78	0.67
ส้ม	0.60	0.67	0.63
กางเกง	0.40	0.44	0.42
มะละกอ	0.54	0.78	0.64

ตาราง 9 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
ดินสอ	0.60	0.67	0.63
สับปะรด	0.44	0.44	0.44
ชมพู่	0.50	0.33	0.40
วุ้น	1.00	0.56	0.71
เสื่อ	0.40	0.22	0.29
อาบน้ำ	0.53	0.89	0.67
งู	0.31	0.56	0.40
ยีน	0.83	0.56	0.67
โต๊ะ	0.83	0.56	0.67
เสื่อ	0.36	0.44	0.40
ลิ้น	0.40	0.22	0.29
ฟัน	0.50	0.44	0.47
แปรงสีฟัน	1.00	0.89	0.94
รดน้ำต้นไม้	0.70	0.78	0.74
คูทีวี	0.78	0.78	0.78
เดิน	0.75	0.33	0.46
แต่งโม	0.60	0.33	0.43
เซ็ดตัว	0.70	0.78	0.74
เขียน	1.00	0.56	0.71
Accuracy			0.57
Macro avg	0.61	0.57	0.57
Weighted avg	0.61	0.57	0.57



ตาราง 10 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
แอปเปิ้ล	0.82	1.00	0.90
แขน	0.64	0.78	0.70
กล้วย	0.70	0.78	0.74
นก	0.70	0.78	0.74
ดำ	1.00	0.78	0.88
น้ำเงิน	0.75	0.33	0.46
น้ำตาล	0.69	1.00	0.82
ไก่	0.60	0.67	0.63
หวี	0.73	0.89	0.80
จระเข้	0.58	0.78	0.67
หมา	0.60	0.67	0.63
เปิด	0.83	0.56	0.67
หู	0.62	0.89	0.73
กิน	0.89	0.89	0.89
ไข่	0.88	0.78	0.82
ข้อศอก	1.00	0.56	0.71
ข้าง	0.75	0.33	0.46
ตา	0.57	0.89	0.70
พัดลม	0.67	0.44	0.53
เท้า	1.00	0.11	0.20
เตะฟุตบอล	0.78	0.78	0.78
แก้ว	0.80	0.44	0.57
เชียว	0.50	0.44	0.47
ป็น	0.50	0.67	0.57
มือ	0.00	0.00	0.00
หมวก	1.00	0.67	0.80
สิงโต	0.75	0.67	0.71
มะม่วง	0.54	0.78	0.64
มังคุด	0.62	0.56	0.59
ลิง	0.29	0.56	0.38
ปาก	0.67	0.67	0.67
จมูก	0.45	0.56	0.50
ส้ม	0.50	0.44	0.47
ส้ม (สี)	0.00	0.00	0.00
กางเกง	0.44	0.44	0.44
มะละกอ	0.71	0.56	0.63

## ตาราง 10 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
ดินสอ	1.00	0.89	0.94
สับปะรด	0.75	0.33	0.46
ม่วง	0.00	0.00	0.00
แดง	0.64	0.78	0.70
ชมพู	0.40	0.67	0.50
วุ้น	0.75	0.67	0.71
เสื่อ	0.75	0.33	0.46
อาบน้ำ	0.80	0.89	0.84
ฟ้า	0.50	0.22	0.31
งู	0.50	0.78	0.61
ยีน	0.62	0.56	0.59
โต๊ะ	0.88	0.78	0.82
เสื่อ	0.62	0.89	0.73
ลิ้น	0.67	0.22	0.33
ฟัน	0.60	0.33	0.43
แปรงสีฟัน	0.88	0.78	0.82
รดน้ำต้นไม้	0.88	0.78	0.82
คูทีวี	0.64	1.00	0.78
เดิน	0.44	0.44	0.44
แดงโม	0.60	0.67	0.63
ขาว	0.41	1.00	0.58
เข็ดด้ว	0.75	1.00	0.86
เขียน	0.64	0.78	0.70
เหลือง	0.36	0.56	0.43
Accuracy			0.62
Macro avg	0.64	0.62	0.61
Weighted avg	0.64	0.62	0.61



ตาราง 11 ผลลัพธ์ของวิธีการ CNN กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
อา	0.42	0.56	0.48
แอปเปิ้ล	1.00	0.67	0.80
แขน	0.36	0.44	0.40
กล้วย	0.78	0.78	0.78
นก	0.83	0.56	0.67
ดำ	0.70	0.78	0.74
น้ำเงิน	0.80	0.44	0.57
น้ำตาล	1.00	0.44	0.62
ไก่	0.50	0.11	0.18
หวี	0.53	0.89	0.67
จระเข้	0.33	0.78	0.47
หมา	0.43	0.67	0.52
เปิด	0.83	0.56	0.67
หู	0.58	0.78	0.67
กิน	0.41	0.78	0.54
อี	0.57	0.44	0.50
ไข่	0.47	0.78	0.58
ข้อศอก	1.00	0.67	0.80
ข้าง	0.20	0.11	0.14
ตา	0.44	0.44	0.44
พัดลม	0.33	0.22	0.27
เท้า	0.50	0.11	0.18
เตะฟุตบอล	0.56	0.56	0.56
แก้ว	0.50	0.33	0.40
เจียว	0.71	0.56	0.63
ปิ่น	0.67	0.67	0.67
มือ	0.22	0.22	0.22

ตาราง 11 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
หมวก	1.00	0.67	0.80
สิงโต	0.58	0.78	0.67
มะม่วง	0.70	0.78	0.74
มังคุด	0.78	0.78	0.78
อีม	0.11	0.11	0.11
ลิง	0.36	0.44	0.40
ปาก	0.75	0.33	0.46
จมูก	0.60	0.67	0.63
อู	0.40	0.22	0.29
ส้ม	0.30	0.33	0.32
ส้ม (สี)	0.20	0.11	0.14
กางเกง	0.62	0.56	0.59
มะละกอ	0.80	0.44	0.57
ดินสอ	0.60	0.67	0.63
สับปะรด	1.00	0.44	0.62
ม่วง	0.00	0.00	0.00
แดง	0.53	0.89	0.67
ชมพู	0.54	0.78	0.64
วีน	1.00	0.67	0.80
ช	0.45	1.00	0.62
เสื่อ	0.75	0.33	0.46
อาบน้ำ	0.50	0.67	0.57
ฟ้า	0.29	0.22	0.25
งู	0.67	0.44	0.53
ช	0.56	0.56	0.56
ยีน	0.60	0.33	0.43

ตาราง 11 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
โต๊ะ	1.00	0.67	0.80
เสื่อ	0.50	0.78	0.61
ลึ้น	0.38	0.33	0.35
พื่น	1.00	0.22	0.36
แปรงสีฟัน	0.78	0.78	0.78
รดน้ำต้นไม้	0.54	0.78	0.64
คูทิว	1.00	0.89	0.94
เดิน	0.50	0.33	0.40
แตงโม	0.45	0.56	0.50
ขาว	0.29	1.00	0.45
เช็ดตัว	0.50	0.89	0.64
เขียน	0.75	0.67	0.71
เหลือง	0.33	0.22	0.27
Accuracy			0.54
Macro avg	0.58	0.54	0.53
Weighted avg	0.58	0.54	0.53



Confusion Matrix visualization showing predicted vs actual classes for a CNN model. The matrix is a grid of small colored cells representing counts for each class combination. The classes listed on the left include: none, none1, none2, none3, none4, none5, none6, none7, none8, none9, none10, none11, none12, none13, none14, none15, none16, none17, none18, none19, none20, none21, none22, none23, none24, none25, none26, none27, none28, none29, none30, none31, none32, none33, none34, none35, none36, none37, none38, none39, none40, none41, none42, none43, none44, none45, none46, none47, none48, none49, none50, none51, none52, none53, none54, none55, none56, none57, none58, none59, none60, none61, none62, none63, none64, none65, none66, none67, none68, none69, none70, none71, none72, none73, none74, none75, none76, none77, none78, none79, none80, none81, none82, none83, none84, none85, none86, none87, none88, none89, none90, none91, none92, none93, none94, none95, none96, none97, none98, none99, none100, none101, none102, none103, none104, none105, none106, none107, none108, none109, none110, none111, none112, none113, none114, none115, none116, none117, none118, none119, none120, none121, none122, none123, none124, none125, none126, none127, none128, none129, none130, none131, none132, none133, none134, none135, none136, none137, none138, none139, none140, none141, none142, none143, none144, none145, none146, none147, none148, none149, none150, none151, none152, none153, none154, none155, none156, none157, none158, none159, none160, none161, none162, none163, none164, none165, none166, none167, none168, none169, none170, none171, none172, none173, none174, none175, none176, none177, none178, none179, none180, none181, none182, none183, none184, none185, none186, none187, none188, none189, none190, none191, none192, none193, none194, none195, none196, none197, none198, none199, none200, none201, none202, none203, none204, none205, none206, none207, none208, none209, none210, none211, none212, none213, none214, none215, none216, none217, none218, none219, none220, none221, none222, none223, none224, none225, none226, none227, none228, none229, none230, none231, none232, none233, none234, none235, none236, none237, none238, none239, none240, none241, none242, none243, none244, none245, none246, none247, none248, none249, none250, none251, none252, none253, none254, none255, none256, none257, none258, none259, none260, none261, none262, none263, none264, none265, none266, none267, none268, none269, none270, none271, none272, none273, none274, none275, none276, none277, none278, none279, none280, none281, none282, none283, none284, none285, none286, none287, none288, none289, none290, none291, none292, none293, none294, none295, none296, none297, none298, none299, none300, none301, none302, none303, none304, none305, none306, none307, none308, none309, none310, none311, none312, none313, none314, none315, none316, none317, none318, none319, none320, none321, none322, none323, none324, none325, none326, none327, none328, none329, none330, none331, none332, none333, none334, none335, none336, none337, none338, none339, none340, none341, none342, none343, none344, none345, none346, none347, none348, none349, none350, none351, none352, none353, none354, none355, none356, none357, none358, none359, none360, none361, none362, none363, none364, none365, none366, none367, none368, none369, none370, none371, none372, none373, none374, none375, none376, none377, none378, none379, none380, none381, none382, none383, none384, none385, none386, none387, none388, none389, none390, none391, none392, none393, none394, none395, none396, none397, none398, none399, none400, none401, none402, none403, none404, none405, none406, none407, none408, none409, none410, none411, none412, none413, none414, none415, none416, none417, none418, none419, none420, none421, none422, none423, none424, none425, none426, none427, none428, none429, none430, none431, none432, none433, none434, none435, none436, none437, none438, none439, none440, none441, none442, none443, none444, none445, none446, none447, none448, none449, none450, none451, none452, none453, none454, none455, none456, none457, none458, none459, none460, none461, none462, none463, none464, none465, none466, none467, none468, none469, none470, none471, none472, none473, none474, none475, none476, none477, none478, none479, none480, none481, none482, none483, none484, none485, none486, none487, none488, none489, none490, none491, none492, none493, none494, none495, none496, none497, none498, none499, none500, none501, none502, none503, none504, none505, none506, none507, none508, none509, none510, none511, none512, none513, none514, none515, none516, none517, none518, none519, none520, none521, none522, none523, none524, none525, none526, none527, none528, none529, none530, none531, none532, none533, none534, none535, none536, none537, none538, none539, none540, none541, none542, none543, none544, none545, none546, none547, none548, none549, none550, none551, none552, none553, none554, none555, none556, none557, none558, none559, none560, none561, none562, none563, none564, none565, none566, none567, none568, none569, none570, none571, none572, none573, none574, none575, none576, none577, none578, none579, none580, none581, none582, none583, none584, none585, none586, none587, none588, none589, none590, none591, none592, none593, none594, none595, none596, none597, none598, none599, none600, none601, none602, none603, none604, none605, none606, none607, none608, none609, none610, none611, none612, none613, none614, none615, none616, none617, none618, none619, none620, none621, none622, none623, none624, none625, none626, none627, none628, none629, none630, none631, none632, none633, none634, none635, none636, none637, none638, none639, none640, none641, none642, none643, none644, none645, none646, none647, none648, none649, none650, none651, none652, none653, none654, none655, none656, none657, none658, none659, none660, none661, none662, none663, none664, none665, none666, none667, none668, none669, none670, none671, none672, none673, none674, none675, none676, none677, none678, none679, none680, none681, none682, none683, none684, none685, none686, none687, none688, none689, none690, none691, none692, none693, none694, none695, none696, none697, none698, none699, none700, none701, none702, none703, none704, none705, none706, none707, none708, none709, none710, none711, none712, none713, none714, none715, none716, none717, none718, none719, none720, none721, none722, none723, none724, none725, none726, none727, none728, none729, none730, none731, none732, none733, none734, none735, none736, none737, none738, none739, none740, none741, none742, none743, none744, none745, none746, none747, none748, none749, none750, none751, none752, none753, none754, none755, none756, none757, none758, none759, none760, none761, none762, none763, none764, none765, none766, none767, none768, none769, none770, none771, none772, none773, none774, none775, none776, none777, none778, none779, none780, none781, none782, none783, none784, none785, none786, none787, none788, none789, none790, none791, none792, none793, none794, none795, none796, none797, none798, none799, none800, none801, none802, none803, none804, none805, none806, none807, none808, none809, none810, none811, none812, none813, none814, none815, none816, none817, none818, none819, none820, none821, none822, none823, none824, none825, none826, none827, none828, none829, none830, none831, none832, none833, none834, none835, none836, none837, none838, none839, none840, none841, none842, none843, none844, none845, none846, none847, none848, none849, none850, none851, none852, none853, none854, none855, none856, none857, none858, none859, none860, none861, none862, none863, none864, none865, none866, none867, none868, none869, none870, none871, none872, none873, none874, none875, none876, none877, none878, none879, none880, none881, none882, none883, none884, none885, none886, none887, none888, none889, none890, none891, none892, none893, none894, none895, none896, none897, none898, none899, none900, none901, none902, none903, none904, none905, none906, none907, none908, none909, none910, none911, none912, none913, none914, none915, none916, none917, none918, none919, none920, none921, none922, none923, none924, none925, none926, none927, none928, none929, none930, none931, none932, none933, none934, none935, none936, none937, none938, none939, none940, none941, none942, none943, none944, none945, none946, none947, none948, none949, none950, none951, none952, none953, none954, none955, none956, none957, none958, none959, none960, none961, none962, none963, none964, none965, none966, none967, none968, none969, none970, none971, none972, none973, none974, none975, none976, none977, none978, none979, none980, none981, none982, none983, none984, none985, none986, none987, none988, none989, none990, none991, none992, none993, none994, none995, none996, none997, none998, none999, 1000. The matrix shows a high density of 'none' predictions, with some scattered values for other classes. The x-axis is labeled 'Predicted class' and the y-axis is labeled 'Actual class'.

ภาพประกอบ 47 Confusion Matrix ของ CNN 66 กลุ่มข้อมูล

ตาราง 12 ผลลัพธ์ Accuracy ของวิธีการ CNN กับฐานข้อมูลทั้งหมด

ประเภทของข้อมูล	Train (%)	Validation (%)	Test (%)
10 กลุ่มข้อมูล	97.62	83.33	77.78
20 กลุ่มข้อมูล	99.88	81.67	73.89
30 กลุ่มข้อมูล	93.97	64.07	65.19
40 กลุ่มข้อมูล	100	64.44	58.89
50 กลุ่มข้อมูล	99.67	58.89	57.11
60 กลุ่มข้อมูล	100	65.19	62.41
66 กลุ่มข้อมูล	96.47	57.41	54.04

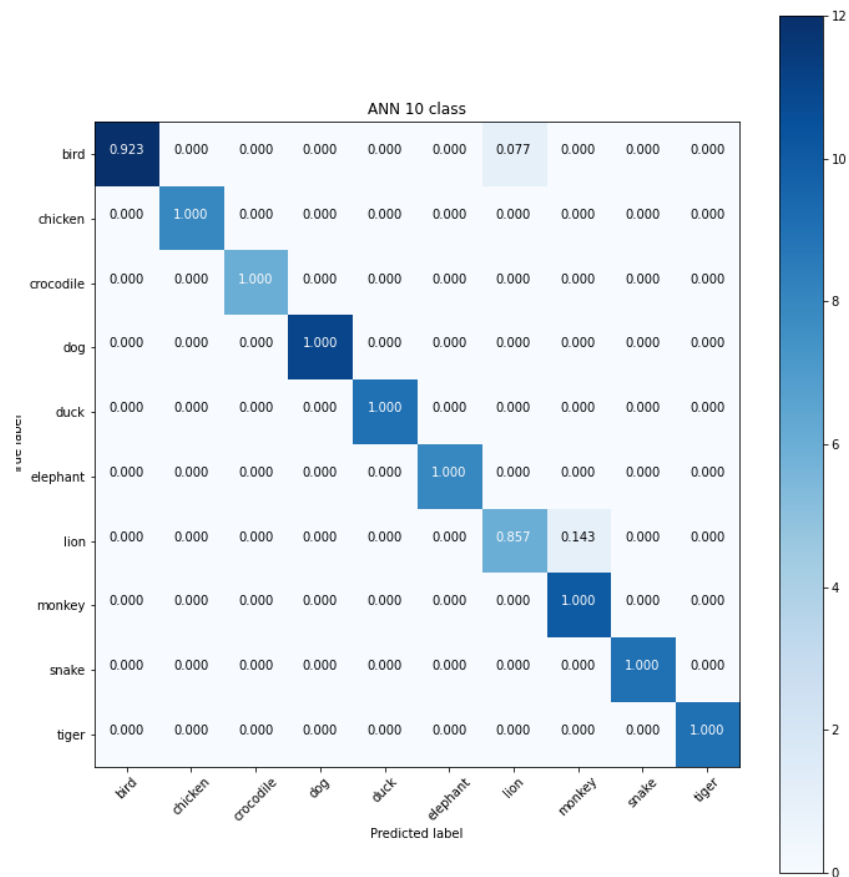
จากตาราง 12 จะเห็นได้ว่าวิธีการ CNN มีความเหมาะสมกับฐานข้อมูลแบบ 10 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy สูงที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 77.78% และมีความไม่เหมาะสมกับฐานข้อมูลแบบ 66 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy น้อยที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 54.04%

#### 4.4 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ ANN

ในขั้นตอนการ train model ของ ANN มีการใช้ Loss function แบบ Categorical crossentropy, Optimizer แบบ Adam และ epochs จำนวน 300 หลังจากได้ model มาแล้ว จะนำ model ที่ได้ไปทดสอบกับ test data ในการประเมินประสิทธิภาพของ model จะใช้ Confusion Matrix ในการหาค่า Accuracy, Sensitivity, F1-score และ Precision ของ model ซึ่งได้ผลลัพธ์ดังตารางต่อไปนี้

ตาราง 13 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	1.00	0.92	0.96
ไก่	1.00	1.00	1.00
จระเข้	1.00	1.00	1.00
หมา	1.00	1.00	1.00
เปิด	1.00	1.00	1.00
ช้าง	1.00	1.00	1.00
สิงโต	0.86	0.86	0.86
ลิง	0.91	1.00	0.95
งู	1.00	1.00	1.00
เสื่อ	1.00	1.00	1.00
Accuracy			0.98
Macro avg	0.98	0.98	0.98
Weighted avg	0.98	0.98	0.98



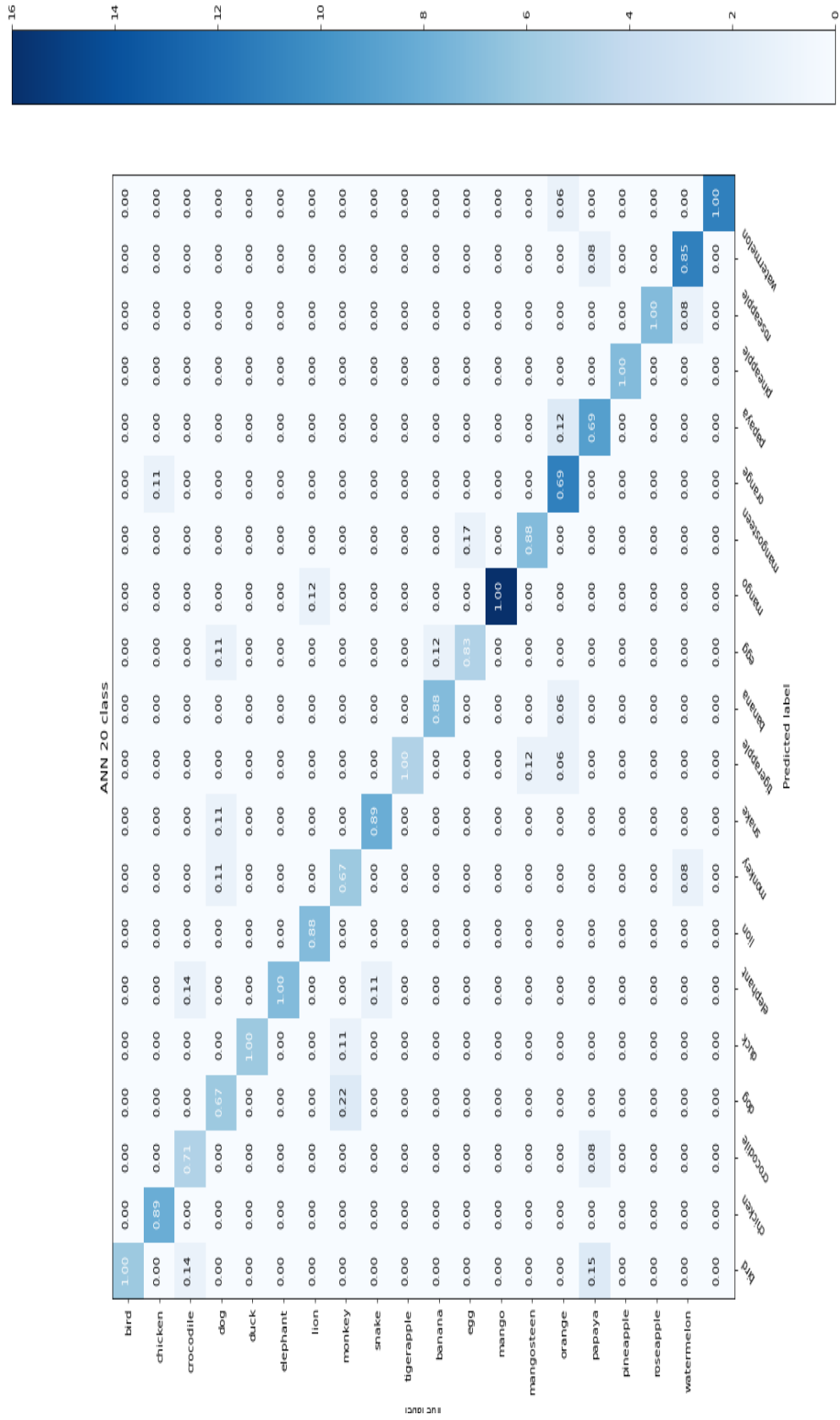
ภาพประกอบ 48 Confusion Matrix ของ ANN 10 กลุ่มข้อมูล

ตาราง 14 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.67	1.00	0.80
ไก่	1.00	0.89	0.94
จระเข้	0.83	0.71	0.77
หมา	0.75	0.67	0.71
เป็ด	0.86	1.00	0.92
ช้าง	0.78	1.00	0.88
สิงโต	1.00	0.88	0.93
ลิง	0.75	0.67	0.71
งู	0.89	0.89	0.89

ตาราง 14 (ต่อ)

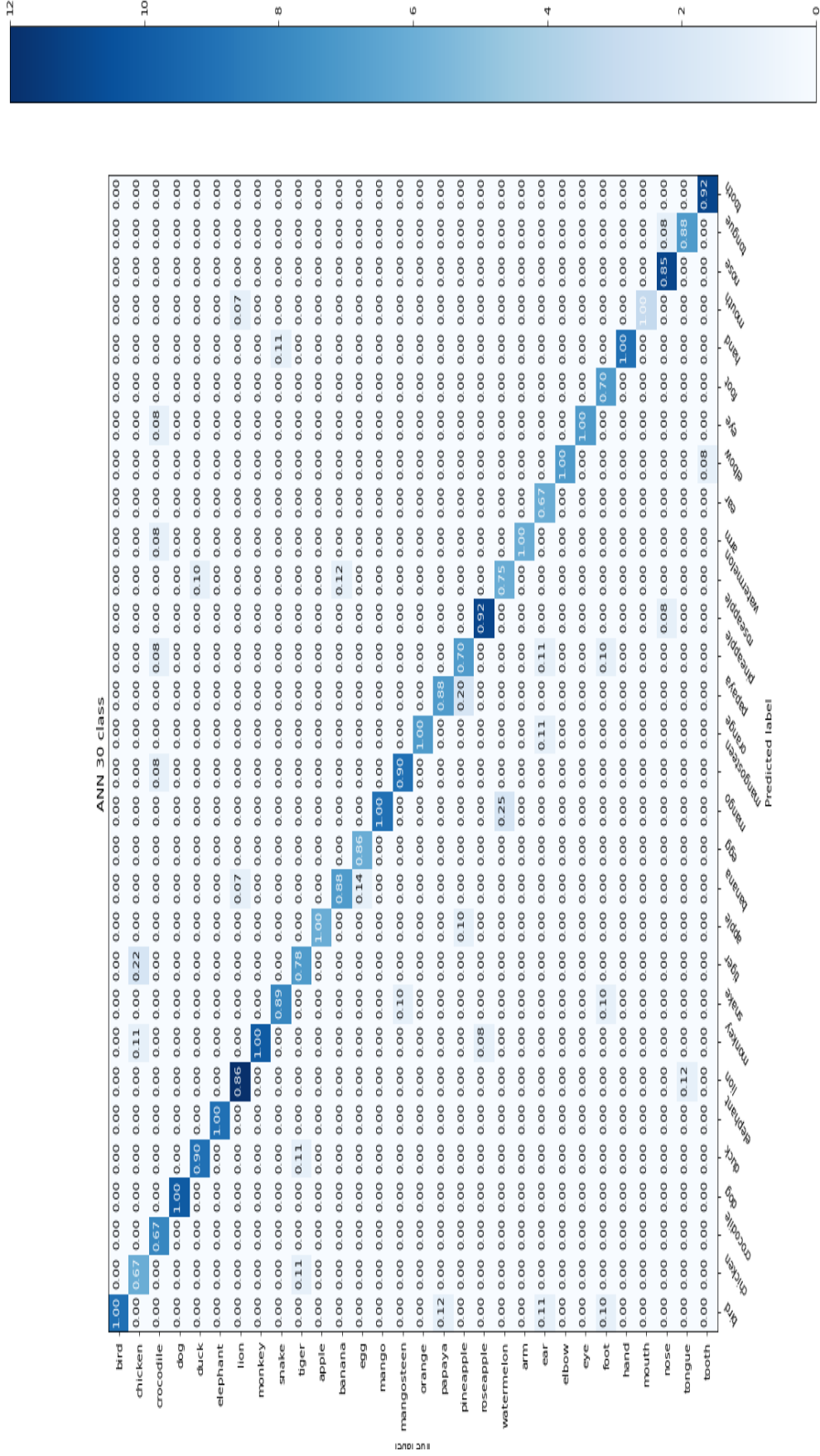
ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
เลือด	0.71	1.00	0.83
แอปเปิ้ล	0.88	0.88	0.88
กล้วย	0.71	0.83	0.77
ไข่	0.94	1.00	0.97
มะม่วง	0.88	0.88	0.88
มังคุด	0.92	0.69	0.79
ส้ม	0.82	0.69	0.75
มะละกอ	1.00	1.00	1.00
สับปะรด	0.88	1.00	0.93
ชมพู	0.92	0.85	0.88
แตงโม	0.92	1.00	0.96
<b>Accuracy</b>			0.86
<b>Macro avg</b>	0.85	0.88	0.86
<b>Weighted avg</b>	0.87	0.86	0.86



ภาพประกอบที่ 49 Confusion Matrix ของ ANN 20 กลุ่มข้อมูล

ตาราง 15 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.75	1.00	0.86
ไก่	0.86	0.67	0.75
จระเข้	1.00	0.67	0.80
หมา	1.00	1.00	1.00
เปิด	0.90	0.90	0.90
ข้าง	1.00	1.00	1.00
สิงโต	0.92	0.86	0.89
ลิง	0.83	1.00	0.91
งู	0.80	0.89	0.84
เสือ	0.78	0.78	0.78
แอปเปิ้ล	0.86	1.00	0.92
กล้วย	0.78	0.88	0.82
ไข่	1.00	0.86	0.92
มะม่วง	0.82	1.00	0.90
มังคุด	0.90	0.90	0.90
ส้ม	0.88	1.00	0.93
มะละกอ	0.78	0.88	0.82
สับปะรด	0.70	0.70	0.70
ชมพู่	0.92	0.92	0.92
แตงโม	0.75	0.75	0.75
ขนุน	0.86	1.00	0.92
ทุ	1.00	0.67	0.80
ข้อศอก	0.88	1.00	0.93
ตา	0.88	1.00	0.93
เท้า	1.00	0.70	0.82
มือ	0.90	1.00	0.95
ปาก	0.75	1.00	0.86
จมูก	1.00	0.85	0.92
ลิ้น	0.88	0.88	0.88
ฟัน	1.00	0.92	0.96
Accuracy			0.88
Macro avg	0.88	0.89	0.88
Weighted avg	0.89	0.88	0.88



ภาพประกอบ 50 Confusion Matrix ของ ANN 30 กลุ่มข้อมูล

ตาราง 16 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	1.00	0.78	0.88
ไก่	0.83	0.83	0.83
จระเข้	0.80	0.67	0.73
หมา	0.91	0.91	0.91
เปิด	0.50	0.90	0.64
ข้าง	1.00	1.00	1.00
สิงโต	1.00	0.82	0.90
ลิง	0.90	0.82	0.86
งู	0.88	0.78	0.82
เสือ	0.71	0.83	0.77
แอปเปิ้ล	0.90	0.69	0.78
กล้วย	0.67	1.00	0.80
ไข่	0.33	0.25	0.29
มะม่วง	0.70	0.70	0.70
มังคุด	1.00	0.76	0.87
ส้ม	1.00	0.73	0.85
มะละกอ	0.70	1.00	0.82
สับปะรด	0.83	0.71	0.77
ชมพู่	0.83	0.83	0.83
แตงโม	0.69	0.82	0.75
ขน	0.90	0.90	0.90
หู	0.75	0.86	0.80
ข้อศอก	0.83	0.71	0.77
ตา	0.58	0.78	0.67
เท้า	0.67	1.00	0.80
มือ	0.64	1.00	0.78
ปาก	0.83	0.71	0.77
จมูก	0.62	0.83	0.71
ลิ้น	0.86	0.86	0.86
ฟัน	0.67	0.57	0.62
หัว	1.00	0.73	0.84

ตาราง 16 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัดลม	0.77	0.91	0.83
แก้ว	0.86	0.86	0.86
ปืน	1.00	0.86	0.92
หมวก	0.86	0.86	0.86
กางเกง	0.80	1.00	0.89
ดินสอ	0.40	0.29	0.33
เสื้อ	1.00	0.64	0.78
โต๊ะ	1.00	0.83	0.91
แปรงสีฟัน	0.90	0.90	0.90
Accuracy			0.80
Macro avg	0.80	0.80	0.79
Weighted avg	0.83	0.80	0.80



ตาราง 17 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	1.00	0.90	0.95
ไก่	1.00	1.00	1.00
จระเข้	1.00	0.73	0.84
หมา	0.73	0.89	0.80
เปิด	0.50	0.50	0.50
ข้าง	0.64	1.00	0.78
สิงโต	0.90	0.90	0.90
ลิง	0.67	0.75	0.71
งู	1.00	1.00	1.00
เสือ	0.75	0.82	0.78
แอปเปิ้ล	1.00	0.33	0.50
กล้วย	0.80	0.50	0.62
ไข่	0.83	0.83	0.83
มะม่วง	0.75	0.67	0.71
มังคุด	0.50	0.83	0.62
ส้ม	0.89	1.00	0.94
มะละกอ	1.00	0.91	0.95
สับปะรด	0.78	0.54	0.64
ชมพู่	0.73	0.89	0.80
แตงโม	0.86	0.75	0.80
เชน	0.83	0.56	0.67
หู	0.67	0.44	0.53
ข้อศอก	0.89	0.89	0.89
ตา	0.73	0.89	0.80
เท้า	0.88	0.88	0.88
มือ	0.75	0.90	0.82
ปาก	1.00	1.00	1.00
จมูก	0.73	1.00	0.84
ลิ้น	0.67	0.50	0.57
ฟัน	0.75	0.86	0.80
หัว	0.64	0.90	0.75

ตาราง 17 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัดลม	0.80	0.57	0.67
แก้ว	0.93	0.87	0.90
ปืน	0.77	0.91	0.83
หมวก	0.75	0.90	0.82
กางเกง	0.56	0.62	0.59
ดินสอ	1.00	0.82	0.90
เสื้อ	1.00	0.75	0.86
โต๊ะ	1.00	0.80	0.89
แปรงสีฟัน	1.00	0.92	0.96
กิน	0.78	0.78	0.78
เตะฟุตบอล	1.00	0.58	0.74
วิ่ง	0.73	0.73	0.73
อาบน้ำ	0.80	0.80	0.80
ยืน	1.00	1.00	1.00
รดน้ำต้นไม้	0.79	1.00	0.88
ดูทีวี	0.75	0.86	0.80
วิ่ง	0.58	0.78	0.67
เช็ดตัว	0.87	0.93	0.90
เขียน	0.88	0.88	0.88
Accuracy			0.80
Macro avg	0.82	0.80	0.80
Weighted avg	0.82	0.80	0.80

ANN 50 Confusion Matrix

	bird	chicken	chickadee	crow	duck	flamingo	heron	kingfisher	magpie	parrot	pelican	phoebe	quail	screech owl	shrike	tiger	trout	warbler	wren	yellowthroat
bird	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
chicken	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
chickadee	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
crow	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
duck	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
flamingo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
heron	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
kingfisher	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
magpie	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
parrot	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
pelican	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
phoebe	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
quail	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
screech owl	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
shrike	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
tiger	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
trout	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
warbler	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
wren	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
yellowthroat	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

ภาพประกอบ 52 Confusion Matrix ของ ANN 50 กลุ่มข้อมูล

ตาราง 18 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.91	0.91	0.91
ไก่	0.90	0.75	0.82
จระเข้	0.78	0.64	0.70
หมา	0.92	0.92	0.92
เป็ด	1.00	0.80	0.89
ช้าง	0.90	1.00	0.95
สิงโต	0.75	0.75	0.75
ลิง	0.62	0.71	0.67
งู	0.78	1.00	0.88
เสือดอ	0.89	0.80	0.84
แอปเปิ้ล	1.00	0.67	0.80
กล้วย	0.80	0.57	0.67
ไข่	0.78	0.78	0.78
มะม่วง	0.88	0.78	0.82
มังคุด	0.56	1.00	0.71
ส้ม	0.59	1.00	0.74
มะละกอ	0.90	0.75	0.82
สับปะรด	0.55	1.00	0.71
ชมพู่	0.75	0.75	0.75
แตงโม	0.62	0.62	0.62
เชน	0.88	0.64	0.74
หูก	1.00	0.83	0.91
ข้อศอก	0.88	0.88	0.88
ตา	1.00	0.89	0.94
เท้า	0.83	0.71	0.77
มือ	0.67	1.00	0.80
ปาก	1.00	0.78	0.88
จมูก	0.40	0.40	0.40
ลิ้น	0.80	0.80	0.80
ฟัน	0.67	1.00	0.80
หวี	1.00	0.90	0.95
พัดลม	0.67	0.80	0.73
แก้ว	0.60	0.27	0.37
ปิ่น	0.50	0.83	0.62
หมวก	0.86	0.86	0.86
กางเกง	0.83	0.56	0.67

ตาราง 18 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
ดินสอ	0.82	0.90	0.86
เสื่อ	0.70	0.64	0.67
โต๊ะ	0.67	0.80	0.73
แปรงสีฟัน	0.78	1.00	0.88
กิน	1.00	1.00	1.00
เตะฟุตบอล	0.80	1.00	0.89
วิ่ง	0.67	0.86	0.75
อาบน้ำ	0.60	0.75	0.67
ยืน	0.60	0.33	0.43
รดน้ำต้นไม้	1.00	0.86	0.92
ดูทีวี	1.00	0.67	0.80
วิ่ง	1.00	0.88	0.93
เช็ดตัว	0.83	0.56	0.67
เขียน	1.00	0.70	0.82
ดำ	0.82	0.90	0.86
น้ำเงิน	0.89	0.57	0.70
น้ำตาล	0.78	0.88	0.82
เขี้ยว	0.93	0.87	0.90
ส้ม (สี)	0.91	1.00	0.95
ม่วง	0.58	0.88	0.70
แดง	0.67	1.00	0.80
ฟ้า	0.78	1.00	0.88
ขาว	1.00	0.75	0.86
เหลือง	0.92	0.92	0.92
Accuracy			0.80
Macro avg	0.80	0.80	0.79
Weighted avg	0.82	0.80	0.80



ตาราง 19 ผลลัพธ์ของวิธีการ ANN กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.78	1.00	0.88
ไก่	0.80	0.89	0.84
จระเข้	0.67	1.00	0.80
หมา	0.62	1.00	0.77
เปิด	0.89	1.00	0.94
ข้าง	1.00	0.88	0.93
สิงโต	0.90	0.82	0.86
ลิง	0.64	1.00	0.78
งู	0.91	1.00	0.95
เสื่อ	0.75	0.67	0.71
แอปเปิ้ล	1.00	0.75	0.86
กล้วย	1.00	0.88	0.93
ไข่	0.83	1.00	0.91
มะม่วง	0.88	0.78	0.82
มังคุด	0.70	0.78	0.74
ส้ม	0.92	1.00	0.96
มะละกอด	1.00	1.00	1.00
สับปะรด	1.00	0.91	0.95
ชมพู่	0.83	0.71	0.77
แตงโม	0.60	1.00	0.75
ขน	0.90	0.60	0.72
หู	0.67	0.60	0.63
ข้อศอก	0.88	0.78	0.82
ตา	0.67	0.86	0.75
เท้า	0.91	0.91	0.91
มือ	0.80	0.62	0.70
ปาก	0.60	0.75	0.67
จมูก	1.00	0.30	0.46
ลิ้น	0.80	0.80	0.80
ฟัน	0.67	0.57	0.62
หัว	0.67	0.89	0.76

ตาราง 19 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัคดลม	1.00	0.77	0.87
แก้ว	0.88	1.00	0.93
ปิ่น	0.83	0.71	0.77
หมวก	0.80	0.73	0.76
กางเกง	0.33	0.33	0.33
ดินสอ	0.00	0.00	0.00
เสื่อ	0.69	0.75	0.72
โต๊ะ	0.82	1.00	0.90
แปรงสีฟัน	0.67	0.67	0.67
กิน	0.86	0.60	0.71
เตะฟุตบอล	0.67	0.80	0.73
วิ่ง	0.67	1.00	0.80
อาบน้ำ	0.86	0.86	0.86
ยืน	0.45	0.62	0.53
รดน้ำต้นไม้	0.83	1.00	0.91
ดูทีวี	0.83	1.00	0.91
วิ่ง	1.00	0.92	0.96
เช็ดตัว	1.00	0.64	0.78
เขียน	0.73	0.67	0.70
ดำ	0.78	0.88	0.82
น้ำเงิน	1.00	0.92	0.96
น้ำตาล	0.89	1.00	0.94
เขียว	0.90	0.90	0.90
ส้ม (สี)	0.57	0.80	0.67
ม่วง	0.88	0.88	0.88
แดง	0.73	0.67	0.70
ฟ้า	0.86	0.55	0.67
ขาว	0.79	0.85	0.81
เหลือง	0.81	0.93	0.87
อา	0.67	0.80	0.73

ตาราง 19 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
อี	0.89	1.00	0.94
อีมี	0.75	0.75	0.75
อู	0.67	0.40	0.50
ช	0.89	0.67	0.76
ช	0.75	0.75	0.75
Accuracy			0.80
Macro avg	0.79	0.80	0.78
Weighted avg	0.81	0.80	0.79



ตาราง 20 ผลลัพธ์ Accuracy ของวิธีการ ANN กับฐานข้อมูลทั้งหมด

ประเภทของข้อมูล	Train (%)	Validation (%)	Test (%)
10 กลุ่มข้อมูล	100	91.22	97.78
20 กลุ่มข้อมูล	100	85.56	86.11
30 กลุ่มข้อมูล	100	86.67	87.78
40 กลุ่มข้อมูล	100	78.61	80
50 กลุ่มข้อมูล	100	78.67	80.44
60 กลุ่มข้อมูล	100	74.26	79.81
66 กลุ่มข้อมูล	100	79.46	79.63

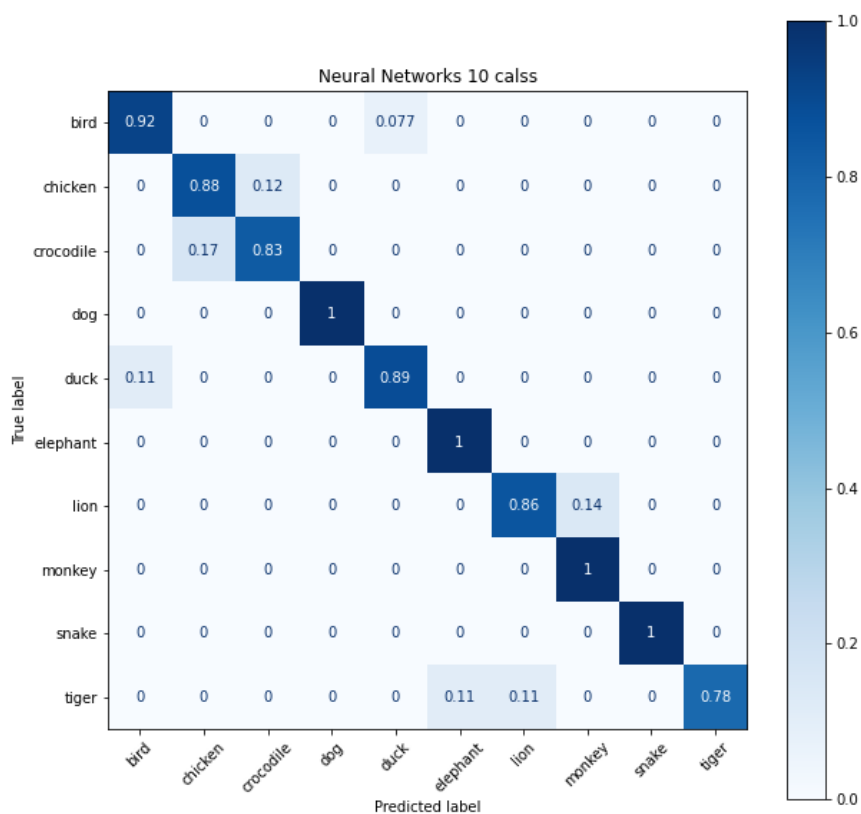
จากตาราง 20 จะเห็นได้ว่าวิธีการ ANN มีความเหมาะสมกับฐานข้อมูลแบบ 10 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy สูงที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 97.78% และมีความไม่เหมาะสมกับฐานข้อมูลแบบ 66 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy น้อยที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 79.63%

#### 4.5 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ Neural Network

ในขั้นตอนการ train model ของ Neural Network มีการใช้ Optimizer แบบ Adam และ epochs จำนวน 200 หลังจากได้ model มาแล้ว จะนำ model ที่ได้ไปทดสอบกับ test data ในการประเมินประสิทธิภาพของ model จะใช้ Confusion Matrix ในการหาค่า Accuracy, Sensitivity, F1-score และ Precision ของ model ซึ่งได้ผลลัพธ์ดังตารางต่อไปนี้

ตาราง 21 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.92	0.92	0.92
ไก่	0.88	0.88	0.88
จระเข้	0.83	0.83	0.83
หมา	1.00	1.00	1.00
เปิด	0.89	0.89	0.89
ข้าง	0.89	1.00	0.94
สิงโต	0.86	0.86	0.86
ลิง	0.91	1.00	0.95
งู	1.00	1.00	1.00
เสือ	1.00	0.78	0.88
Accuracy			0.92
Macro avg	0.92	0.92	0.91
Weighted avg	0.92	0.92	0.92



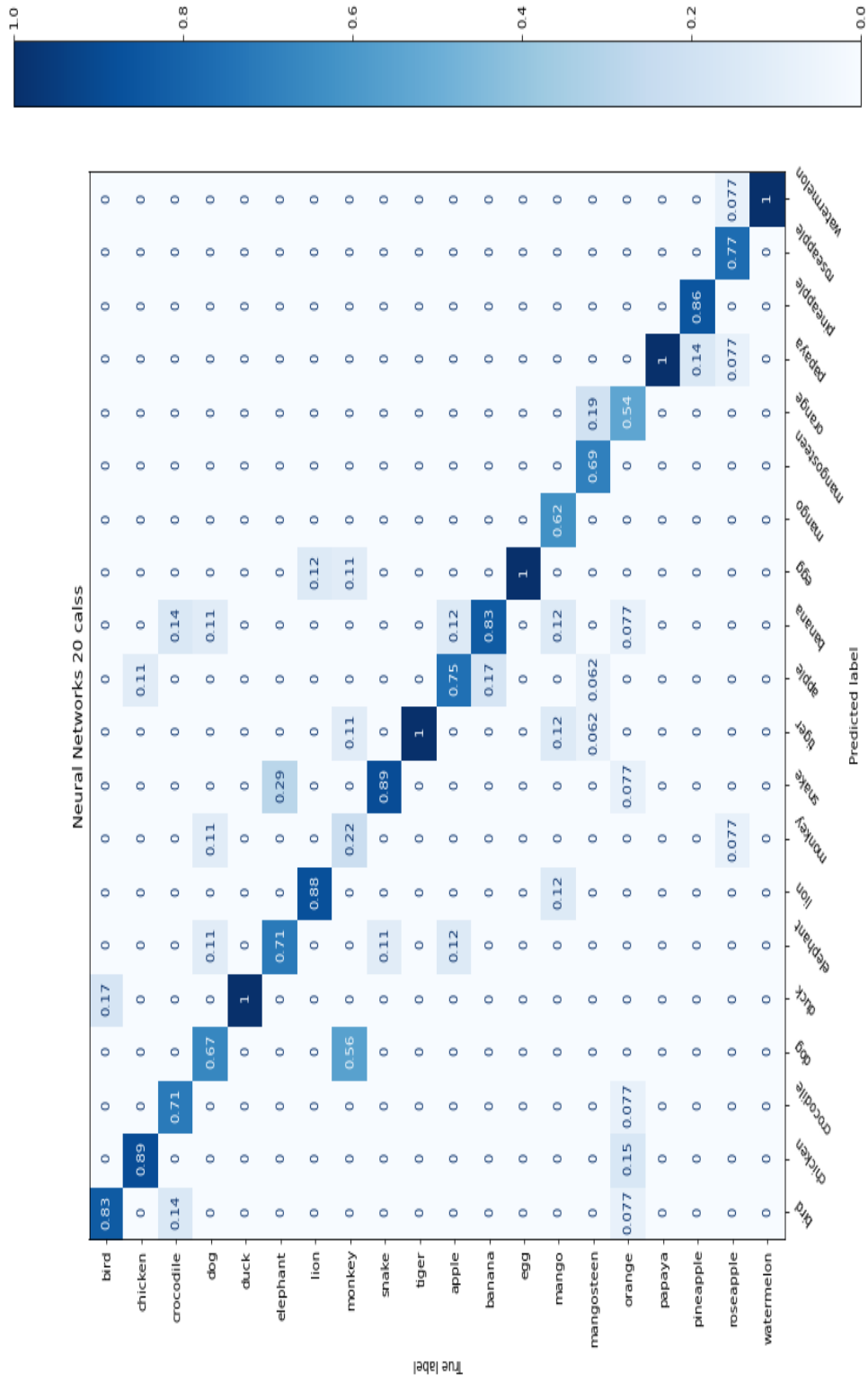
ภาพประกอบ 55 Confusion Matrix ของ Neural Network 10 กลุ่มข้อมูล

ตาราง 22 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.71	0.83	0.77
ไก่	0.80	0.89	0.84
จระเข้	0.83	0.71	0.77
หมา	0.55	0.67	0.60
เป็ด	0.86	1.00	0.92
ช้าง	0.62	0.71	0.67
สิงโต	0.88	0.88	0.88
ลิง	0.50	0.22	0.31
งู	0.73	0.89	0.80
เสือ	0.62	1.00	0.77

ตาราง 22 (ต่อ)

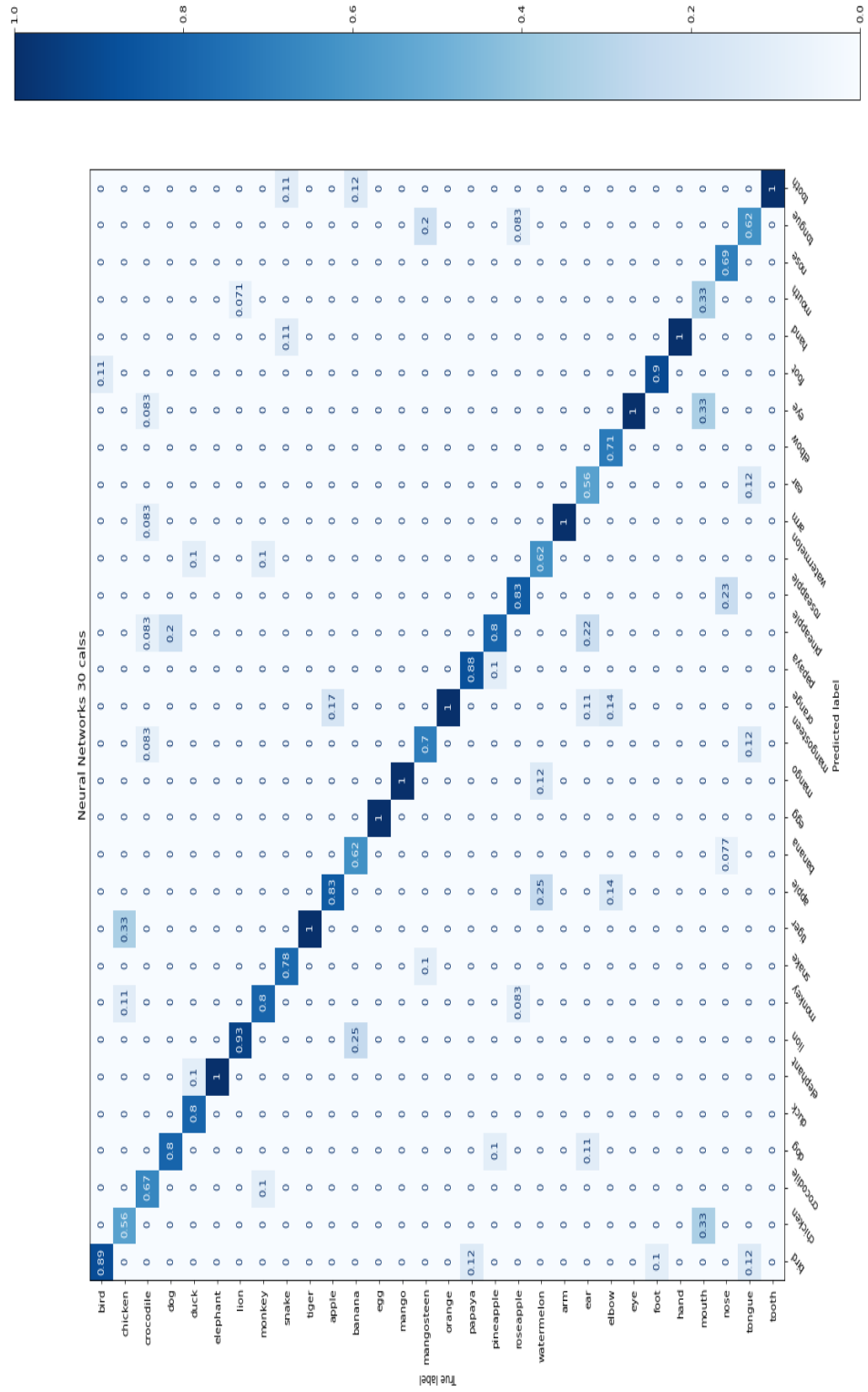
ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
แอปเปิ้ล	0.67	0.75	0.71
กล้วย	0.50	0.83	0.62
ไข่	0.89	1.00	0.94
มะม่วง	1.00	0.62	0.77
มังคุด	1.00	0.69	0.81
ส้ม	0.70	0.54	0.61
มะละกอ	0.78	1.00	0.88
สับปะรด	1.00	0.86	0.92
ชมพู่	1.00	0.77	0.87
แตงโม	0.92	1.00	0.96
<b>Accuracy</b>			0.78
<b>Macro avg</b>	0.78	0.79	0.77
<b>Weighted avg</b>	0.80	0.78	0.78



ภาพประกอบ 56 Confusion Matrix ของ Neural Network 20 กลุ่มข้อมูล

ตาราง 23 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.73	0.89	0.80
ไก่	0.83	0.56	0.67
จระเข้	0.89	0.67	0.76
หมา	0.80	0.80	0.80
เปิด	1.00	0.80	0.89
ข้าง	0.90	1.00	0.95
สิงโต	0.87	0.93	0.90
ลิง	0.80	0.80	0.80
งู	0.88	0.78	0.82
เสื่อ	0.75	1.00	0.86
แอปเปิ้ล	0.62	0.83	0.71
กล้วย	0.83	0.62	0.71
ไข่	1.00	1.00	1.00
มะม่วง	0.90	1.00	0.95
มังคุด	0.78	0.70	0.74
ส้ม	0.70	1.00	0.82
มะละกอ	0.88	0.88	0.88
สับปะรด	0.62	0.80	0.70
ชมพู่	0.77	0.83	0.80
แตงโม	0.71	0.62	0.67
เชน	0.86	1.00	0.92
หูก	0.83	0.56	0.67
ข้อศอก	1.00	0.71	0.83
ตา	0.78	1.00	0.88
เท้า	0.90	0.90	0.90
มือ	0.90	1.00	0.95
ปาก	0.50	0.33	0.40
จมูก	1.00	0.69	0.82
ลิ้น	0.62	0.62	0.62
ฟัน	0.86	1.00	0.92
Accuracy			0.82
Macro avg	0.82	0.81	0.80
Weighted avg	0.83	0.82	0.81



ภาพประกอบที่ 57 Confusion Matrix ของ Neural Network 30 กลุ่มข้อมูล

ตาราง 24 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.78	0.78	0.78
ไก่	0.88	0.70	0.78
จระเข้	0.56	0.83	0.67
หมา	0.90	1.00	0.95
เปิด	0.78	0.64	0.70
ข้าง	0.92	0.92	0.92
สิงโต	1.00	1.00	1.00
ลิง	0.71	0.71	0.71
งู	0.50	0.60	0.55
เสื่อ	0.75	0.64	0.69
แอปเปิ้ล	0.67	0.33	0.44
กล้วย	0.78	1.00	0.88
ไข่	0.78	0.64	0.70
มะม่วง	0.86	1.00	0.92
มังคุด	0.44	0.67	0.53
ส้ม	0.88	0.78	0.82
มะละกอ	0.88	0.88	0.88
สับปะรด	0.88	0.70	0.78
ชมพู่	0.83	0.56	0.67
แตงโม	0.42	0.56	0.48
ขน	0.82	1.00	0.90
หู	0.70	0.64	0.67
ข้อศอก	0.89	0.89	0.89
ตา	0.44	0.57	0.50
เท้า	0.71	0.83	0.77
มือ	0.83	0.77	0.80
ปาก	0.73	0.80	0.76

ตาราง 24 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
จุ่มก	0.60	1.00	0.75
ลึ้น	1.00	0.83	0.91
พึ้น	0.57	0.50	0.53
หวี	0.80	0.80	0.80
พัดลม	0.75	0.75	0.75
แก้ว	0.40	1.00	0.57
ปึ้น	0.80	0.71	0.75
หมวก	0.71	0.62	0.67
กางเกง	0.75	0.55	0.63
ดินสอ	0.73	0.53	0.62
เสื่อ	0.80	0.67	0.73
โต๊ะ	0.83	1.00	0.91
แปรงสีฟัน	0.82	0.90	0.86
Accuracy			0.74
Macro avg	0.75	0.76	0.74
Weighted avg	0.76	0.74	0.74



ตาราง 25 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.82	0.90	0.86
ไก่	0.57	0.80	0.67
จระเข้	0.50	0.45	0.48
หมา	0.89	0.89	0.89
เปิด	0.50	0.38	0.43
ข้าง	0.70	1.00	0.82
สิงโต	0.80	0.80	0.80
ลิง	0.38	0.38	0.38
งู	0.86	0.86	0.86
เสือ	0.69	0.82	0.75
แอปเปิ้ล	0.80	0.67	0.73
กล้วย	0.50	0.38	0.43
ไข่	0.83	0.83	0.83
มะม่วง	0.62	0.56	0.59
มังคุด	0.50	0.67	0.57
ส้ม	0.56	0.62	0.59
มะละกอ	0.75	0.82	0.78
สับปะรด	0.80	0.62	0.70
ชมพู่	0.62	0.89	0.73
แตงโม	0.56	0.62	0.59
เชน	0.67	0.44	0.53
หู	0.71	0.56	0.63
ข้อศอก	0.88	0.78	0.82
ตา	0.78	0.78	0.78
เท้า	0.67	0.75	0.71
มือ	0.67	0.60	0.63
ปาก	0.71	0.83	0.77
จมูก	0.40	0.50	0.44
ลิ้น	1.00	0.50	0.67
ฟัน	0.50	0.43	0.46
หวี	0.78	0.70	0.74

ตาราง 25 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัคดม	0.71	0.71	0.71
แก้ว	0.83	0.67	0.74
ปิ่น	0.73	1.00	0.85
หมวก	0.53	0.80	0.64
กางเกง	0.80	0.50	0.62
ดินสอ	0.73	0.73	0.73
เสื่อ	0.78	0.58	0.67
โต๊ะ	0.67	0.80	0.73
แปรงสีฟัน	1.00	0.75	0.86
กิน	0.62	0.56	0.59
เตะฟุตบอล	0.75	0.50	0.60
วิ่ง	0.50	0.55	0.52
อาบน้ำ	0.83	1.00	0.91
ยื่น	0.80	1.00	0.89
รดน้ำต้นไม้	0.77	0.91	0.83
ดูทีวี	0.75	0.86	0.80
วิ่ง	0.70	0.78	0.74
เช็ดตัว	0.67	0.71	0.69
เขียน	1.00	0.75	0.86
Accuracy			0.70
Macro avg	0.70	0.70	0.69
Weighted avg	0.71	0.70	0.69



ตาราง 26 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.78	0.64	0.70
ไก่	0.70	0.58	0.64
จระเข้	0.71	0.45	0.56
หมา	1.00	0.67	0.80
เปิด	1.00	0.70	0.82
ข้าง	0.78	0.78	0.78
สิงโต	0.55	0.75	0.63
ลิง	1.00	0.43	0.60
งู	0.75	0.86	0.80
เสือ	0.82	0.90	0.86
แอปเปิ้ล	1.00	0.50	0.67
กล้วย	0.71	0.71	0.71
ไข่	0.80	0.89	0.84
มะม่วง	0.62	0.56	0.59
มังคุด	0.71	1.00	0.83
ส้ม	0.58	0.70	0.64
มะละกอ	0.77	0.83	0.80
สับปะรด	0.71	0.83	0.77
ชมพู่	0.70	0.88	0.78
แตงโม	0.50	0.38	0.43
แตง	1.00	0.64	0.78
ทุ	1.00	0.58	0.74
ข้อศอก	0.75	0.75	0.75
ตา	0.90	1.00	0.95
เท้า	0.83	0.71	0.77
มือ	0.50	0.67	0.57
ปาก	0.88	0.78	0.82
จมูก	0.25	0.20	0.22
ลิ้น	0.73	0.80	0.76
ฟัน	0.62	1.00	0.76
หวี	0.78	0.70	0.74
พดลัม	0.50	0.80	0.62
แก้ว	0.33	0.09	0.14
ปิ่น	0.29	0.67	0.40
หมวก	0.71	0.71	0.71
กางเกง	0.80	0.89	0.84

ตาราง 26 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
ดินสอ	0.50	0.60	0.55
เสื่อ	0.78	0.64	0.70
โต๊ะ	0.60	0.60	0.60
แปลงสีพื้น	0.56	0.71	0.63
กิน	0.75	1.00	0.86
เตะฟุตบอล	0.73	1.00	0.84
วิ่ง	0.50	0.71	0.59
อาบน้ำ	0.83	0.62	0.71
ยืน	0.56	0.56	0.56
รดน้ำต้นไม้	0.86	0.86	0.86
ดูทีวี	0.67	0.67	0.67
วิ่ง	0.87	0.81	0.84
เช็ดตัว	0.64	0.78	0.70
เขียน	0.86	0.60	0.71
ดำ	0.55	0.60	0.57
น้ำเงิน	0.80	0.57	0.67
น้ำตาล	0.67	1.00	0.80
เขี้ยว	0.89	0.53	0.67
ส้ม (สี)	0.82	0.90	0.86
ม่วง	0.64	0.88	0.74
แดง	0.71	1.00	0.83
ฟ้า	0.58	1.00	0.74
ขาว	0.75	0.75	0.75
เหลือง	0.89	0.62	0.73
Accuracy			0.71
Macro avg	0.72	0.72	0.70
Weighted avg	0.73	0.71	0.70



ตาราง 27 ผลลัพธ์ของวิธีการ Neural Network กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.64	0.70	0.67
ไก่	0.90	0.69	0.78
จระเข้	0.60	0.60	0.60
หมา	0.18	0.40	0.25
เปิด	0.73	1.00	0.84
ข้าง	0.42	0.56	0.48
สิงโต	0.83	0.91	0.87
ลิง	0.86	0.86	0.86
งู	0.57	0.80	0.67
เสือ	0.77	0.83	0.80
แอปเปิ้ล	1.00	0.79	0.88
กล้วย	0.62	0.45	0.53
ไข่	1.00	0.91	0.95
มะม่วง	0.78	0.78	0.78
มังคุด	0.64	0.88	0.74
ส้ม	0.89	0.80	0.84
มะละกอ	0.29	0.29	0.29
สับปะรด	0.88	0.78	0.82
ชมพู่	0.92	0.73	0.81
แตงโม	0.60	0.67	0.63
เชน	0.50	0.71	0.59
หู	0.73	0.73	0.73
ข้อศอก	0.93	0.93	0.93
ตา	1.00	0.20	0.33
เท้า	0.67	0.33	0.44
มือ	0.86	0.86	0.86
ปาก	0.55	0.75	0.63
จมูก	0.64	1.00	0.78
ลิ้น	0.57	0.67	0.62
ฟัน	0.77	0.91	0.83
หวี	0.78	0.70	0.74

ตาราง 27 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัสดม	0.78	0.78	0.78
แก้ว	0.67	0.86	0.75
ปิ่น	0.86	0.86	0.86
หมวก	0.67	0.80	0.73
กางเกง	0.43	0.75	0.55
ดินสอ	0.50	0.30	0.37
เสื่อ	0.38	0.45	0.42
โต๊ะ	0.50	0.83	0.62
แปรงสีฟัน	0.71	0.83	0.77
กิน	0.86	0.60	0.71
เตะฟุตบอล	0.67	0.67	0.67
วิ่ง	0.67	0.75	0.71
อาบน้ำ	1.00	0.60	0.75
ยื่น	0.75	0.50	0.60
รดน้ำต้นไม้	0.73	0.92	0.81
ดูทีวี	0.88	0.88	0.88
วิ่ง	0.67	0.80	0.73
เข็ดตัว	0.75	0.50	0.60
เขียน	0.57	0.44	0.50
ดำ	0.70	0.54	0.61
น้ำเงิน	0.92	1.00	0.96
น้ำตาล	1.00	0.70	0.82
เขียว	0.73	1.00	0.84
ส้ม (สี)	0.71	0.50	0.59
ม่วง	0.90	0.69	0.78
แดง	0.86	0.60	0.71
ฟ้า	0.86	0.86	0.86
ขาว	0.89	0.73	0.80
เหลือง	0.62	0.71	0.67
อา	0.75	1.00	0.86

ตาราง 27 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
สี่	0.80	0.73	0.76
หิม	0.29	0.40	0.33
คู	0.44	0.44	0.44
ช	0.25	0.33	0.29
ช	0.71	0.50	0.59
Accuracy			0.70
Macro avg	0.71	0.70	0.69
Weighted avg	0.73	0.70	0.70



ตาราง 28 ผลลัพธ์ Accuracy ของวิธีการ Neural Network กับฐานข้อมูลทั้งหมด

ประเภทของข้อมูล	Train (%)	Validation (%)	Test (%)
10 กลุ่มข้อมูล	100	90	92.22
20 กลุ่มข้อมูล	100	78.89	78.33
30 กลุ่มข้อมูล	100	78.89	81.85
40 กลุ่มข้อมูล	100	70.28	74.44
50 กลุ่มข้อมูล	100	72.22	69.78
60 กลุ่มข้อมูล	100	66.48	70.93
66 กลุ่มข้อมูล	100	70.54	70.37

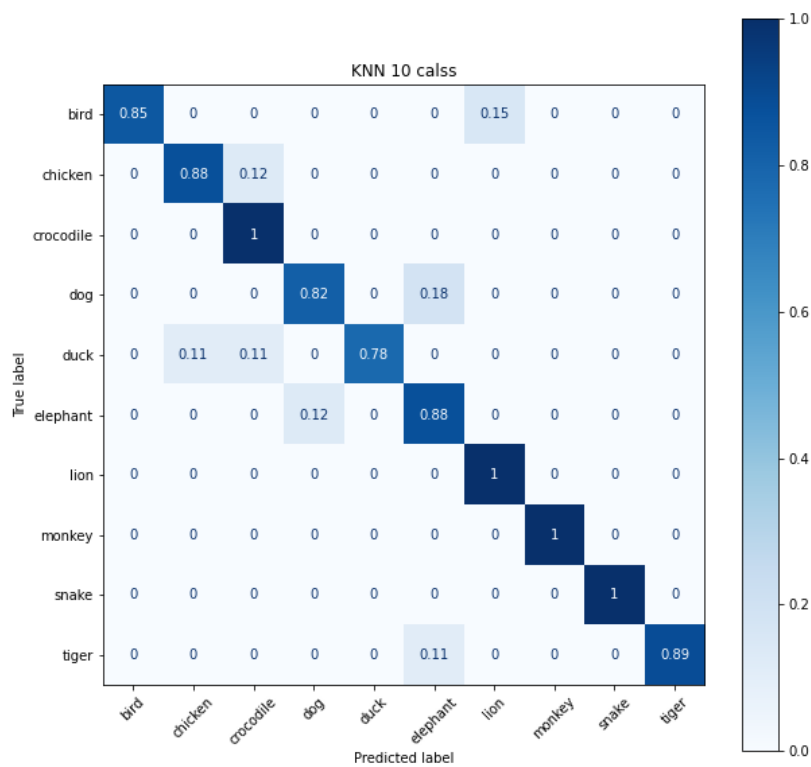
จากตาราง 28 จะเห็นได้ว่าวิธีการ Neural Network มีความเหมาะสมกับฐานข้อมูลแบบ 10 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy สูงที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 92.22% และมีความไม่เหมาะสมกับฐานข้อมูลแบบ 50 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy น้อยที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 69.78%

#### 4.6 ผลลัพธ์ของการประเมินประสิทธิภาพของระบบรู้จำเสียงโดยใช้วิธีการ KNN

ในขั้นตอนการ train model ของ KNN มีการใช้ค่า K หรือจำนวนข้อมูลที่ใช้ในการตรวจสอบข้อมูลเท่ากับ 1, สมการหาระยะทางแบบ Euclidean ในการคำนวณหาระยะทางของข้อมูลฝึกและข้อมูลทดสอบ และ epochs จำนวน 200 หลังจากได้ model มาแล้ว จะนำ model ที่ได้ไปทดสอบกับ test data ในการประเมินประสิทธิภาพของ model จะใช้ Confusion Matrix ในการหาค่า Accuracy, Sensitivity, F1-score และ Precision ของ model ซึ่งได้ผลลัพธ์ดังตารางต่อไปนี้

ตาราง 29 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 10 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	1.00	0.85	0.92
ไก่	0.88	0.88	0.88
จระเข้	0.75	1.00	0.86
หมา	0.90	0.82	0.86
เปิด	1.00	0.78	0.88
ข้าง	0.70	0.88	0.78
สิงโต	0.78	1.00	0.88
ลิง	1.00	1.00	1.00
งู	1.00	1.00	1.00
เสือ	1.00	0.89	0.94
Accuracy			0.90
Macro avg	0.90	0.91	0.90
Weighted avg	0.92	0.90	0.90



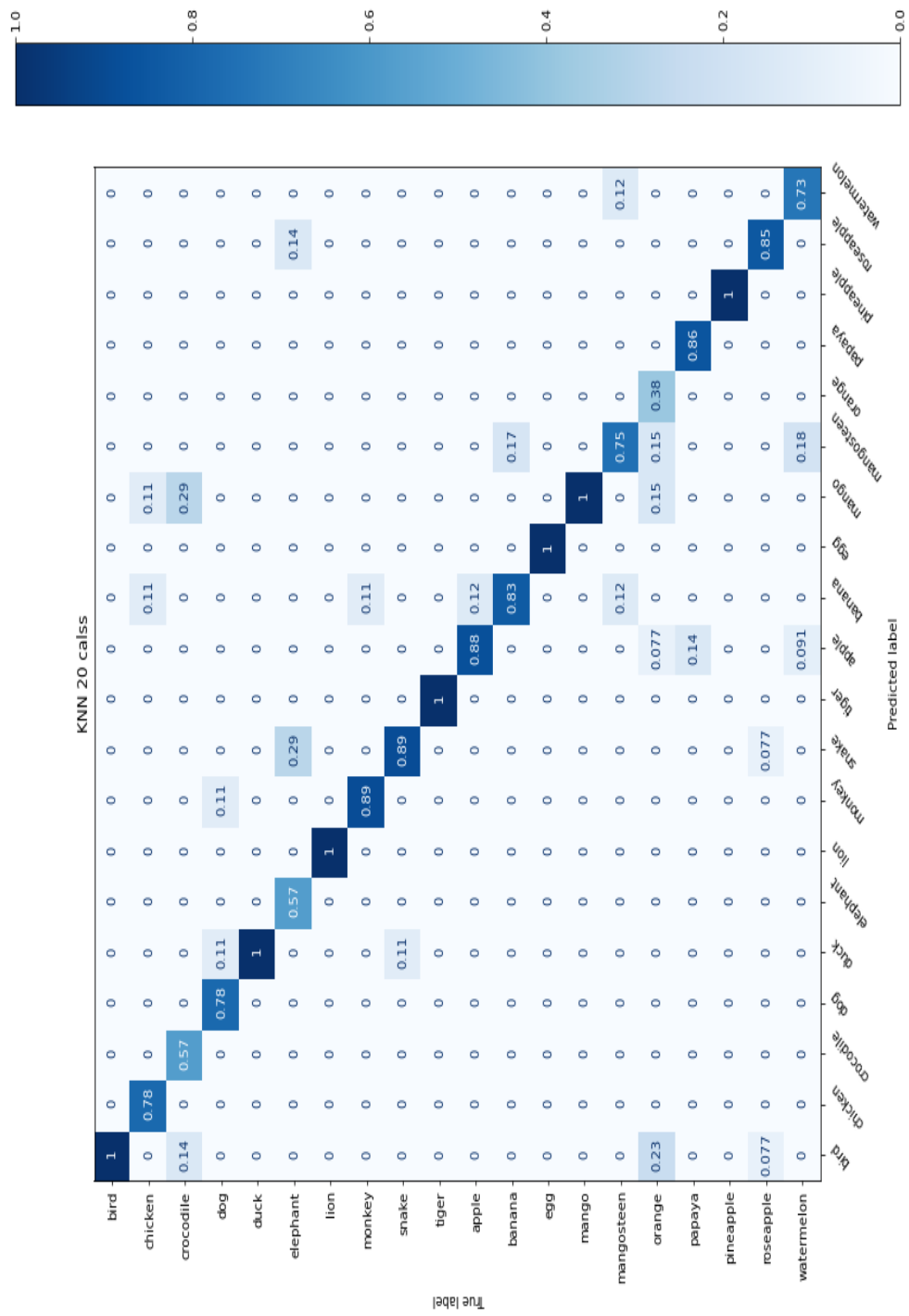
ภาพประกอบ 62 Confusion Matrix ของ KNN 10 กลุ่มข้อมูล

ตาราง 30 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 20 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.55	1.00	0.71
ไก่	1.00	0.78	0.88
จระเข้	1.00	0.57	0.73
หมา	1.00	0.78	0.88
เป็ด	0.75	1.00	0.86
ช้าง	1.00	0.57	0.73
สิงโต	1.00	1.00	1.00
ลิง	0.89	0.89	0.89
งู	0.73	0.89	0.80
เสือดอ	1.00	1.00	1.00
แอปเปิ้ล	0.70	0.88	0.78

ตาราง 30 (ต่อ)

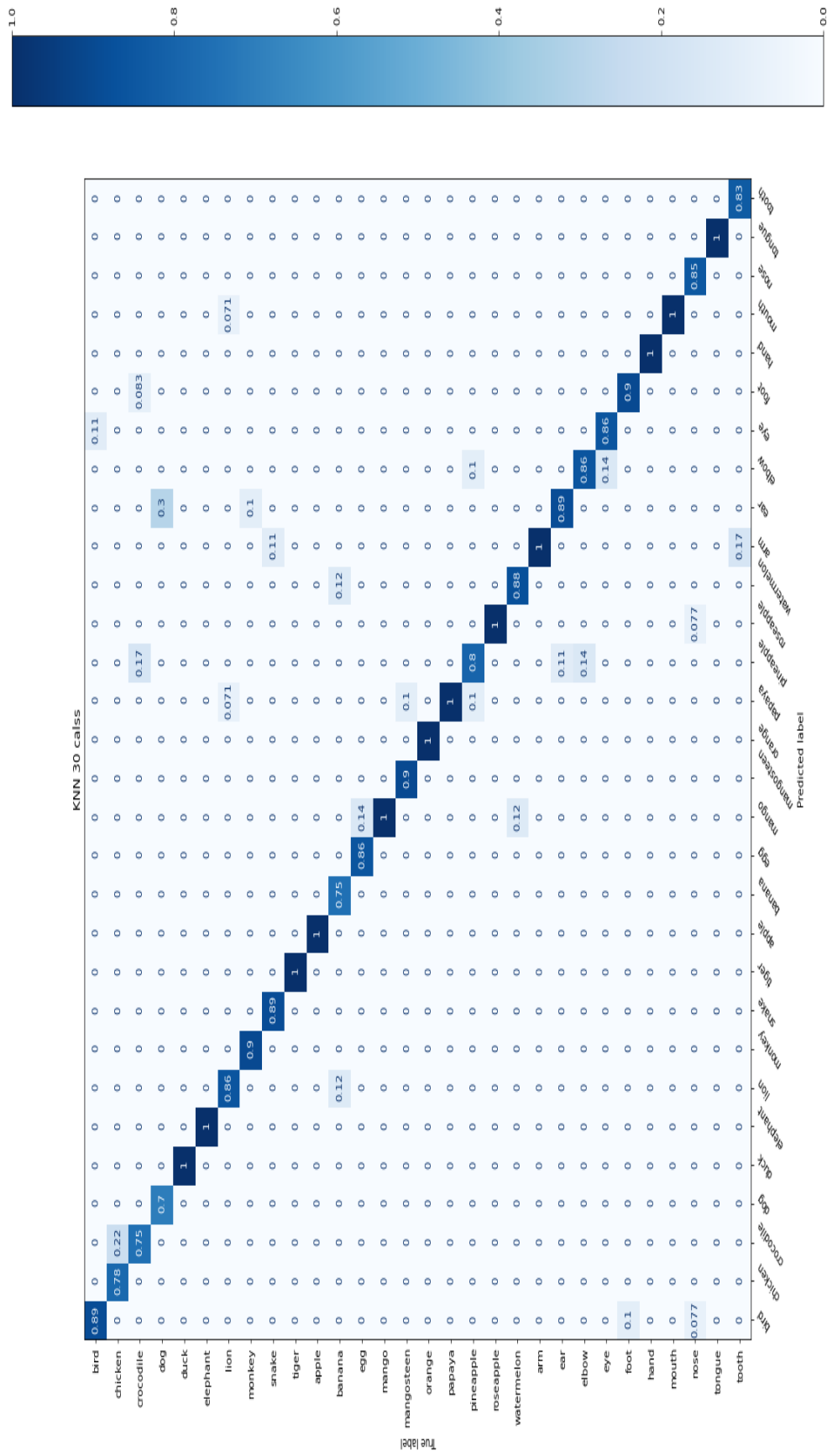
ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
กล้วย	0.50	0.83	0.62
ไข่	1.00	1.00	1.00
มะม่วง	0.62	1.00	0.76
มังคุด	0.71	0.75	0.73
ส้ม	1.00	0.38	0.56
มะละกอ	1.00	0.86	0.92
สับปะรด	1.00	1.00	1.00
ชมพู่	0.92	0.85	0.88
แตงโม	0.80	0.73	0.76
Accuracy			0.82
Macro avg	0.86	0.84	0.82
Weighted avg	0.87	0.82	0.82



ภาพประกอบ 63 Confusion Matrix ของ KNN 20 กลุ่มข้อมูล

ตาราง 31 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 30 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.80	0.89	0.84
ไก่	1.00	0.78	0.88
จระเข้	0.82	0.75	0.78
หมา	1.00	0.70	0.82
เบ็ด	1.00	1.00	1.00
ช้าง	1.00	1.00	1.00
สิงโต	0.92	0.86	0.89
ลิง	1.00	0.90	0.95
งู	1.00	0.89	0.94
เสื่อ	1.00	1.00	1.00
แอปเปิ้ล	1.00	1.00	1.00
กล้วย	1.00	0.75	0.86
ไข่	1.00	0.86	0.92
มะม่วง	0.82	1.00	0.90
มังคุด	1.00	0.90	0.95
ส้ม	1.00	1.00	1.00
มะละกอ	0.73	1.00	0.84
สับปะรด	0.67	0.80	0.73
ชมพู่	0.92	1.00	0.96
แตงโม	0.88	0.88	0.88
ขนุน	0.67	1.00	0.80
ทุเรียน	0.67	0.89	0.76
ข้าวสอก	0.75	0.86	0.80
ตา	0.86	0.86	0.86
เท้า	0.90	0.90	0.90
มือ	1.00	1.00	1.00
ปาก	0.75	1.00	0.86
จมูก	1.00	0.85	0.92
ลิ้น	1.00	1.00	1.00
ฟัน	1.00	0.83	0.91
Accuracy			0.90
Macro avg	0.90	0.90	0.90
Weighted avg	0.91	0.90	0.90



ภาพประกอบ 64 Confusion Matrix ของ KNN 30 กลุ่มข้อมูล

ตาราง 32 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 40 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.64	1.00	0.78
ไก่	0.91	1.00	0.95
จระเข้	0.43	0.50	0.46
หมา	1.00	0.56	0.71
เปิด	0.80	0.73	0.76
ข้าง	0.92	1.00	0.96
สิงโต	0.43	0.43	0.43
ลิง	0.83	0.71	0.77
งู	1.00	0.80	0.89
เสือ	1.00	0.71	0.83
แอปเปิ้ล	1.00	0.50	0.67
กล้วย	0.88	1.00	0.93
ไข่	0.89	0.73	0.80
มะม่วง	0.71	0.83	0.77
มังคุด	0.67	0.67	0.67
ส้ม	1.00	0.78	0.88
มะละกอ	0.89	1.00	0.94
สับปะรด	1.00	0.70	0.82
ชมพู่	0.70	0.78	0.74
แตงโม	0.56	0.56	0.56
ขน	1.00	1.00	1.00
หู	0.89	0.73	0.80
ข้อศอก	0.62	0.89	0.73
ตา	0.70	1.00	0.82
เท้า	0.86	1.00	0.92
มือ	0.85	0.85	0.85
ปาก	0.80	0.80	0.80
จมูก	0.86	1.00	0.92
ลิ้น	1.00	1.00	1.00
ฟัน	0.56	0.62	0.59
หัว	0.67	0.40	0.50

ตาราง 32 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัสดม	1.00	0.92	0.96
แก้ว	1.00	1.00	1.00
ปิ่น	0.82	0.82	0.82
หมวก	0.80	1.00	0.89
กางเกง	1.00	0.91	0.95
ดินสอ	1.00	0.80	0.89
เสื้อ	0.90	0.75	0.82
โต๊ะ	0.59	1.00	0.74
แปรงสีฟัน	0.91	1.00	0.95
Accuracy			0.82
Macro avg	0.83	0.81	0.81
Weighted avg	0.84	0.82	0.82



ตาราง 33 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 50 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.82	0.90	0.86
ไก่	1.00	1.00	1.00
จระเข้	1.00	0.45	0.62
หมา	0.70	0.78	0.74
เปิด	0.67	0.50	0.57
ข้าง	0.88	1.00	0.93
สิงโต	0.89	0.80	0.84
ลิง	0.86	0.75	0.80
งู	1.00	1.00	1.00
เสือ	0.85	1.00	0.92
แอปเปิ้ล	0.86	1.00	0.92
กล้วย	0.86	0.75	0.80
ไข่	0.92	1.00	0.96
มะม่วง	1.00	0.89	0.94
มังคุด	1.00	1.00	1.00
ส้ม	0.80	1.00	0.89
มะละกอ	0.82	0.82	0.82
สับปะรด	0.54	0.54	0.54
ชมพู่	1.00	0.89	0.94
แตงโม	1.00	0.62	0.77
ขน	0.83	0.56	0.67
หู	1.00	0.56	0.71
ข้อศอก	0.75	1.00	0.86
ตา	0.67	0.89	0.76
เท้า	0.75	0.75	0.75
มือ	0.69	0.90	0.78
ปาก	0.71	0.83	0.77
จมูก	0.58	0.88	0.70
ลิ้น	0.80	1.00	0.89
ฟัน	0.55	0.86	0.67
หัว	0.80	0.80	0.80

ตาราง 33 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พัสดม	0.78	1.00	0.88
แก้ว	0.87	0.87	0.87
ปิ่น	0.83	0.91	0.87
หมวก	0.90	0.90	0.90
กางเกง	1.00	0.62	0.77
ดินสอ	0.83	0.91	0.87
เสื่อ	1.00	0.58	0.74
โต๊ะ	0.83	1.00	0.91
แปรงสีฟัน	1.00	0.92	0.96
กิน	0.80	0.89	0.84
เตะฟุตบอล	1.00	0.58	0.74
วิ่ง	1.00	0.73	0.84
อาบน้ำ	0.83	1.00	0.91
ยืน	1.00	1.00	1.00
รดน้ำต้นไม้	0.79	1.00	0.88
ดูทีวี	0.75	0.86	0.80
วิ่ง	0.90	1.00	0.95
เช็ดตัว	0.69	0.79	0.73
เขียน	1.00	0.75	0.86
Accuracy			0.83
Macro avg	0.85	0.84	0.83
Weighted avg	0.85	0.83	0.82



ตาราง 34 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 60 กลุ่มข้อมูล

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	0.90	0.82	0.86
ไก่	1.00	1.00	1.00
จระเข้	0.82	0.82	0.82
หมา	0.91	0.83	0.87
เบ็ด	0.91	1.00	0.95
ช้าง	0.82	1.00	0.90
สิงโต	0.64	0.88	0.74
ลิง	1.00	0.71	0.83
งู	0.64	1.00	0.78
เสื่อ	0.82	0.90	0.86
แอปเปิ้ล	1.00	1.00	1.00
กล้วย	0.86	0.86	0.86
ไข่	1.00	0.78	0.88
มะม่วง	0.89	0.89	0.89
มังคุด	0.83	1.00	0.91
ส้ม	1.00	0.80	0.89
มะละกอ	0.92	1.00	0.96
สับปะรด	0.86	1.00	0.92
ชมพู่	0.80	1.00	0.89
แตงโม	1.00	0.75	0.86
แคน	0.56	0.91	0.69
หู	0.92	0.92	0.92
ข้อศอก	1.00	0.88	0.93
ตา	0.78	0.78	0.78
เท้า	0.86	0.86	0.86
มือ	0.67	1.00	0.80
ปาก	0.89	0.89	0.89
จมูก	0.60	0.60	0.60
ลิ้น	0.90	0.90	0.90
ฟัน	0.89	1.00	0.94
หวี	1.00	0.90	0.95
พัดลม	1.00	1.00	1.00
แก้ว	1.00	0.36	0.53
ปิ่น	0.46	1.00	0.63
หมวก	0.36	0.57	0.44
กางเกง	0.71	0.56	0.63

ตาราง 34 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
ดินสอ	0.62	1.00	0.77
เส้น	0.89	0.73	0.80
โต๊ะ	0.78	0.70	0.74
แปรงสีฟัน	1.00	1.00	1.00
กิน	1.00	1.00	1.00
เตะฟุตบอล	1.00	1.00	1.00
วิ่ง	0.78	1.00	0.88
อาบน้ำ	0.86	0.75	0.80
ยื่น	0.89	0.89	0.89
รดน้ำต้นไม้	1.00	0.86	0.92
ดูทีวี	1.00	1.00	1.00
วิ่ง	0.93	0.88	0.90
เช็ดตัว	1.00	0.56	0.71
เขียน	1.00	0.80	0.89
ดำ	1.00	0.90	0.95
น้ำเงิน	0.80	0.57	0.67
น้ำตาล	1.00	0.88	0.93
เขี้ยว	1.00	0.80	0.89
ส้ม (สี)	0.88	0.70	0.78
ม่วง	1.00	1.00	1.00
แดง	0.91	1.00	0.95
ฟ้า	1.00	1.00	1.00
ขาว	1.00	0.50	0.67
เหลือง	0.93	1.00	0.96
Accuracy			0.85
Macro avg	0.87	0.86	0.85
Weighted avg	0.88	0.85	0.85



ตาราง 35 ผลลัพธ์ของวิธีการ KNN กับฐานข้อมูลแบบ 66 กลุ่มข้อมูล

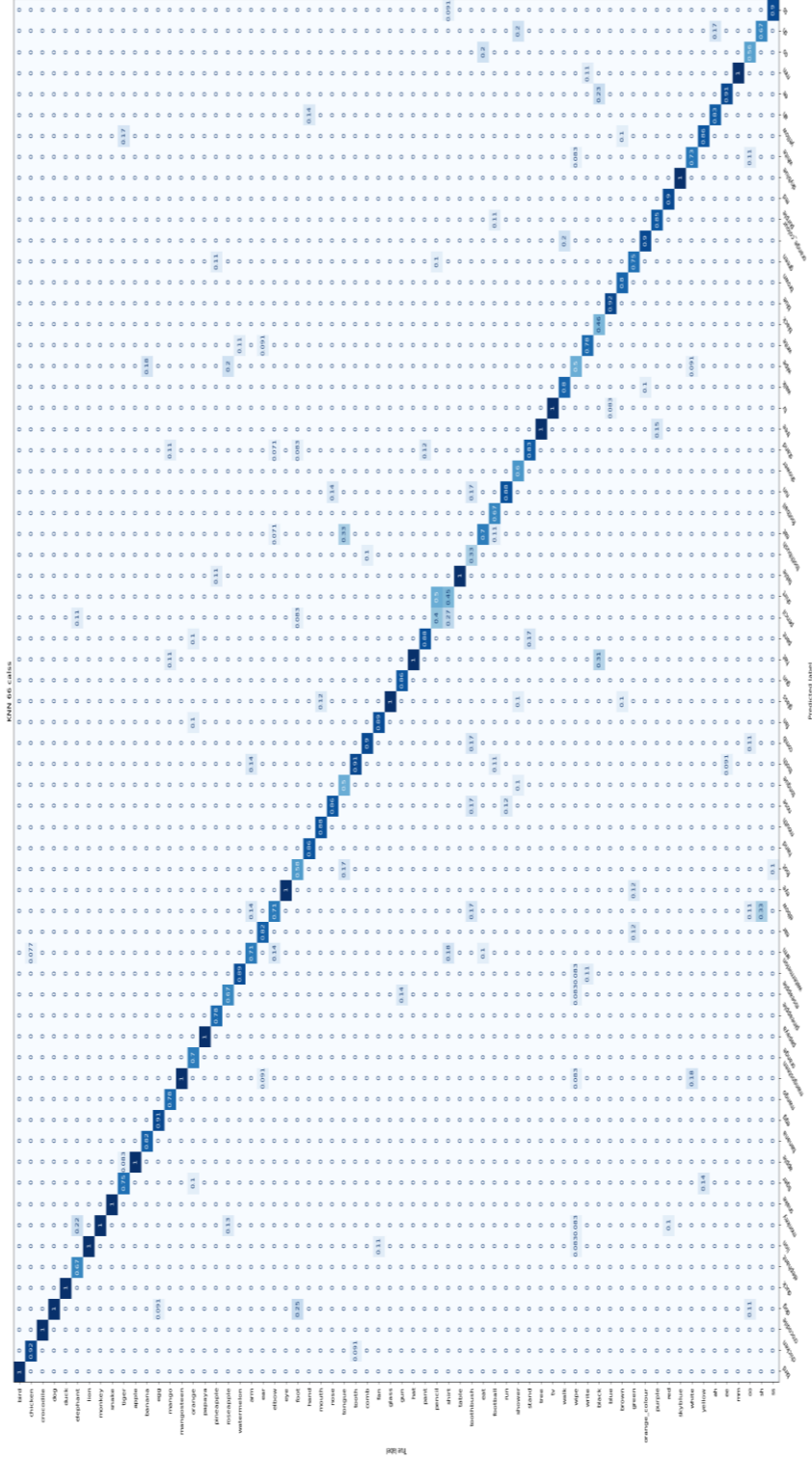
ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
นก	1.00	1.00	1.00
ไก่	0.92	0.92	0.92
จระเข้	1.00	1.00	1.00
หมา	0.50	1.00	0.67
เปิด	1.00	1.00	1.00
ข้าง	1.00	0.67	0.80
สิงโต	0.85	1.00	0.92
ลิง	0.54	1.00	0.70
งู	1.00	1.00	1.00
เสือ	0.82	0.75	0.78
แอปเปิ้ล	0.93	1.00	0.97
กล้วย	1.00	0.82	0.90
ไข่	1.00	0.91	0.95
มะม่วง	1.00	0.78	0.88
มังคุด	0.67	1.00	0.80
ส้ม	1.00	0.70	0.82
มะละกอ	1.00	1.00	1.00
สับปะรด	1.00	0.78	0.88
ชมพู่	0.83	0.67	0.74
แตงโม	0.80	0.89	0.84
แตง	0.45	0.71	0.56
ทุเรียน	0.90	0.82	0.86
ข้อศอก	0.71	0.71	0.71
ตา	0.83	1.00	0.91
เท้า	0.78	0.58	0.67
มือ	1.00	0.86	0.92
ปาก	1.00	0.88	0.93
จมูก	0.75	0.86	0.80
ลิ้น	0.75	0.50	0.60
ฟัน	0.77	0.91	0.83
หัว	0.82	0.90	0.86

ตาราง 35 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
พุดลม	0.89	0.89	0.89
แก้ว	0.70	1.00	0.82
ปิ่น	1.00	0.86	0.92
หมวก	0.50	1.00	0.67
กางเกง	0.78	0.88	0.82
ดินสอ	0.44	0.40	0.42
เสื้อ	0.50	0.45	0.48
โต๊ะ	0.86	1.00	0.92
แปรงสีฟัน	0.67	0.33	0.44
กิ้ง	0.64	0.70	0.67
ตะปูตบอล	1.00	0.67	0.80
วุ้น	0.78	0.88	0.82
อาบน้ำ	1.00	0.60	0.75
ยีน	0.56	0.83	0.67
รดน้ำต้นไม้	0.86	1.00	0.92
คูพีวี	0.89	1.00	0.94
วุ้น	0.89	0.80	0.84
เช็ดตัว	0.50	0.50	0.50
เขียน	0.78	0.78	0.78
ดำ	1.00	0.46	0.63
น้ำเงิน	1.00	0.92	0.96
น้ำตาล	1.00	0.80	0.89
เชียว	0.75	0.75	0.75
ส้ม (สี)	0.82	0.90	0.86
ม่วง	0.92	0.85	0.88
แดง	1.00	0.90	0.95
ฟ้า	1.00	1.00	1.00
ขาว	0.80	0.73	0.76
เหลือง	0.67	0.86	0.75
อา	0.83	0.83	0.83

ตาราง 35 (ต่อ)

ชื่อกลุ่มข้อมูล	Precision	Sensitivity	F1-score
สี่	0.77	0.91	0.83
สี่ม	0.83	1.00	0.91
คู	0.71	0.56	0.63
ช	0.40	0.67	0.50
ช	0.90	0.90	0.90
Accuracy			0.81
Macro avg	0.82	0.82	0.81
Weighted avg	0.83	0.81	0.81



ภาพประกอบ 68 Confusion Matrix ของ KNN 66 กลุ่มข้อมูล

ตาราง 36 ผลลัพธ์ Accuracy ของวิธีการ KNN กับฐานข้อมูลทั้งหมด

ประเภทของข้อมูล	Train (%)	Validation (%)	Test (%)
10 กลุ่มข้อมูล	100	92.22	90.00
20 กลุ่มข้อมูล	100	87.22	82.22
30 กลุ่มข้อมูล	100	87.41	89.63
40 กลุ่มข้อมูล	100	85.83	81.67
50 กลุ่มข้อมูล	100	84.22	82.67
60 กลุ่มข้อมูล	100	83.15	85.37
66 กลุ่มข้อมูล	100	80.47	80.98

จากตาราง 36 จะเห็นได้ว่าวิธีการ KNN มีความเหมาะสมกับฐานข้อมูลแบบ 30 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy สูงที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 89.63% และมีความไม่เหมาะสมกับฐานข้อมูลแบบ 66 กลุ่มข้อมูลมากที่สุด เนื่องจากมี Accuracy น้อยที่สุดเมื่อเปรียบเทียบกับฐานข้อมูลแบบอื่นๆ คือ 80.98%

## บทที่ 5

### สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในการวิจัยเรื่องการพัฒนาาระบบสารสนเทศเพื่อการบริหารจัดการงานปริญญาโท/สตรนิพนธ์ บัณฑิตวิทยาลัยมหาวิทยาลัยศรีนครินทรวิโรฒ ผู้วิจัยได้ทำการประเมินประสิทธิภาพของระบบ โดยการประเมินผลการวิจัย จาก ผู้เชี่ยวชาญและผู้ใช้ระบบ เพื่อวิเคราะห์หาประสิทธิภาพของระบบ หลังจากได้ผลการดำเนินงานแล้ว สามารถสรุปผลการดำเนินงาน โดยแบ่งหัวข้อในการสรุปผลได้ดังต่อไปนี้

1 สรุปผลการวิจัย

2 อภิปรายผลการวิจัย

#### 5.1 สรุปผลการวิจัย

งานวิจัยนี้นำเสนอระบบรู้จำเสียงสำหรับผู้พิการทางการได้ยิน โดยใช้ฐานข้อมูลจากการบันทึกเสียง Ling 6 sound และคำศัพท์ในชีวิตประจำวัน 60 คำ จากผู้ชาย 10 คน และผู้หญิง 10 คน ที่มีการผ่าตัดประสาทหูเทียมทั้งสองข้าง โดยได้รับคำปรึกษาจากนักฝึกการพูดของศูนย์การแพทย์สมเด็จพระเทพรัตนราชสุดาฯ สยามบรมราชกุมารีในด้านของคำศัพท์ในชีวิตประจำวัน ในการสร้างระบบรู้จำเสียง ได้นำฐานข้อมูลจากกลุ่มตัวอย่างมาหาลักษณะสำคัญของเสียงด้วยเทคนิค MFCC แล้วนำมาสร้างโมเดลด้วยวิธีการต่างๆ ได้แก่ Convolutional Neural Network (CNN), Artificial Neuron Network (ANN), Neural Network (NN) และ K-Nearest Neighbors (KNN) เพื่อหาวิธีการที่เหมาะสมกับชุดข้อมูลมากที่สุด

จากการทดลองนำข้อมูลที่แบ่งเป็น 7 กลุ่มได้แก่ 10 กลุ่มข้อมูล, 20 กลุ่มข้อมูล, 30 กลุ่มข้อมูล, 40 กลุ่มข้อมูล, 50 กลุ่มข้อมูล, 60 กลุ่มข้อมูล และ 66 กลุ่มข้อมูล ทดลองใน 4 วิธีการ ได้แก่ CNN, ANN, Neural Network และ KNN โดยแบ่งข้อมูลออกเป็น train, validation และ test ได้ผลดังตาราง 37

ตาราง 37 สรุปผล Accuracy จากการสอนและทดสอบโมเดล

ประเภทของข้อมูล	วิธีการ	Train (%)	Validation (%)	Test (%)
10 กลุ่มข้อมูล	CNN	97.62	83.33	77.78
10 กลุ่มข้อมูล	ANN	100	91.22	97.78
10 กลุ่มข้อมูล	Neural Network	100	90.00	92.22

ตาราง 37 (ต่อ)

ประเภทของข้อมูล	วิธีการ	Train (%)	Validation (%)	Test (%)
10 กลุ่มข้อมูล	KNN	100	92.22	90.00
20 กลุ่มข้อมูล	CNN	99.88	81.67	73.89
20 กลุ่มข้อมูล	ANN	100	85.56	86.11
20 กลุ่มข้อมูล	Neural Network	100	78.89	78.33
20 กลุ่มข้อมูล	KNN	100	87.22	82.22
30 กลุ่มข้อมูล	CNN	93.97	64.07	65.19
30 กลุ่มข้อมูล	ANN	100	86.67	87.78
30 กลุ่มข้อมูล	Neural Network	100	78.89	81.85
30 กลุ่มข้อมูล	KNN	100	87.41	89.63
40 กลุ่มข้อมูล	CNN	100	64.44	58.89
40 กลุ่มข้อมูล	ANN	100	78.61	80.00
40 กลุ่มข้อมูล	Neural Network	100	70.28	74.44
40 กลุ่มข้อมูล	KNN	100	85.83	81.67
50 กลุ่มข้อมูล	CNN	99.67	58.89	57.11
50 กลุ่มข้อมูล	ANN	100	78.67	80.44
50 กลุ่มข้อมูล	Neural Network	100	72.22	69.78
50 กลุ่มข้อมูล	KNN	100	84.22	82.67
60 กลุ่มข้อมูล	CNN	100	65.19	62.41
60 กลุ่มข้อมูล	ANN	100	74.26	79.81
60 กลุ่มข้อมูล	Neural Network	100	66.48	70.93
60 กลุ่มข้อมูล	KNN	100	83.15	85.37
66 กลุ่มข้อมูล	CNN	96.47	57.41	54.04
66 กลุ่มข้อมูล	ANN	100	79.46	79.63
66 กลุ่มข้อมูล	Neural Network	100	70.54	70.37
66 กลุ่มข้อมูล	KNN	100	80.47	80.98

จากตาราง 37 แสดงให้เห็นว่าจำนวนของคำศัพท์มีผลต่อประสิทธิภาพของโมเดลซึ่งชุดข้อมูลแบบ 10 กลุ่มข้อมูล มีประสิทธิภาพที่ดีที่สุดในการเรียนรู้ของโมเดลเมื่อเทียบกับชุดข้อมูลอื่นๆ โดยในการทดสอบโมเดลด้วยชุดข้อมูลแบบ 10 กลุ่มข้อมูล และ 20 กลุ่มข้อมูล วิธีการ ANN

ได้ accuracy สูงที่สุดเมื่อเปรียบเทียบกับวิธีการอื่นๆ คือ 97.78% และ 86.11% ตามลำดับ สำหรับ ชุดข้อมูลแบบ 30 กลุ่มข้อมูล, 40 กลุ่มข้อมูล, 50 กลุ่มข้อมูล, 60 กลุ่มข้อมูล และ 66 กลุ่มข้อมูล วิธีการ KNN ได้ accuracy สูงที่สุดเมื่อเปรียบเทียบกับวิธีการอื่นๆ คือ 89.63%, 81.67%, 82.67%, 85.37% และ 80.98% ตามลำดับ แสดงให้เห็นว่าวิธีการ ANN และ KNN มีความยืดหยุ่นกับชุด ข้อมูลที่ยังไม่เคยเจอมาก่อนได้ดีเมื่อเทียบกับวิธีการที่เหลือ เนื่องจากทั้ง ANN และ KNN ได้ ผลลัพธ์จากการทดสอบโมเดลที่มีความใกล้เคียงกัน

## 5.2 อภิปรายผลการวิจัย

ผลลัพธ์จากการเรียนรู้ของโมเดลนอกจากสามารถจำแนกคำศัพท์ต่างๆ ได้แล้ว ยังได้ ไฟล์โมเดล ซึ่งสามารถนำไปสร้างอุปกรณ์หรือเครื่องมือที่นำไปใช้จริงในอนาคต สำหรับการฟื้นฟู การสื่อสารของผู้ที่พิการทางการได้ยิน เช่น นำไฟล์โมเดลไปใส่ในแอปพลิเคชันของระบบแอน ดรอยด์บนอุปกรณ์สื่อสาร หรือ นำไฟล์โมเดลไปสร้างระบบฝังตัว (Embedded System) เป็นต้น โดยระบบนี้มีความเหมาะสมกับผู้พิการประเภทผู้ที่มีปัญหาสูญเสียการได้ยินเกิดขึ้นโดยที่ยังไม่ เคยมีภาษาพูดมาก่อน (Pre-lingual Deafness) เนื่องจากผู้พิการกลุ่มนี้ไม่เคยมีภาษาพูดมาก่อน ทำให้มีปัญหาในการฟื้นฟูการสื่อสารมากกว่าผู้พิการที่เคยมีภาษาพูดมาก่อนแล้ว (Post-lingual Deafness) แต่หากนำระบบนี้ไปใช้กับผู้พิการที่เคยมีภาษาพูดมาก่อนแล้วจะทำให้ผู้พิการกลุ่มนี้ สามารถฟื้นฟูการสื่อสารได้เร็วกว่าผู้พิการที่ยังไม่เคยมีภาษาพูดมาก่อน เนื่องจากเคยมีพื้นฐานใน ด้านของภาษามาก่อน จึงทำให้สามารถเรียนรู้และฝึกฝนได้เร็วกว่าผู้พิการที่ยังไม่เคยมีภาษาพูด มาก่อน โดยระบบนี้สามารถเพิ่มการพัฒนาการสื่อสารของผู้พิการทั้ง 2 กลุ่ม ดังกล่าว ส่งผลให้ ระยะเวลาในการฟื้นฟูการสื่อสารใช้เวลาน้อยกว่าระยะเวลาในการฟื้นฟูการสื่อสารในรูปแบบปกติ

ตาราง 38 ตารางเปรียบเทียบค่าความถูกต้อง (Accuracy) ระหว่างงานวิจัยนี้กับงานวิจัยที่ เกี่ยวข้องกับระบบรู้จำเสียงภาษาไทย

การเผยแพร่	ฐานข้อมูล	วิธีหา ลักษณะสำคัญ	วิธีการจำแนก	ค่าความถูก ต้อง (%)
K. Srijiranon et al. (8)	9 กลุ่มข้อมูล	Perceptual Linear Predictive (PLP)	Neuro-Fuzzy	78.99

ตาราง 38 (ต่อ)

การเผยแพร่	ฐานข้อมูล	วิธีหา ลักษณะสำคัญ	วิธีการจำแนก	ค่าความถูกต้อง (%)
P. Sangsai and P. Detyong (17)	6 กลุ่มข้อมูล	Mel-Frequency Cepstral Coefficients (MFCC)	CMUSphinx	81.67
		Support Energy Spectral Density (ESD)	Vector Machine (SVM)	64.45 (ชาย) 69.23 (หญิง)
S. Tantisatirapong et al. (21)	10 กลุ่ม ข้อมูล	Power Spectral Density (PSD)	SVM	68.38 (ชาย) 77.11 (หญิง)
		Spectrogram (SPT)	SVM	58.81 (ชาย) 65.09 (หญิง)
		MFCC	SVM	89.34 (ชาย) 93.81 (หญิง)
งานวิจัยนี้	66 กลุ่ม ข้อมูล	MFCC	KNN	80.98

จากตาราง 38 เมื่อเปรียบเทียบงานวิจัยนี้กับงานวิจัยที่ผ่านมาจะให้เห็นได้ว่าจุดเด่นของงานวิจัยนี้ได้แก่ ด้านของจำนวนคำที่สามารถจำแนกได้โดยงานวิจัยนี้มีจำนวนคำที่สามารถจำแนกได้มากกว่างานวิจัยอื่นๆ และงานวิจัยนี้ใช้คำศัพท์ที่ใช้สำหรับการฟื้นฟูการสื่อสารของผู้พิการทางการได้ยิน และข้อจำกัดของงานวิจัยนี้ได้แก่ ยังไม่มีการทดสอบกับการใช้งานจริงกับผู้พิการทางการได้ยิน และสำหรับการจำแนกเสียงของ 66 คำศัพท์ ยังไม่ได้ประสิทธิภาพที่ดีเพียงพอ

### 5.3 ข้อเสนอแนะ

จากตาราง 37 จะเห็นได้ว่ายิ่งจำนวนคำศัพท์มากประสิทธิภาพของโมเดลจะยิ่งลดลง ทั้งนี้อาจเป็นเพราะจำนวนข้อมูลในฐานข้อมูลแต่ละคำศัพท์ไม่เพียงพอต่อการเรียนรู้ของโมเดลที่

จะต้องจำแนกกลุ่มข้อมูลที่มีจำนวนมาก รวมถึงผลลัพธ์ของวิธีการ CNN และ ANN เนื่องจากเป็นวิธีการการเรียนรู้แบบเชิงลึก ซึ่งจะต้องใช้จำนวนข้อมูลจำนวนมากในการเรียนรู้ ทำให้ได้ผลลัพธ์ในการจำแนกกลุ่มข้อมูลออกมาไม่ดีเท่าที่ควร สำหรับการสร้างโมเดลที่มีคำศัพท์ทั้ง 66 คำ ผู้วิจัยควรต้องเก็บข้อมูลจากกลุ่มตัวอย่างที่มากขึ้นสำหรับการเรียนรู้ของโมเดล เพื่อเพิ่มประสิทธิภาพการจำแนกกลุ่มข้อมูลของโมเดล

งานวิจัยนี้มีข้อเสนอแนะสำหรับกลุ่มผู้พิการทางการได้ยินเพื่อนำไปประยุกต์ใช้ประโยชน์ในอนาคต โดยอธิบายรายละเอียดให้กลุ่มต่างๆ ได้แก่

1. กลุ่มผู้ใช้งานส่วนบุคคล ได้แก่ บุคคลที่มีปัญหาสูญเสียการได้ยินเกิดขึ้นโดยที่ยังไม่เคยมีภาษาพูดมาก่อน และที่เคยมีภาษามาก่อนแล้ว โดยได้รับการผ่าตัดประสาทหูเทียมในช่วงอายุ 5 – 80 ปี ได้ให้คำแนะนำโดยนำโมเดลที่ได้จากงานวิจัยนี้ไปสร้างแอปพลิเคชันบนมือถือถึงสมาร์ตโฟนหรือแท็บเล็ตสำหรับฟื้นฟูการสื่อสารโดยสามารถฝึกฝนการสื่อสารด้วยตนเอง เนื่องจากผู้ช่วยส่วนใหญ่ไม่ได้ฝึกฝนทักษะการสื่อสารเพิ่มเติมด้วยตนเอง ทำให้พัฒนาการการสื่อสารเป็นไปได้ช้า แอปพลิเคชันนี้จึงเป็นตัวช่วยเสริมที่ให้ผู้ช่วยใช้ฝึกทักษะการพูดด้วยตนเองที่บ้านหรือสถานที่อื่นๆ ได้

2. กลุ่มผู้ใช้งานในส่วนที่เป็นองค์กร ได้แก่ โรงพยาบาล โรงเรียนสอนคนหูหนวก สถานคุ้มครองและพัฒนาคนพิการ ศูนย์พัฒนาการเด็กพิการ ได้ให้คำแนะนำโดยการนำโมเดลที่ได้จากงานวิจัยไปสร้างแอปพลิเคชันบนมือถือถึงสมาร์ตโฟนหรือแท็บเล็ต หรือสร้างระบบฝังตัว (Embedded System) บนอุปกรณ์ สำหรับเป็นสื่อการเรียนการสอน รวมถึงการฝึกฝนทักษะการฟัง และการสื่อสารของผู้พิการทางการได้ยิน เพื่อให้ผู้พิการสามารถมีพัฒนาการด้านการสื่อสารให้เหมือนหรือใกล้เคียงกับบุคคลปกติ

3. กลุ่มผู้ใช้งานภาคเอกชน อุตสาหกรรม และธุรกิจ ได้ให้คำแนะนำการนำโมเดลจากงานวิจัยนี้ไปสร้างแอปพลิเคชันบนมือถือถึงสมาร์ตโฟนหรือแท็บเล็ต หรือสร้างระบบฝังตัว (Embedded System) บนอุปกรณ์และพัฒนาในเชิงอุตสาหกรรม ได้แก่ การทดสอบมาตรฐานแอปพลิเคชันสำหรับอุปกรณ์เคลื่อนที่ มาตรฐานผลิตภัณฑ์อุตสาหกรรม (มอก.) เช่น มอก./ISO 13485 การนำแอปพลิเคชันหรืออุปกรณ์ไปจำหน่ายในเชิงพาณิชย์ เช่น ขายแอปพลิเคชันใน Google Play Store บนระบบปฏิบัติการแอนดรอยด์ หรือ App Store บนระบบปฏิบัติการไอโอเอส ขายขาดให้บุคคล และขายหรือเช่าให้กับหน่วยงานหรือองค์กรต่างๆ เป็นต้น

4. กลุ่มผู้ใช้งานที่มีหน้าที่กำหนดนโยบาย เช่น กระทรวงสาธารณสุข กรมส่งเสริมและพัฒนาคุณภาพชีวิตคนพิการ กระทรวงการพัฒนาระบบสังคมและความมั่นคงของมนุษย์ ได้ให้

คำแนะนำการนำโมเดลที่ได้จากงานวิจัยนี้ไปกำหนดเป็นนโยบายของประเทศ เพื่อให้เกิดการพัฒนาคุณภาพชีวิตและความเป็นอยู่ของผู้พิการทางการได้ยินในประเทศไทย



## บรรณานุกรม

1. WHO global estimates on prevention of deafness and hearing loss [Internet] 2018 [Available from: <https://www.who.int/deafness/estimates/en/>].
2. Morton NE. Genetic epidemiology of hearing impairment. *Ann N Y Acad Sci*. 1991;630:16-31.
3. Kemper AR, Downs S. A Cost-effectiveness Analysis of Newborn Hearing Screening Strategies. *Archives of pediatrics & adolescent medicine*. 2000;154:484-8.
4. World Health Organization ROFS-EA. Situation review and update on deafness, hearing loss and intervention programmes. New Delhi: WHO Regional Office for South-East Asia; 2007 2007.
5. สำนักงานสถิติแห่งชาติ. สํารวจความพิการ [Internet] 2560 [Available from: <http://www.nso.go.th/sites/2014/Pages/Statistics%20from%20major%20Survey.aspx>].
6. Geers A, Brenner C, Davidson L. Factors Associated with Development of Speech Perception Skills in Children Implanted by Age Five. *Ear and hearing*. 2003;24:24S-35S.
7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60(6):84–90.
8. Srijanon K, Eiamkanitchat N, editors. Thai speech recognition using Neuro-fuzzy system. 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON); 2015 24-27 June 2015.
9. Gupta K, Gupta D, editors. An analysis on LPC, RASTA and MFCC techniques in Automatic Speech recognition system. 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence); 2016 14-15 Jan. 2016.
10. Hardt D, Fellbaum K. Spectral subtraction and RASTA-filtering in text-dependent HMM-based speaker verification1997. 867-70 vol.2 p.
11. Wijoyo S, editor *Speech Recognition Using Linear Predictive Coding and Artificial Neural Network for Controlling Movement of Mobile Robot*2011.

12. Mansour E, Sayed M, Moselhy AM, Abdelnaiem AA, editors. LPC and MFCC Performance Evaluation with Artificial Neural Network for Spoken Language Identification 2013.
13. Patil UG, Shirbahadurkar SD, Paithane AN, editors. Automatic Speech Recognition of isolated words in Hindi language using MFCC. 2016 International Conference on Computing, Analytics and Security Trends (CAST); 2016 19-21 Dec. 2016.
14. Imtiaz MA, Raja G, editors. Isolated word Automatic Speech Recognition (ASR) System using MFCC, DTW & KNN. 2016 Asia Pacific Conference on Multimedia and Broadcasting (APMediaCast); 2016 17-19 Nov. 2016.
15. Qian Y, Bi M, Tan T, Yu K. Very Deep Convolutional Neural Networks for Noise Robust Speech Recognition. IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2016;24(12):2263-76.
16. Adiwijaya, Aulia MN, Mubarak MS, Novia WU, Nhita F, editors. A comparative study of MFCC-KNN and LPC-KNN for hijaiyyah letters pronunciation classification system. 2017 5th International Conference on Information and Communication Technology (ICoICT7); 2017 17-19 May 2017.
17. Sangsai P, Detyong P. The Development of Aural Habilitation/Rehabilitation on Android Mobile Phone [Independent Study B.Eng.]. Nakhon Nayok: Bachelor of Engineering Srinakharinwirot University; 2017.
18. Swedia ER, Mutiara AB, Subali M, Ernastuti, editors. Deep Learning Long-Short Term Memory (LSTM) for Indonesian Speech Digit Recognition using LPC and MFCC Feature. 2018 Third International Conference on Informatics and Computing (ICIC); 2018 17-18 Oct. 2018.
19. Mini PP, Thomas T, Gopikakumari R, editors. Feature Vector Selection of Fusion of MFCC and SMRT Coefficients for SVM Classifier Based Speech Recognition System. 2018 8th International Symposium on Embedded Computing and System Design (ISED); 2018 13-15 Dec. 2018.
20. Liu Z, Wu Z, Li T, Li J, Shen C. GMM and CNN Hybrid Method for Short Utterance Speaker Recognition. IEEE Transactions on Industrial Informatics. 2018;14(7):3244-52.

21. Tantisatirapong S, Prasoproek C, Phothisonothai M, editors. Comparison of Feature Extraction for Accent Dependent Thai Speech Recognition System. 2018 IEEE Seventh International Conference on Communications and Electronics (ICCE); 2018 18-20 July 2018.
22. Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K. Speech Recognition Using Deep Neural Networks: A Systematic Review. IEEE Access. 2019;7:19143-65.
23. Puengsiricharoen K, Sirirattawan J, Sueaseenak D. An analysis of the Ling Six Sounds using the MFCC and DTW  
for a Speech-Rehabilitation System in Patients with Cochlear Implants. International Journal of Applied Biomedical Engineering. 2021;14:15-21.
24. Intimex. ประเภทของการสูญเสียการได้ยิน 2018 [Available from: <https://intimexchiangmai.com/types-of-hearing-loss/>].
25. Scollie S, Glista D, Tenhaaf J, Dunn A, Malandrino A, Keene K, et al. Stimuli and Normative Data for Detection of Ling-6 Sounds in Hearing Level. American journal of audiology. 2012;21.
26. กาญจนมงคล ณ. การฟื้นฟูการได้ยิน 2020 [Available from: <https://hearingaidsbestprice.com/2501/>].
27. Bionics A. The Ling Six Sound Check [Available from: <https://advancedbionics.com/content/dam/advancedbionics/Documents/libraries/Tools-for-Toddlers/tools-for-parents/The-Ling-Six-Sound-Check.pdf>].
28. Fast Fourier transform [Available from: <https://cp-algorithms.com/algebra/fft.html>].
29. Rao K, k e M. Speech Recognition Using Articulatory and Excitation Source Features2017.
30. Saha S. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way 2018 [Available from: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>].
31. Nicholson C. A Beginner's Guide to Neural Networks and Deep Learning [Available from: <https://wiki.pathmind.com/neural-network>].

32. Science D. Machine Learning ตอนที่ 2: Neural Network แบบจำลองของสมอง 2018 [Available from: <https://codeonthehill.com/machine-learning-2-neural-network/>].
33. L MP. Deep Learning แบบฉบับคนสามัญชน EP 1 : Neural Network History 2019 [Available from: <https://medium.com/mmp-li/deep-learning-แบบฉบับคนสามัญชน-ep-1-neural-network-history-f7789236a9a3>].
34. Artificial Neural Networks for Machine Learning – Every aspect you need to know about [Available from: <https://data-flair.training/blogs/artificial-neural-networks-for-machine-learning/>].
35. Singh A. Artificial Neural Network | Types | Feed Forward | Feedback | Structure | Perceptron | Machine Learning | Applications 2018 [Available from: <https://msatechnosoft.in/blog/artificial-neural-network-types-feed-forward-feedback-structure-perceptron-machine-learning-applications/>].
36. Salim AP, Laksitowening KA, Asror I, editors. Time Series Prediction on College Graduation Using KNN Algorithm. 2020 8th International Conference on Information and Communication Technology (ICoICT); 2020 24-26 June 2020.
37. Audacity. How Audacity Noise Reduction Works 2015 [Available from: [https://wiki.audacityteam.org/wiki/How\\_Audacity\\_Noise\\_Reduction\\_Works](https://wiki.audacityteam.org/wiki/How_Audacity_Noise_Reduction_Works)].
38. Sharma S. Activation Functions in Neural Networks 2017 [Available from: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>].
39. Wood T. What is the Softmax Function [Available from: <https://deeptai.org/machine-learning-glossary-and-terms/softmax-layer>].
40. Narkhede S. Understanding Confusion Matrix 2018 [Available from: <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>].
41. Categorical crossentropy [Available from: <https://peltarion.com/knowledge-center/documentation/modeling-view/build-an-ai-model/loss-functions/categorical-crossentropy>].
42. L MP. Deep Learning แบบฉบับสามัญชน EP 2 Optimization & Activation Function เรียบกันสบายๆสไตล์ชิคๆ 2019 [Available from: <https://medium.com/mmp-li/deep-learning-เรียบกันสบายๆสไตล์ชิคๆ>].

แบบฉบับสามัญชน-ep-2-optimization-activation-function-เรียนกันสบายๆสไตล์ศิลป์-  
9feb5a87e3b2.



## ประวัติผู้เขียน

ชื่อ-สกุล	คุณตม บึงศิริเจริญ
วัน เดือน ปี เกิด	25 กรกฎาคม 2540
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	พ.ศ. 2562 วิศวกรรมศาสตรบัณฑิต สาขาวิชาวิศวกรรมชีวการแพทย์ จาก มหาวิทยาลัยศรีนครินทรวิโรฒ
ที่อยู่ปัจจุบัน	11/88 ถนน ราชพฤกษ์ แขวง บางเขินกหนึ่ง เขต ดลิ่งชัน กรุงเทพมหานคร 10170

