

การระบุและรู้จำสายพันธุ์ของนกโดยใช้ทัศนียภาพของเสียง
แบบต่อเนื่องจากการบันทึกข้อมูลเสียงทางชีวภาพรอบทิศทาง
IDENTIFYING AND RECOGNIZING BIRD SPECIES
USING CONTINUOUS SOUNDSCAPES FROM
OMNIDIRECTIONAL BIOACOUSTIC RECORDINGS

โดย

นายเขตโสภณ ขุนพารเพิง
KATESOPON KUNPANPERNG

นางสาวบุษบงก์ ใจตรง
BUDSABONG JAITRONG

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ภาคเรียนที่ 1 ปีการศึกษา 2562

การระบุและรู้จำสายพันธุ์ของนกโดยใช้ทัศนียภาพของเสียง
แบบต่อเนื่องจากการบันทึกข้อมูลเสียงทางชีวภาพรอบทิศทาง
IDENTIFYING AND RECOGNIZING BIRD SPECIES
USING CONTINUOUS SOUNDSCAPES FROM
OMNIDIRECTIONAL BIOACOUSTIC RECORDINGS

โดย

นายเขตโสภณ ชุนพารเพิง

นางสาวบุษบงก์ ใจตรง

อาจารย์ที่ปรึกษา

รองศาสตราจารย์ ดร. อีรพงศ์ ลีลานุกภาพ

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

คณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 1 ปีการศึกษา 2562

**IDENTIFYING AND RECOGNIZING BIRD SPECIES
USING CONTINUOUS SOUNDSCAPES FROM
OMNIDIRECTIONAL BIOACOUSTIC RECORDINGS**

KATESOPON KUNPANPERNG

BUDSABONG JAITRONG

**A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

1/2020

ใบรับรองปริญญาโท ประจำปีการศึกษา 2562
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การระบุและรู้จำสายพันธุ์ของนกโดยใช้ทัศนียภาพของเสียง
แบบต่อเนื่องจากการบันทึกข้อมูลเสียงทางชีวภาพรอบทิศทาง
IDENTIFYING AND RECOGNIZING BIRD SPECIES
USING CONTINUOUS SOUNDSCAPES FROM
OMNIDIRECTIONAL BIOACOUSTIC RECORDINGS

ผู้จัดทำ

1. นายเขตโสภณ ชุนพารเพ็ง รหัสประจำตัว 60070127
2. นางสาวบุษบงก์ ใจตรง รหัสประจำตัว 60070145

..... อาจารย์ที่ปรึกษา
รองศาสตราจารย์ ดร. ชีรพงศ์ ลีลานุภาพ

ใบรับรองโครงการ (PROJECT)

เรื่อง

การระบุและรู้จำสายพันธุ์ของนกโดยใช้ทัศนียภาพของเสียงแบบต่อเนื่องจา
การบันทึกข้อมูลเสียงทางชีวภาพรอบทิศทาง

**IDENTIFYING AND RECOGNIZING BIRD SPECIES USING
CONTINUOUS SOUNDSCAPES FROM OMNIDIRECTIONAL
BIOACOUSTIC RECORDINGS**

นายเขตโสภณ ขุนพารเพิง รหัสประจำตัว 60070127
นางสาวบุษบงก์ ใจตรง รหัสประจำตัว 60070145

ขอรับรองว่ารายงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกมาจากที่ใด
รายงานฉบับนี้ได้รับการตรวจสอบและอนุมัติให้เป็นส่วนหนึ่งของ
การศึกษาวิชาโครงการ หลักสูตรวิทยาศาสตรบัณฑิต (เทคโนโลยีสารสนเทศ)
ภาคเรียนที่ 1 ปีการศึกษา 2562

.....

นายเขตโสภณ ขุนพารเพิง

.....

นางสาวบุษบงก์ ใจตรง

หัวข้อโครงงาน	การระบุและรู้จำสายพันธ์ของนกโดยใช้ทัศนียภาพของเสียงแบบต่อเนื่องจาการบันทึกข้อมูลเสียงทางชีวภาพรอบทิศทาง
นักศึกษา	นายเขตโสภณ ขุนพารเพิง นางสาวบุษบงก์ ใจตรง
ปริญญา	วิทยาศาสตรบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
ปีการศึกษา	2562
อาจารย์ที่ปรึกษา	รองศาสตราจารย์ ดร. อีรพงศ์ ลีลานุภาพ

บทคัดย่อ

บทคัดย่อ

Project Title	IDENTIFYING AND RECOGNIZING BIRD SPECIES USING CONTINUOUS SOUNDSCAPES FROM OMNIDIRECTIONAL BIOACOUSTIC RECORDINGS
Student	Katesopon kunpanperng Budsabong Jaitrong
Degree	Bachelor of Science
Program	Information Technology
Academic Year	2020
Advisor	Assoc. Prof. Teerapong Leelanupab(Ph.D.)

Abstract

Abstract eng

กิตติกรรมประกาศ

แก้ไขกิตติกรรมประกาศในไฟล์ acknowledgement.tex

นายเขตโสภณ ขุนพารเพิง
นางสาวบุษบงก์ ใจตรง

สารบัญ

	หน้า
บทคัดย่อ	I
บทคัดย่อ ภาษาอังกฤษ	II
กิตติกรรมประกาศ	III
สารบัญ	IV
สารบัญตาราง	V
สารบัญรูป	VI
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญ	1
1.2 วัตถุประสงค์	2
1.3 วิธีการดำเนินการ	2
1.4 ขอบเขตของงาน	3
1.5 ประโยชน์ที่คาดว่าจะได้รับ	4
บทที่ 2 การทบทวนวรรณกรรมที่เกี่ยวข้อง	5
2.1 เสียงและสัญญาณ (Sound and Signal)	5
2.2 ทฤษฎีการประมวลผลภาพดิจิทัล (Digital Image Processing)	7
2.3 การเรียนรู้ของเครื่อง (Machine Learning)	7
2.4 โครงข่ายประสาทเทียม (Artificial Neural Network)	9
2.5 สถาปัตยกรรมต่าง ๆ ของโครงข่ายประสาทแบบคอนโวลูชันและการกำหนดองค์ประกอบ (CNN Architecture and Configuration)	16
2.6 เมตริกที่ใช้ประเมินผลแบบจำลอง (Evaluation Metrics)	16
บทที่ 3 วิธีการดำเนินการวิจัย	19
3.1 บทนำและงานวิจัยที่เกี่ยวข้อง	19
3.2 การจัดเตรียมการทดลอง	21
บทที่ 4 ผลการทดลองเบื้องต้นหรือระบบต้นแบบ	24
บทที่ 5 บทสรุป	25
บรรณานุกรม	26
ภาคผนวก ก เรื่องที่หนึ่ง	28
ประวัติผู้เขียน	30

สารบัญตาราง

ตารางที่

หน้า

สารบัญรูป

รูปที่	หน้า
2.1 แสดงกราฟของฟังก์ชันไซน์	5
2.2 แสดงสัญญาณซ้ำคาบ	6
2.3 แสดงตัวอย่างภาพสเปกโตรแกรม	7
2.4 แสดงความแตกต่างระหว่างการเขียนโปรแกรมในอิตาลีกับการเรียนรู้ของเครื่อง	8
2.5 แสดงโครงสร้างของโครงข่ายประสาทเทียม	9
2.6 แสดงการทำงานของเซลล์ประสาทเทียบกับ Activation Function	10
2.7 รูปเปรียบเทียบระหว่าง Sigmoid และ ReLU Function	10
2.8 แสดงโครงสร้างของโครงข่ายประสาทเชิงลึก	11
2.9 แสดงโครงข่ายประสาทเทียม	11
2.10 แสดงสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชัน	13
2.11 แสดงตัวกรองขนาด 3 X 3	13
2.12 แสดงตัวอย่างของข้อมูลภาพขาเข้า ตัวกรอง และฟังก์ชันลักษณะ	14
2.13 แสดงตัวอย่างลักษณะการเคลื่อนที่ของตัวกรองเพื่อให้ได้ผลลัพธ์ดังภาพ 2.12	14
2.14 แสดงการเคลื่อนที่ของตัวกรองเมื่อ Stride มีค่าเท่ากับ 1 และแสดงผลลัพธ์ของฟังก์ชันลักษณะที่ได้	14
2.15 แสดงการเคลื่อนที่ของตัวกรองเมื่อ Stride มีค่าเท่ากับ 2 และแสดงผลลัพธ์ของฟังก์ชันลักษณะที่ได้	15
2.16 แสดงข้อมูลรูปภาพที่มีการทำ Padding และแสดงของฟังก์ชันลักษณะที่ได้หลังจากการทำ Padding	15
2.17 แสดงการทำ Max Pooling ที่มีขนาดตัวกรองเท่ากับ 2 x 2 และ stride เท่ากับ 2	16
2.18 แสดงความสัมพันธ์ของ AUC และ ROC Curve	17
3.1 โฟล์เดอร์ที่ใช้ในการเก็บไฟล์เสียง	22
3.2 ตัวอย่างไฟล์เสียงที่ถูกเก็บไว้ในโฟล์เดอร์	22
3.3 ไฟล์เนลยการทำงานแบบจำลอง	23

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญ

การสำรวจและตรวจสอบสายพันธุ์ของนกโดยอาศัยเสียงของนกมาเป็นตัวบ่งบอกถึงความแตกต่างทางสายพันธุ์นั้น มีความสำคัญอย่างมากต่อจุดประสงค์ของการวิจัยทางด้านวิทยาศาสตร์ในหลายๆด้าน ซึ่งองค์ความรู้ที่ใช้ในการระบุสายพันธุ์ของนกอย่างแม่นยำนั้น สามารถช่วยให้เราเข้าใจถึงการกระจายทางภูมิศาสตร์และเรื่องการของวิวัฒนาการของสายพันธุ์นกได้ด้วย และเรียกได้ว่าเป็นสิ่งสำคัญจำเป็นสำหรับการอนุรักษ์ความหลากหลายทางชีวภาพที่มีให้คงอยู่อย่างยั่งยืน ซึ่งผู้ที่สามารถนำองค์ความรู้ตรงนี้ไปใช้ต่อยอดได้มีอยู่อย่างมากมาย ไม่ว่าจะเป็นผู้ที่มีความต้องการในการค้นคว้าเรื่องความหลากหลายทางชีวภาพ ผู้ที่มีความสนใจในความต่างของเสียงนกแต่ละสายพันธุ์ และแน่นอนนักปักษีวิทยาย่อมเป็นหนึ่งในผู้ที่สามารถใช้องค์ความรู้ในส่วนนี้ได้ อย่างมีประสิทธิภาพมากที่สุด

การระบุสายพันธุ์ของนกโดยใช้การรวบรวมข้อมูลโดยใช้เสียงของนกแต่ละสายพันธุ์มารวบรวมถือว่าเป็นวิธีที่ดี ง่าย และทำได้ง่ายมีประสิทธิภาพมากกว่าการใช้รูปถ่ายของนกมาก [1] เนื่องจากการถ่ายภาพนกเก็บไว้ตามสถานที่ต่างๆ ให้ข้อมูลที่เยอะและแม่นยำซึ่งถือว่าเป็นเรื่องที่ยากและท้าทายมาก และในการเก็บข้อมูลเสียงของนกนั้นมีความเป็นไปได้ที่สูงกว่าในการเก็บข้อมูลเสียงของนกที่มีความละเอียด ไม่เปลืองทรัพยากรในการเก็บ และได้รับข้อมูลมาอย่างครอบคลุมกว่าการถ่ายภาพนกเก็บไว้อย่างแน่นอน

การมีส่วนร่วมกันของโครงการทางวิทยาศาสตร์หลายๆโครงการ อย่างเช่น Xeno-Canto [2] มหาวิทยาลัยเทคโนโลยีเคมีนิทซ์ (Chemnitz University of Technology) และที่อื่นๆ ได้มุ่งเน้นไปที่การบันทึกเสียงของนกไว้เป็นจำนวนมากใหญ่ เพื่อเพิ่มความเป็นไปได้โดยรวมทั้งหมดในการรับรู้สายพันธุ์ของนกผ่านการฟังเสียงของพวกมัน และเพื่อที่จะนำข้อมูลที่ได้ไปใช้ในการสร้างแบบจำลองเชิงลึก (Deep Learning Model) เพื่อทำให้กระบวนการในการรับรู้และจำแนกเสียงนกสามารถดำเนินการไปได้แบบอัตโนมัติ เช่นการสร้างแอปพลิเคชันในการระบุเสียงนกที่ได้รับข้อมูลเข้ามา การคงไว้ซึ่งความหลากหลายทางชีวภาพ และการอนุรักษ์นกที่ใกล้สูญพันธุ์ผ่านการ ฟังเสียงและติดตามว่ายังคงมีนกชนิดนั้น ๆอยู่

ความท้าทายในการจำแนกเสียงนกในอดีตได้รับข้อมูลเสียงนกในการเดินทางแบบทิศทางเดียว (Mono-directional recording) จากเครื่องมือที่ใช้ในการบันทึกเสียงในขณะนั้น [3] และองค์ความรู้และระบบการระบุเสียงนกในครั้งนั้นสามารถระบุออกมาได้เป็นอย่างดีและได้ถูกนำไปพัฒนาต่อยอดเป็นแอปพลิเคชันอย่างมากมายในปัจจุบัน ตัวอย่างเช่นแอปพลิเคชันที่ชื่อ BirdGenie [4] ที่มีความสามารถในการบันทึกเสียงของนกในขณะนั้น และบอกมาว่านกชนิดนั้นคือนกสายพันธุ์อะไรซึ่งมีจำนวนนกที่บอกได้ทั้งหมดมากกว่า 100 สายพันธุ์ด้วย แต่อย่างไรก็ตามก็ยังคงมีการให้ความสนใจอย่างมากในการระบุสายพันธุ์ของนกจากข้อมูลเสียงนกแบบหลายทิศทาง (Omnidirectional) ซึ่งสิ่งนี้จะช่วยให้มีการตรวจสอบสภาพแวดล้อมและเสียงของสิ่งรอบข้างได้อย่างแม่นยำมากขึ้น และข้อดีของวิธีการนี้คือจำนวนของความเป็นอคติที่มีอยู่ในการสุ่มตัวอย่าง (Sampling bias) จะน้อยกว่าอคติที่มาจากการสำรวจและเก็บข้อมูลโดยนักวิทยาศาสตร์ทั่วไป อย่างไรก็ตามการรับรู้เสียงนกในสภาพแวดล้อมที่มีกิจกรรมจากเสียงรอบข้างอยู่ มีสัญญาณเสียงที่ซ้อน

ทับกัน และมีเสียงรบกวนในระดับสูงอาจทำให้มีความยุ่งยากในรับรู้และระบุสายพันธุ์นกได้

จากที่กล่าวมาข้างต้นผู้วิจัยจะสร้างแบบจำลองในการคาดคะเนสายพันธุ์ของนกแต่ละตัวขึ้นมา โดยใช้ลักษณะต่าง ๆ ที่สามารถสกัดมาได้จากข้อมูลเสียงของนก เพื่อนำมาเรียนรู้และสร้างแบบจำลองโครงข่ายประสาทเชิงลึกแบบคอนโวลูชัน (Deep Convolutional Neural Network) ซึ่งมีความเป็นไปได้ว่าจะให้ประสิทธิภาพในการคาดคะเนสายพันธุ์นกได้อย่างแม่นยำ

1.2 วัตถุประสงค์

1. เพื่อศึกษาโครงสร้างของแบบจำลองที่ใช้ในการจำแนกชนิดของนกด้วยไฟล์ที่ได้ทำการบันทึกเสียงและระบุสายพันธุ์ของนกไว้
2. เพื่อจำแนกและระบุชนิดของนกจากไฟล์ที่บันทึกเสียงของนกไว้เพื่อให้ง่ายต่อการติดตามและเป็นประโยชน์ต่อบัณฑิตศึกษา หรือเป็นประโยชน์ต่อนักชีววิทยาผู้ซึ่งมีความสนใจและค้นคว้าเกี่ยวกับเสียงของนก
3. เพื่อใช้ในการปรับปรุงและพัฒนาระบบตรวจจับเสียงอัตโนมัติที่สามารถจดจำ และจำแนกเสียงของนกที่มาจากการบันทึกเสียงในรูปแบบที่หลากหลายได้

1.3 วิธีการดำเนินการ

1. วางแผนการดำเนินงาน
 - กำหนดวัตถุประสงค์และขอบเขตของโครงการงาน
 - กำหนดแบบจำลองที่ใช้ในการทำนาย
 - แบบจำลองแบบอินเซปชัน (Inception model)
 - ImageNet [5]
 - กำหนดเครื่องมือที่ใช้ในการสร้างแบบจำลอง
2. ศึกษาเครื่องมือที่ใช้และทฤษฎีที่เกี่ยวข้อง
 - ศึกษาทฤษฎีต่างๆที่เกี่ยวข้องกับการสร้างแบบจำลอง
 - ทฤษฎีการเรียนรู้เชิงลึก (Deep Learning) โครงข่ายประสาทเชิงลึก (Deep Neural Network) โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network)
 - โครงข่ายประสาทแบบซ่อนๆกัน (Recurrent neural network) หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory)
 - ศึกษาหลักการทำงานของเครื่องมือที่เลือกใช้
 - PyTorch [6]
 - Keras
 - Librosa [7]
 - Sklearn
3. ดำเนินการจัดการข้อมูลและสร้างแบบจำลอง

- รวบรวมข้อมูลและจัดการข้อมูลให้อยู่ในรูปแบบที่เหมาะสม
- สำรวจและตรวจค้นข้อมูล (Data Exploration)
- ดำเนินการแยกคุณสมบัติที่สามารถหาได้จากข้อมูลที่เป็นเสียง (Feature Extraction)

4. ประเมินผลแบบจำลอง

1.4 ขอบเขตของงาน

ขอบเขตของงานถูกแบ่งออกเป็น 3 ส่วนคือ ขอบเขตของข้อมูลที่ใช้ในการสร้างและฝึกแบบจำลอง ขอบเขตของข้อมูลที่ใช้สำหรับการประเมินประสิทธิภาพของแบบจำลอง และขอบเขตการวิจัยในการสร้างแบบจำลองโครงข่ายประสาทเชิงลึกแบบคอนโวลูชัน (Deep Convolutional Neural Network)

1. ขอบเขตของข้อมูลที่นำมาใช้ในการสร้างและฝึกแบบจำลอง

- ชุดข้อมูลตัวเก่าที่ได้มาจากปี 2019 ที่ได้ถูกนำไปเพิ่มคุณสมบัติของชุดข้อมูล โดยมีเครือข่าย Xeno-canto [2] เป็นผู้มีส่วนร่วมในการช่วยจัดหาชุดข้อมูลและเพิ่มส่วนขยายทางภูมิศาสตร์ภายในตัวในชุดข้อมูลด้วย
 - ชุดข้อมูลที่ใช้ในการฝึกและสอนแบบจำลอง ประกอบไปด้วยเสียงบันทึกของนกชนิดต่าง ๆ จากทั้งอเมริกาเหนือ อเมริกาใต้และยุโรป โดยได้รับการสนับสนุนจาก Xeno-canto เป็นเสียงบันทึกที่มีคุณภาพสูงมากกว่า 70,000 รายการครอบคลุม 961 สายพันธุ์ และภายในแต่ละเสียงบันทึกจะมีข้อมูลเกี่ยวกับตำแหน่งที่บันทึก วันที่และคำอธิบายอื่น ๆ ของผู้บันทึกด้วย
 - มีเมตาดาต้า (metadata) ประกอบคู่กันกับไฟล์เสียงของนกที่ได้มาด้วย ซึ่งภายในเมตาดาต้าจะประกอบด้วยข้อมูลคร่าว ๆ เกี่ยวกับการเก็บข้อมูลในครั้งนั้น ๆ มา เช่น สถานที่ที่ใช้ในการอัดเสียงนกที่ได้ วันที่ที่เก็บ และคุณภาพของเสียงที่ได้

2. ขอบเขตของข้อมูลที่ใช้สำหรับการประเมินประสิทธิภาพของแบบจำลอง

- ข้อมูลที่ใช้ในการตรวจสอบประสิทธิภาพและการทำงานของแบบจำลอง (Validation data)
 - ชุดข้อมูลที่ใช้ในการตรวจสอบประสิทธิภาพของแบบจำลอง ประกอบไปด้วยไฟล์เสียงที่บันทึกในประเทศเปรู และสหรัฐอเมริกาทั้งหมด 12 ไฟล์และแต่ละไฟล์จะมีรายละเอียดต่อวินาที (sample rate) อยู่ที่ 32 kHz และแต่ละไฟล์จะมีระยะเวลาถึง 10 นาที
- ข้อมูลที่ใช้ในการทดสอบแบบจำลอง
 - ชุดข้อมูลที่ใช้ในการทดสอบประกอบด้วยเสียงบันทึกจำนวน 153 ไฟล์ที่บันทึกในประเทศสหรัฐอเมริกาและเยอรมนี โดยแต่ละเสียงบันทึกจะมีระยะเวลา 10 นาที และมีการซ้อนทับกันของเสียงนกในปริมาณมาก

3. ขอบเขตการวิจัยในการสร้างแบบจำลองโครงข่ายประสาทเชิงลึกแบบคอนโวลูชัน

- ผู้สร้างแบบจำลองต้องการที่จะให้แบบจำลองที่สร้างขึ้นมานั้น สามารถที่จะระบุสายพันธุ์ของนก ตามเสียงที่ได้มาอย่างถูกต้องและใกล้เคียงกับความเป็นจริงให้ได้มากที่สุด และด้วยความที่มีข้อมูลเสียงยกอยู่อย่างเป็นจำนวนมาก จึงมีความจำเป็นที่จะต้องแยกคุณสมบัติของเสียงออกมาจาก ไฟล์เสียงที่ได้เป็นคุณสมบัติต่างๆไม่ว่าจะเป็นขนาดของเสียงตามระยะเวลาที่ได้ และการตัดภาพจากกราฟเสียงที่ได้ออกมาเป็นส่วนใหญ่ ๆ ผ่านการใช้ Librosa [7] เพื่อนำไปเรียนรู้บนแบบจำลองที่สร้างบนสถาปัตยกรรมต่างๆ ตามโครงข่ายประสาทเชิงลึกแบบคอนโวลูชันไม่ว่าจะเป็น ImageNet Inception และอื่น ๆ โดยแบบจำลองจะถูกสร้างโดยใช้ PyTorch และ Keras เพื่อให้ได้ผลลัพธ์ออกมาดีที่สุด และได้สร้างแบบจำลองออกมาเป็นจำนวนมาก เพื่อเปรียบเทียบประสิทธิภาพการทำงานของแต่ละแบบจำลอง และนำแบบจำลองที่ได้ผลลัพธ์ดีที่สุดมาใช้ต่อไป

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ระบบสามารถจดจำเสียงนกและแยกประเภทของนกได้อย่างถูกต้อง
2. สามารถเลือกแบบจำลอง และสถาปัตยกรรมเพื่อนำมาใช้กับข้อมูลที่มีความแตกต่างออกไปตามการใช้งานได้
3. สามารถนำข้อมูลนกที่แยกประเภทจากระบบแล้วไปใช้ประโยชน์ในการติดตามและตรวจสอบสุขภาพของระบบนิเวศ หรือเป็นประโยชน์ในการศึกษาของนักนกวิทยา
4. ระบบและแบบจำลองที่สร้างขึ้น สามารถนำไปแยกสายพันธุ์ของนกจากข้อมูลไฟล์เสียงบันทึกใหม่ๆ ที่ถูกนำเข้าไปได้ในตัวระบบได้

บทที่ 2

การทบทวนวรรณกรรมที่เกี่ยวข้อง

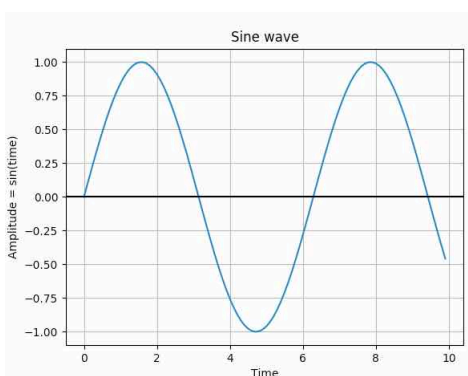
ในบทที่ 2 จะอธิบายถึงทฤษฎีที่เกี่ยวข้องในการสร้างแบบจำลองการระบุสายพันธุ์นก ซึ่งประกอบด้วย ทฤษฎีเสียงและสัญญาณ (Sound and Signal) ทฤษฎีการประมวลผลภาพดิจิทัล (Digital Image Processing) ทฤษฎีการเรียนรู้ของเครื่อง (Machine Learning) โครงข่ายประสาทเทียม (Artificial Neural Network) โครงข่ายประสาทเชิงลึก (Deep Neural Network) โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network, CNN) สถาปัตยกรรมต่าง ๆ ของโครงข่ายประสาทแบบคอนโวลูชันและการกำหนดองค์ประกอบ (CNN Architecture and Configuration) เมตริกที่ใช้ประเมินผลแบบจำลอง (Evaluation Metrics) และหลักการทำงานของเครื่องมือต่าง ๆ ที่ใช้ในงานวิจัย

2.1 เสียงและสัญญาณ (Sound and Signal)

เสียงเป็นพลังงานรูปแบบหนึ่งที่เกิดจากการสั่นสะเทือนของแหล่งกำเนิดเสียงผ่านตัวกลางและรับรู้ได้ด้วยหู ส่วนสัญญาณเป็นการแสดงถึงปริมาณที่แปรผันตามเวลา ซึ่งมีค่าจำกัดความที่ค่อนข้างเป็นนามธรรม และสัญญาณเสียงจะแสดงถึงความแปรปรวนของความกดอากาศเมื่อเวลาผ่านไป โดยมีไมโครโฟนที่ทำหน้าที่ในการวัดความแปรผันและสร้างสัญญาณไฟฟ้าที่เป็นตัวแทนของเสียง และมีลำโพงที่ทำหน้าที่ในการรับสัญญาณไฟฟ้าและทำให้เกิดเสียง ซึ่งทั้งไมโครโฟนและลำโพงจะถูกเรียกว่า ทรานสดิวเซอร์ (transducer) เนื่องจากทั้งไมโครโฟนและลำโพงเป็นอุปกรณ์ที่ทำหน้าที่รับพลังงานจากรูปแบบหนึ่ง แล้วแปลงไปให้อยู่ในอีกรูปแบบหนึ่ง

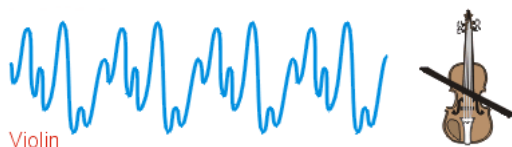
2.1.1 สัญญาณซ้ำคาบ (Periodic signals)

เป็นลักษณะของสัญญาณที่มีรูปแบบการเกิดที่ซ้ำกันในช่วงเวลาที่เท่า ๆ กัน เรียกการวนครบรูปแบบของสัญญาณหนึ่งครั้งว่ารอบ (cycle) และเรียกระยะเวลาในแต่ละรอบว่าคาบ (period) สัญญาณนี้จะมีลักษณะคล้ายกับกราฟของฟังก์ชันไซน์ ดังภาพ 2.1



รูปที่ 2.1 แสดงกราฟของฟังก์ชันไซน์

ความถี่ของสัญญาณมีหน่วยเป็นจำนวนของรอบต่อวินาที ซึ่งเป็นค่าที่ตรงกันข้ามกับคาบ ซึ่งมีหน่วยคือ รอบต่อวินาที หรือ เฮิรตซ์ (Hz) ในเครื่องดนตรีส่วนใหญ่จะให้สัญญาณซ้ำคาบเพียงแต่ไม่ได้อยู่ในรูปแบบของสัญญาณไซน์ (sinusoidal) ตัวอย่างเช่นสัญญาณของเสียงไวโอลิน



รูปที่ 2.2 แสดงสัญญาณซำคาบ

เราเรียกรูปร่างของสัญญาณซำคาบเหล่านี้ว่า รูปแบบของคลื่น (waveform) เครื่องดนตรีส่วนใหญ่จะสร้างรูปแบบของคลื่นที่ซับซ้อนมากกว่าสัญญาณไซน์ รูปแบบของคลื่นจะเป็นเครื่องกำหนดลักษณะของเสียงร้องหรือเสียงดนตรี ดังภาพ 2.2 ซึ่งเป็นการรับรู้เกี่ยวกับคุณภาพของเสียง และมนุษย์มักจะรับรู้เสียงที่มีความซับซ้อนมากกว่ารูปแบบคลื่นของไซน์

2.1.2 สัญญาณไม่ซำคาบ (Non-periodic signals)

เป็นสัญญาณที่มีการเปลี่ยนแปลงที่ไม่สามารถระบุรูปแบบที่แน่นอนของสัญญาณได้ เช่นเสียงการพูดคุยของมนุษย์ สัญญาณเหล่านี้สามารถมองเห็นได้โดยการนำเสนอในรูปแบบของสเปกโตรแกรม (spectrograms)

2.1.3 การสุ่มตัวอย่าง (Sampling)

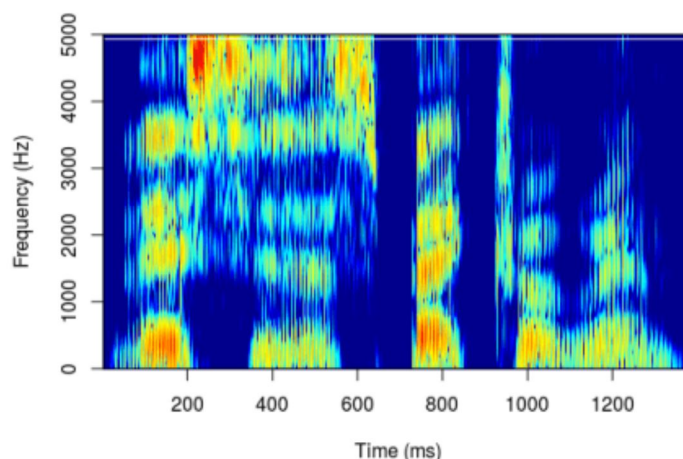
การสุ่มตัวอย่างคือการเปลี่ยนแปลงสัญญาณที่มีความต่อเนื่องทางเวลา (Continuous signal) ให้อยู่ในรูปแบบที่ไม่ต่อเนื่องทางเวลา (Discrete signal) ด้วยการสุ่มเก็บตัวอย่างของสัญญาณในช่วงเวลาที่ห่างเท่า ๆ กัน ซึ่งช่วงเวลาดังกล่าวจะถูกเรียกว่า อัตราสุ่มสัญญาณ (Sampling rate) ในการสุ่มสัญญาณจำเป็นต้องเลือกอัตราการสุ่มให้เหมาะสมกับความถี่ของสัญญาณนั้น ๆ เนื่องจากในเวลาที่ต้องการแปลงสัญญาณกลับไปเป็นสัญญาณที่มีความต่อเนื่องทางเวลาจะได้สัญญาณต้นฉบับที่ถูกต้องและครบถ้วน [8]

2.1.4 สเปกโตรแกรม (Spectrograms)

สเปกโตรแกรมเป็นภาพหรือแผนภาพของสเปกตรัม (Spectrum) โดยที่สเปกตรัมเป็นแถบคลื่นความถี่ของเสียงที่รวมกับแสงที่มนุษย์สามารถมองเห็นได้ โดยสเปกโตรแกรมทำให้สามารถมองเห็นผลลัพธ์ของการแบ่งเสียงออกเป็นส่วน ๆ ได้ โดยเรียกการเปลี่ยนของเสียงในแต่ละส่วนละส่วนเป็นสเปกโตรแกรมนี้จะถูกเรียกว่า Short-Time Fourier Transform (STFT) โดยที่แกน x ของสเปกโตรแกรมจะแสดงเวลาและแสดงความถี่บนแกน y โดยในแต่ละคอลัมน์จะแสดงสเปกตรัมของส่วนนั้น ๆ โดยใช้สีในการแสดงแอมพลิจูด (Amplitude) หรือความเข้มของเสียงนั้น

2.1.5 สัญญาณรบกวน (Noise)

เป็นเสียงหรือสัญญาณที่ไม่พึงประสงค์ที่อาจเกิดจากการทำงานผิดพลาดของอุปกรณ์ที่ใช้ในการเก็บข้อมูล หรือสร้างขึ้นเพื่อศึกษาลักษณะ รูปแบบ และทำความเข้าใจกับสัญญาณรบกวนเหล่านี้ได้ โดยสัญญาณรบกวนที่สร้างง่ายที่สุดถูกเรียกว่า เสียงสัญญาณรบกวนที่ไม่เกี่ยวข้องกันแบบสม่ำเสมอ (Uncorrelated Uniform Noise) หรือ UU Noise โดยที่ Uniform หมายถึงสัญญาณจะมีการสุ่มค่าจากการกระจายที่สม่ำเสมอ นั่นคือทุก ๆ ค่าในช่วงนั้นมีโอกาสที่จะเท่ากัน และ uncorrelated ที่หมายถึงไม่เกี่ยวข้องกันหรือเป็นอิสระต่อกัน กล่าวคือการรู้ค่าเพียงค่าเดียวไม่สามารถให้ข้อมูลอื่น ๆ ได้ และมี 3 สิ่งที่ควรทราบเกี่ยวกับสัญญาณรบกวนและสเปกตรัมคือ 1.



รูปที่ 2.3 แสดงตัวอย่างภาพสเปกโตรแกรม

การกระจายของสัญญาณสุ่ม 2. ความสัมพันธ์ของแต่ละค่าในสัญญาณเป็นอิสระจากกันหรือไม่ และ 3. ความสัมพันธ์ระหว่างกำลัง (Power) กับความถี่ จากสามสิ่งนี้ที่ควรทราบนี้จะสามารถสร้างสัญญาณรบกวนได้ 3 ลักษณะคือ สัญญาณรบกวนบราวน์ (Brownian noise) หรือสัญญาณรบกวนสีแดง (Red noise) สัญญาณรบกวนสีชมพู (Pink Noise) และ Gaussian Noise หรือสัญญาณรบกวนสีขาว สัญญาณรบกวนโดยปกติจะถูกสังเคราะห์ขึ้นจากสมการ 2.1

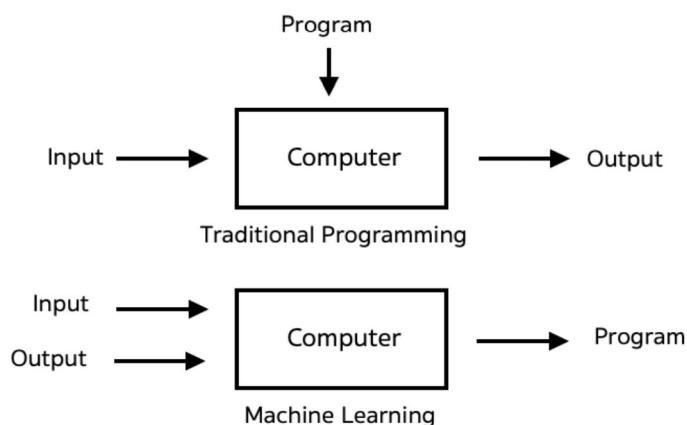
$$P = \frac{K}{f^\beta} \quad (2.1)$$

โดยที่ P คือกำลัง f เป็นค่าของความถี่และ K คือจุดตัดของเส้นตรง จากสมการข้างต้น เราสามารถทราบความสัมพันธ์ระหว่างกำลังและความถี่ของสัญญาณรบกวนต่าง ๆ ได้ คือเมื่อ $P = K/f^2$ ผลลัพธ์ของสัญญาณรบกวนนี้จะตกอยู่ในช่วงของสเปกตรัมแสงสีแดง ซึ่งมีความถี่ที่ต่ำที่สุด ในขณะที่ถ้าเราเปลี่ยนค่าของ β ให้เท่ากับ 0 สัญญาณรบกวนที่เกิดขึ้นจะตกอยู่ในช่วงของสเปกตรัมแสงสีขาว ทำให้เกิดเป็นสัญญาณรบกวนสีขาว และเมื่อค่า β มีค่าอยู่ในระหว่าง 0 ถึง 2 ผลลัพธ์ของสัญญาณรบกวนจะอยู่ระหว่างช่วงสเปกตรัมแสงสีขาวและสีแดง จึงเรียกสัญญาณรบกวนที่อยู่ในช่วงนี้ว่าสัญญาณรบกวนสีชมพู

2.2 ทฤษฎีการประมวลผลภาพดิจิทัล (Digital Image Processing)

2.3 การเรียนรู้ของเครื่อง (Machine Learning)

การเรียนรู้ของเครื่องคือการที่คอมพิวเตอร์สามารถที่จะเรียนรู้ได้ด้วยตัวเอง โดยที่ไม่ต้องเขียนโปรแกรมสั่งให้คอมพิวเตอร์ทำงาน ซึ่งจะแตกต่างกับการเขียนโปรแกรมในสมัยก่อน ซึ่งในสมัยก่อนโปรแกรมเมอร์ต้องทำการโปรแกรมไปยังคอมพิวเตอร์เพื่อให้ได้ผลลัพธ์ออกมา แต่ในการเรียนรู้ของเครื่องผู้ใช้งานเพียงแค่อัปโหลดและระบุผลลัพธ์ที่ต้องการเท่านั้น คอมพิวเตอร์จะทำการเรียนรู้ด้วยตนเองแล้วแสดงผลที่ได้ออกมาเป็นโปรแกรมที่สามารถนำไปใช้ในการทำนายผลลัพธ์ของข้อมูลได้ดังรูป 2.4



รูปที่ 2.4 แสดงความแตกต่างระหว่างการเขียนโปรแกรมในอดีตกับการเรียนรู้ของเครื่อง

โดยการเรียนรู้ของเครื่องนั้นจะประกอบไปด้วยข้อมูลและเครื่องมือทางสถิติที่ใช้ในการทำนายผลลัพธ์หรือหารูปแบบของข้อมูลที่เข้าไป เครื่องจะทำการเรียนรู้ที่มีลักษณะคล้ายกันกับมนุษย์คือ เรียนรู้จากประสบการณ์ ยิ่งมีประสบการณ์มากก็จะมีผู้เชี่ยวชาญมากทำให้ง่ายต่อการทำนายผลลัพธ์ที่เกิดขึ้นจากสถานการณ์ที่คล้าย ๆ กันได้ และเพื่อเพิ่มความแม่นยำในการทำนายผลลัพธ์ เครื่อง (Machine) ก็ควรจะมีการฝึกด้วยข้อมูลจำนวนมากเพียงพอ ที่เครื่องจะสามารถค้นพบผลลัพธ์ผ่านรูปแบบหรือแบบแผนที่ซ้ำ ๆ กันได้ และการเรียนรู้ของเครื่องนั้นสามารถแบ่งรูปแบบการเรียนรู้ได้เป็น 2 ลักษณะใหญ่ ๆ คือ

2.3.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning)

การเรียนรู้แบบมีผู้สอนคือการที่คอมพิวเตอร์สามารถหาผลลัพธ์ของข้อมูลได้ด้วยตัวเองหลังจากที่มีการเรียนรู้จากชุดข้อมูลตัวอย่างไปแล้วระยะเวลาหนึ่ง โดยเริ่มต้นโปรแกรมเมอร์จะเขียนโปรแกรมให้คอมพิวเตอร์สร้างแบบจำลอง (Model) ขึ้นมาจากชุดข้อมูลและผลลัพธ์ตัวอย่าง โดยที่ถ้าชุดข้อมูลตัวอย่างมีความหลากหลายและมีจำนวนมากอาจทำให้ผลลัพธ์ที่ได้จากการเรียนรู้นั้นมีความแม่นยำมากขึ้น กระบวนการนี้จะถูกเรียกว่าการฝึกอบรม(Train) เมื่อผลลัพธ์ที่ได้มีความแม่นยำมากขึ้น จะกล่าวได้ว่าแบบจำลองที่ได้จากการฝึกอบรมนี้มีความสามารถในการทำนายผลลัพธ์ได้ถูกต้องมากขึ้น การเรียนรู้แบบมีผู้สอนสามารถแบ่งออกเป็นประเภทย่อยได้อีก 2 ประเภทย่อยคือ

- การแยกประเภท (Classification)

แบบจำลองการแยกประเภทเป็นแบบจำลองที่ใช้ในการแบ่งข้อมูลออกเป็นกลุ่ม โดยข้อมูลที่นำมาใช้ต้องมีกลุ่มเป้าหมายของข้อมูลกำกับอยู่ด้วย และคอมพิวเตอร์จะเรียนรู้จากลักษณะของข้อมูลที่ใส่เข้ามา เช่นการทำนายว่าลูกค้าอยู่ในกลุ่มลูกค้าชั้นดีหรือลูกค้าชั้นแย่ จากข้อมูลต่าง ๆ ของลูกค้าเช่น เพศ, อายุ, เงินเดือน เป็นต้น

- การทำนายผลข้อมูล (Regression)

แบบจำลองการทำนายผลข้อมูลจะทำการคำนวณผลลัพธ์ของข้อมูลด้วยปัจจัยต่าง ๆ ที่เกี่ยวข้องกับผลลัพธ์ เช่น การทำนายราคาที่ดินจากทำเลที่ตั้ง, ขนาดของที่ดิน, สาธารณูปโภค และหลักฐานการรับรองสิทธิ์ เป็นต้น

ความแตกต่างระหว่างการแยกประเภทกับการทำนายผลข้อมูลคือ การแยกประเภทจะให้ผลลัพธ์เป็นข้อมูลที่ไม่ต่อเนื่องมีลักษณะเป็นกลุ่ม เช่น หมวดหมู่ ใช่หรือไม่ใช่ และระดับความเสี่ยงเป็นต้น ในขณะที่การทำนายผลข้อมูลจะให้ผลลัพธ์เป็นข้อมูลแบบต่อเนื่อง เช่น ตัวเลขที่ได้จากการคำนวณของแบบจำลองที่มีค่าระหว่าง 0 ถึง 100 เป็นต้น

2.3.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning)

การเรียนรู้แบบไม่มีผู้สอนคือการที่คอมพิวเตอร์สามารถเรียนรู้ได้ด้วยตัวเองจากข้อมูลที่ใส่เข้าไป โดยข้อมูลนั้นไม่จำเป็นต้องมีผลลัพธ์กำกับไว้ ประเภทหลัก ๆ ของการเรียนรู้แบบไม่มีผู้สอนคือ การแบ่งกลุ่มข้อมูล (Clustering)

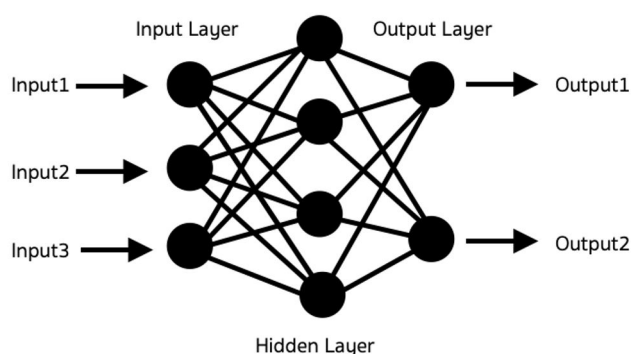
- การแบ่งกลุ่มข้อมูล (Clustering)

การแบ่งกลุ่มข้อมูลคือการที่คอมพิวเตอร์สามารถที่จะแบ่งข้อมูลออกเป็นกลุ่ม ๆ ได้ด้วยตัวเอง โดยที่ข้อมูลที่มีลักษณะใกล้เคียงกันจะอยู่ในกลุ่มเดียวกัน และข้อมูลที่แตกต่างกันจะอยู่คนละกลุ่มกัน เช่น การจัดกลุ่มของลูกค้าจากพฤติกรรมการซื้อสินค้าของลูกค้า โดยที่ถ้าลูกค้ามีลักษณะการซื้อสินค้าที่คล้ายกันจะถูกจัดอยู่ในกลุ่มเดียวกัน และลูกค้าที่มีลักษณะการซื้อสินค้าต่างกันจะอยู่คนละกลุ่มกัน

ปัจจุบันการเรียนรู้ของเครื่องได้ถูกนำไปประยุกต์ใช้อย่างแพร่หลายเช่น การคัดแยกจดหมายอิเล็กทรอนิกส์ การตรวจจับใบหน้าของมนุษย์ และการทำให้รถสามารถเคลื่อนที่ได้โดยไร้คนขับ

2.4 โครงข่ายประสาทเทียม (Artificial Neural Network)

โครงข่ายประสาทเทียมเป็นการจำลองระบบประสาทของสิ่งมีชีวิตขึ้นมา เพื่อให้คอมพิวเตอร์สามารถคำนวณผลลัพธ์ออกมาได้ จะประกอบไปด้วยชั้น (Layer) ต่าง ๆ 3 ชั้นที่สำคัญคือ ชั้นขาเข้า (Input Layer) ชั้นซ่อน (Hidden Layer) และสุดท้ายคือชั้นผลลัพธ์ขาออก (Output Layer) เมื่อมีชุดข้อมูลเข้ามาที่ชั้นขาเข้า ชุดข้อมูลนี้จะถูกประมวลผลในชั้นซ่อน และนำเสนอผลลัพธ์ที่ได้ผ่านชั้นผลลัพธ์ขาออก นอกจากนี้ยังสามารถอธิบายการทำงานของโครงข่ายประสาทเทียมผ่านฟังก์ชันทางคณิตศาสตร์ได้

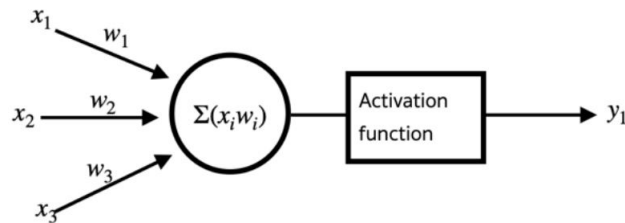


รูปที่ 2.5 แสดงโครงสร้างของโครงข่ายประสาทเทียม

การหาค่าผิดพลาด (Error) ของโครงข่ายประสาทเทียมจะใช้อัลกอริทึมที่ชื่อว่า feed-forward neural networks คือข้อมูลจะถูกส่งไปข้างหน้าเพียงหนึ่งทิศทางเท่านั้น กล่าวคือข้อมูลจะถูกส่ง

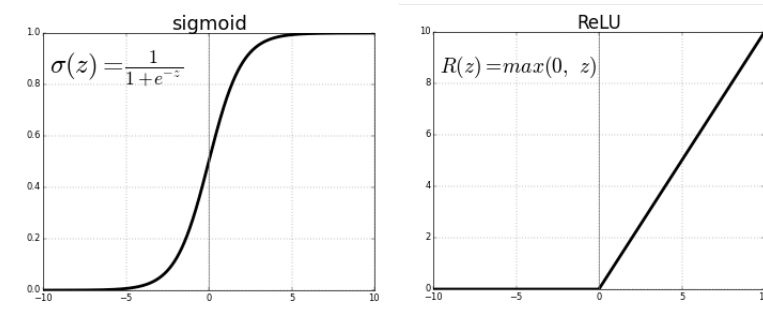
จากชั้นขาเข้าไปยังชั้นซ่อนและส่งต่อไปยังชั้นผลลัพธ์ขาออกเท่านั้นและเมื่อได้ผลลัพธ์ออกมาแล้วจะนำผลลัพธ์ที่ได้มาคำนวณกับค่าเป้าหมายเพื่อหาค่าผิดพลาด จากนั้นนำค่าผิดพลาดที่คำนวณได้ไปใช้ในการปรับค่าน้ำหนักของข้อมูล

ข้อมูลที่ได้รับเข้ามาจะถูกคูณด้วยค่าน้ำหนัก (Weight) ก่อนจะถูกส่งต่อไปยังเซลล์ประสาทในชั้นซ่อน โดยในหนึ่งเซลล์ประสาทในชั้นซ่อนนั้นจะประกอบไปด้วยฟังก์ชันทางคณิตศาสตร์หรือ Activation function ที่ใช้ในการพิจารณาผลลัพธ์ดังภาพ 2.6



รูปที่ 2.6 แสดงการทำงานของเซลล์ประสาทเทียบกับ Activation Function

ซึ่ง Activation function ที่เป็นที่นิยมได้แก่ Sigmoid, Softmax และ ReLU Activation Function โดยที่ Sigmoid Function จะเป็นการเปลี่ยนผลรวมของข้อมูลที่เข้ามาให้มีค่าอยู่ระหว่าง 0 ถึง 1 เท่านั้น จึงเหมาะที่จะถูกนำไปใช้ในงานที่ต้องการ ผลลัพธ์ หรือ output ที่มีความน่าจะเป็น (Probability) ส่วน Softmax Function จะให้ผลลัพธ์เป็น 0 ถึง 1 เหมือนกันแต่สามารถนำไปใช้งานได้หลากหลายกว่า เพราะเป็นการแสดงผลลัพธ์ของหลาย ๆ ข้อมูลที่เข้ามารวมกันเป็นหลาย ๆ ผลลัพธ์ซึ่งเหมาะมากกับการนำไปใช้ในการจำแนกประเภทหลายคลาส (Multi-Class Classification) ส่วน ReLU Function คือฟังก์ชันที่ได้รับความนิยมสูงสุดในขณะนี้ โดยที่ถ้าได้ผลลัพธ์ต่ำกว่า 0 จะได้ผลลัพธ์เป็น 0 โดยทันที แต่หากได้ค่ามากกว่าหรือเท่ากับ 0 ก็จะได้ผลลัพธ์เป็นค่านั้น ๆ ซึ่ง ReLU นั้นนิยมใช้มากในงานโครงข่ายประสาทแบบคอนโวลูชันและโครงข่ายประสาทเชิงลึก

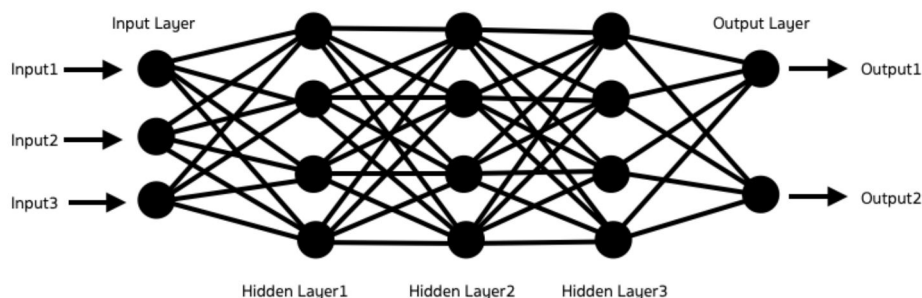


รูปที่ 2.7 รูปเปรียบเทียบระหว่าง Sigmoid และ ReLU Function

2.4.1 โครงข่ายประสาทเชิงลึก (Deep Neural Network)

โครงข่ายประสาทเชิงลึกเป็นการต่อยอดมาจากโครงข่ายประสาทเทียม (Artificial Neural Network) และเป็นส่วนหนึ่งที่อยู่ภายใต้ศาสตร์การเรียนรู้ของเครื่อง (Machine Learning) กล่าวคือในชั้นซ่อนของโครงข่ายประสาทเชิงลึกจะมีจำนวนมากกว่าจำนวนชั้นซ่อนของโครงข่ายประสาทเทียม และใช้อัลกอริทึมในการปรับค่าน้ำหนักที่ต่างกันด้วย

ในโครงข่ายประสาทเทียมเชิงลึกจะใช้อัลกอริทึม Backward Propagation ในการปรับค่าน้ำ

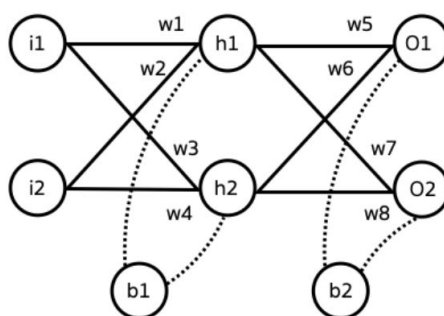


รูปที่ 2.8 แสดงโครงสร้างของโครงข่ายประสาทเชิงลึก

นัก ในชั้นแรกจะทำ feed-forward เมื่อเสร็จสิ้นแล้วจะทำการปรับค่าน้ำหนักของโครงข่ายประสาทเทียมเชิงลึกโดยใช้ใช้อัลกอริทึม Backward Propagation และใช้อัลกอริทึม Gradient Descent ในการหาค่าน้ำหนักที่เหมาะสมที่สุด

- การส่งค่าย้อนกลับ (Backward Propagation or Back Propagation)

ก่อนที่จะมีการส่งค่าย้อนกลับนั้น ต้องทำการคำนวณหาค่าผิดพลาด (Error) จากผลลัพธ์ที่ได้มาจากโครงข่ายประสาทเทียมก่อน โดยนำผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมมาเปรียบเทียบกับผลลัพธ์เป้าหมาย เมื่อได้ค่าผิดพลาดมาแล้วจะทำการส่งค่าผิดพลาดนี้กลับไปยังพารามิเตอร์ที่มีส่วนเกี่ยวข้องหรือก็คือค่าของน้ำหนักนั่นเอง เนื่องจากค่าผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมนั้น ถูกคำนวณมาจากค่าของน้ำหนัก สิ่งที่เราต้องทราบคือค่าน้ำหนักแต่ละตัวนั้นมี ผลกระทบต่อค่าผิดพลาดมากน้อยแค่ไหน โดยคำนวณจากการหาอนุพันธ์ของค่าผิดพลาดเทียบกับน้ำหนัก



รูปที่ 2.9 แสดงโครงข่ายประสาทเทียม

เพื่อให้ง่ายต่อการอธิบายจึงยกตัวอย่างของโครงข่ายประสาทเทียมดังภาพ 2.9 มาช่วยในการอธิบาย โดยเซลล์ i คือเซลล์ประสาทเทียมในชั้นขาเข้า เซลล์ h แทนเซลล์ประสาทเทียมในชั้นซ่อน และเซลล์ O แทนเซลล์ประสาทเทียมในชั้นผลลัพธ์ โดยผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียมนี้จะสามารถคำนวณได้ตามสมการตัวอย่าง

$$O1 = h1 \cdot w_5 + h2 \cdot w_6 + b_2 \quad (2.2)$$

จากสมการ 2.2 ทำให้ทราบว่าผลลัพธ์ที่ได้จากเซลล์ผลลัพธ์จะเท่ากับ ผลลัพธ์ของเซลล์ในชั้นซ่อนก่อนหน้าคูณด้วยน้ำหนักแล้วบวกด้วยค่าอคติหรือ bias ในทำนองเดียวกันการหา

ผลลัพธ์ของเซลล์ในชั้นซ่อนจะหาจากค่าของข้อมูลในชั้นขาเข้าคูณด้วยน้ำหนักและบวกด้วยค่าอคติดังสมการ 2.3

$$h1 = i1 \cdot w_1 + i2 \cdot w_2 + b_1 \quad (2.3)$$

หลังจากที่ทราบค่าในแต่ละเซลล์ผลลัพธ์ O1 และ O2 แล้วจะทำการหาผลรวมของค่าผิดพลาด (Error) จากสมการ 2.4

$$E_{total} = \sum_{i=1}^n \frac{1}{2} (y_i - f(x_i))^2 \quad (2.4)$$

จากสมการแสดงการคำนวณค่าผิดพลาด โดยที่ y_{target} คือผลลัพธ์เป้าหมายและ y_{out} คือผลลัพธ์ที่ได้จากโครงข่ายประสาทเทียม โดยที่เอาค่าที่ได้จากเซลล์ O1 ไปคำนวณด้วย Activation function แล้วนำมาเปรียบเทียบกับผลลัพธ์เป้าหมาย หลังจากที่ได้ค่าผลรวมความผิดพลาดมาแล้ว เราจะทำการหาอนุพันธ์ของค่าผิดพลาดเทียบกับน้ำหนักทีละตัว เพื่อให้ได้ค่าเกรเดียน (Gradient) ของน้ำหนักเพื่อนำไปใช้ในการปรับน้ำหนักดังสมการ 2.5

$$\frac{\partial Error(w)}{\partial W} \quad (2.5)$$

โดยที่ E_{total} คือผลรวมของค่าความผิดพลาด และ w_i คือน้ำหนักตัวที่ i และเนื่องจากสมการ 2.5 เป็นสมการเชิงอนุพันธ์ย่อย (Partial Differential) จึงต้องใช้กฎลูกโซ่ (Chain Rule) เข้ามาช่วยในการคำนวณ ตัวอย่างเช่นสมการต่อไปนี้จะแสดงการหาอนุพันธ์ระหว่างผลรวมของค่าผิดพลาดกับค่าน้ำหนักตัวที่ 5 (w_5)

$$\frac{\partial E_{total}}{\partial w_5} = \frac{\partial E_{total}}{\partial y_{out1}} x \frac{\partial y_{out1}}{\partial O_1} x \frac{\partial O_1}{\partial w_5} \quad (2.6)$$

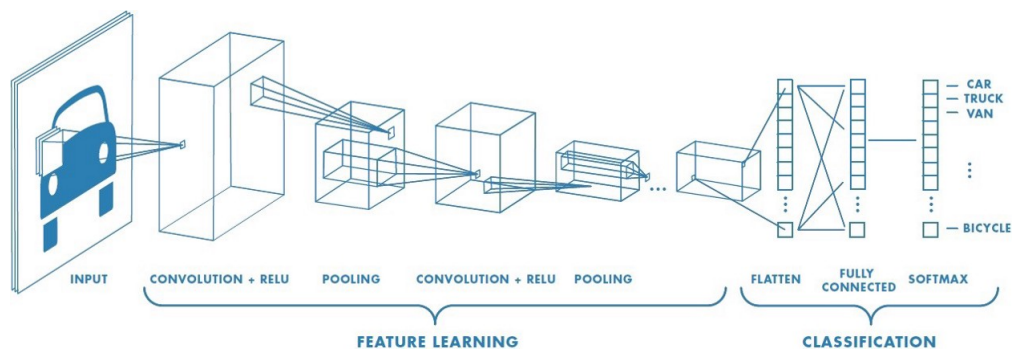
หลังจากที่คำนวณสมการข้างต้นแล้ว จะได้ค่าของเกรเดียนของน้ำหนักตัวที่ 5 หรืออัตราการเปลี่ยนแปลงของน้ำหนักตัวที่ 5 หลังจากนั้นเราจะนำค่าเกรเดียนที่ได้ไปใช้ในการปรับค่าน้ำหนักเพื่อที่จะลดค่าผิดพลาดให้มีย่าน้อยที่สุด โดยใช้สมการดังนี้

$$W_{new} = W_{old} - \alpha * \frac{\partial Error(w)}{\partial W} \quad (2.7)$$

โดยที่ α คือค่าอัตราการเรียนรู้ (Learning Rate) เป็นค่าไฮเปอร์พารามิเตอร์ (Hyperparameter) ที่ควบคุมการปรับค่าน้ำหนักของ โครงข่ายประสาทเทียมว่ามากหรือน้อยแค่ไหนในหนึ่งขั้นตอน (Step) ของการฝึกอบรมแบบจำลอง ซึ่งเป็นค่าคงที่ และหลังจากทำการปรับน้ำหนักแล้ว จะเริ่มทำการคำนวณแบบเดียวกันกับค่าน้ำหนักตัวอื่น ๆ ในชั้นเดียวกันจนครบทุกตัว ขั้นตอนต่อไปคือการคำนวณในทำนองเดียวกันกับชั้นก่อนหน้า โดยที่จะคำนวณอย่างนี้ซ้ำ ไปจนกระทั่งน้ำหนักทุกตัวได้รับการปรับปรุงจนหมด

2.4.2 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network, CNN)

โครงข่ายประสาทแบบคอนโวลูชันเป็นโครงข่ายประสาทเทียมหนึ่งในกลุ่มของ bio-inspired โดยจะจำลองการมองเห็นของมนุษย์ที่จะมองพื้นที่เป็นย่อย ๆ และนำพื้นที่ย่อย ๆ มารวมกัน เพื่อค้นหาสิ่งที่มนุษย์มองอยู่นั้นคืออะไร ซึ่งในการมองพื้นที่ย่อย ๆ ของมนุษย์นั้น จะมีการคัดแยกคุณลักษณะ (Feature Extraction) ของพื้นที่นั้น ๆ เช่นลายเส้น สี การรวมกันของลายเส้นและสี การมองภาพผ่านดวงตาของมนุษย์นั้นจะแบ่งการมองออกเป็นสองจุดใหญ่ ๆ คือจุดที่โฟกัสหรือสนใจและบริเวณต่าง ๆ โดยรอบของจุดนั้น ๆ เมื่อนำมาประกอบกันจะทำให้สมองประมวลผลสิ่งที่เห็นและจะทำให้ทราบว่าสิ่ง ๆ นั้นคืออะไร



รูปที่ 2.10 แสดงสถาปัตยกรรมของโครงข่ายประสาทเทียมแบบคอนโวลูชัน

- การคัดแยกคุณลักษณะเด่น (Feature Extraction) ในโครงข่ายประสาทแบบคอนโวลูชัน
เพื่อให้โครงข่ายประสาทเทียมมีการคำนวณที่สอดคล้องกับแนวคิดของตัวมันเองนั้น จึงมีการนำคณิตศาสตร์เข้ามาช่วยในการคำนวณโดยใช้หลักการเดียวกับ คอนโวลูชันเชิงพื้นที่ (Spatial Convolution) ในงานด้านการประมวลผลภาพดิจิทัล (Digital Image Processing) การคำนวณนี้จะมีการกำหนดค่าของตัวกรอง (Filter) ที่จะช่วยในการดึงคุณลักษณะที่ใช้ในการเรียนรู้และจดจำวัตถุออกมา โดยปกติตัวกรองหนึ่งตัวจะดึงคุณลักษณะที่สนใจออกมาได้เพียงหนึ่งคุณลักษณะเท่านั้น ทำให้จำเป็นต้องมีตัวกรองหลายตัวเพื่อหาคุณลักษณะทางพื้นที่หลาย ๆ อย่างมาประกอบกัน

- ลักษณะของตัวกรอง

ตัวกรองสำหรับภาพดิจิทัลนั้นโดยทั่วไปจะเป็นตาราง 2 มิติที่มีขนาดภาพตามพื้นที่ย่อย ๆ ที่ต้องพิจารณา

1	0	0
0	1	0
0	0	1

รูปที่ 2.11 แสดงตัวกรองขนาด 3 X 3

โดยที่ช่องที่แรเงาตรงกลางจะถูกเรียกว่า Anchor มีหน้าที่เอาไว้ทาบกับพิกเซลของข้อมูลภาพขาเข้า เริ่มแรกตัวกรองจะถูกวางทาบลงไปบนพิกเซลแรกของข้อมูลภาพ และหลังจากนั้นจะถูกเลื่อนไปในตำแหน่งของพิกเซลถัดไปที่ละพิกเซลจนครบทุกพิกเซลในภาพ ในขั้น

ตอนนี้มีข้อจำกัดคือ ตัวกรองจะไม่สามารถถูกนำไปทาบทับพิกเซลที่อยู่ใกล้กับขอบภาพได้ เนื่องจากตัวกรองจะเกินออกไปนอกภาพ เมื่อตัวกรองถูกเลื่อนไปจนครบทุกตำแหน่งที่สามารถเลื่อนไปได้แล้ว สิ่งที่ได้จากขั้นตอนนี้ จะถูกเรียกว่า พังคุณลักษณะ (Feature Map)

- การทำคอนโวลูชัน

การทำคอนโวลูชันเป็นการคูณเมทริกซ์ระหว่างข้อมูลภาพขาเข้า กับตัวกรอง เมื่อนำตัวกรองไปทาบทับบนข้อมูลภาพขาเข้าแล้ว ค่าในทุก ๆ ตำแหน่งที่ตรงกันจะทำการคูณกัน หลังจากนั้นจะรวมผลลัพธ์ที่ได้ทั้งหมดเข้าด้วยกันในตำแหน่งที่เรียกว่า Anchor หลังจากนั้นจะเลื่อนตัวกรองไปยังตำแหน่งถัดไป เมื่อทำครบทุกพิกเซลแล้วจะได้ผลลัพธ์ที่เป็นพังคุณลักษณะออกมา

1	1	0	1	0
0	0	0	1	1
0	1	1	0	0
1	0	0	0	0
0	1	0	1	0

1	0	0
0	1	0
0	0	1

2	1	1
1	1	0
0	2	1

รูปที่ 2.12 แสดงตัวอย่างของข้อมูลภาพขาเข้า ตัวกรอง และพังคุณลักษณะ

1	0	0	1	0
0	1	0	1	1
0	0	1	0	0
1	0	0	0	0
0	1	0	1	0

1	1	0	0	0
0	0	1	0	1
0	0	0	1	0
1	0	0	0	0
0	1	0	1	0

1	1	1	0	0
0	0	0	1	0
0	1	0	0	1
1	0	0	0	0
0	1	0	1	0

รูปที่ 2.13 แสดงตัวอย่างลักษณะการเคลื่อนที่ของตัวกรองเพื่อให้ได้ผลลัพธ์ดังภาพ 2.12

- Stride และ Padding

Stride เปรียบเสมือนระยะห่างในการเลื่อนพิกเซลของตัวกรอง Stride สามารถกำหนดให้มากขึ้นได้ ถ้าต้องการคำนวณหาคุณลักษณะที่มีพื้นที่ทับซ้อนกันน้อยลง แต่ในขณะเดียวกันเมื่อค่า Stride มากขึ้นพังคุณลักษณะจะมีขนาดเล็กลง

1	0	0	1	0
0	1	0	1	1
0	0	1	0	0
1	0	0	0	0
0	1	0	1	0

1	1	0	0	0
0	0	1	0	1
0	0	0	1	0
1	0	0	0	0
0	1	0	1	0

รูปที่ 2.14 แสดงการเคลื่อนที่ของตัวกรองเมื่อ Stride มีค่าเท่ากับ 1 และแสดงผลลัพธ์ของพังคุณลักษณะที่ได้

1	0	0	1	0
0	1	0	1	1
0	0	1	0	0
1	0	0	0	0
0	1	0	1	0

1	1	1	0	0
0	0	0	1	0
0	1	0	0	1
1	0	0	0	0
0	1	0	1	0

รูปที่ 2.15 แสดงการเคลื่อนที่ของตัวกรองเมื่อ Stride มีค่าเท่ากับ 2 และแสดงผลลัพธ์ของฟังก์ชันลักษณะที่ได้

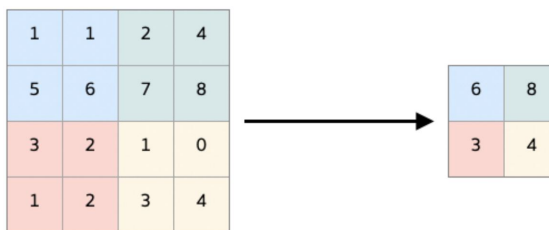
Padding จากขั้นตอนการทำคอนโวลูชันที่ผ่านมาจะเห็นได้ว่าฟังก์ชันลักษณะจะมีขนาดเล็กกว่าข้อมูลภาพขาเข้า หากต้องการให้ฟังก์ชันลักษณะมีขนาดเท่ากับข้อมูลภาพขาเข้าก็จะต้องมีการเติม 0 หรือค่าต่าง ๆ เข้าไปรอบ ๆ ข้อมูลภาพขาเข้า วิธีนี้จะช่วยแก้ปัญหาในกรณีที่ขอบของภาพมีความสำคัญที่ส่งผลต่อการตัดสินใจดังรูป 2.16

	1	1	0	1	0
	0	0	0	1	1
	0	1	1	0	0
	1	0	0	0	0
	0	1	0	1	0

รูปที่ 2.16 แสดงข้อมูลรูปภาพที่มีการทำ Padding และแสดงของฟังก์ชันลักษณะที่ได้หลังจากการทำ Padding

• Max Pooling

มนุษย์เราจำแนกวัตถุโดยอาศัยการดูรายละเอียดเล็ก ๆ และมองแบบคร่าว ๆ บนพื้นที่ใหญ่ ๆ จึงทำให้เกิดปัญหาหากต้องใช้ข้อมูลที่ยากหรือละเอียดอย่างใดอย่างหนึ่งในการจำแนกวัตถุ ดังนั้นในการฝึกอบรมแบบจำลองควรมีข้อมูลทั้งหายาและละเอียดควบคู่กันไป เพื่อให้สามารถจัดการกับปัญหานี้ได้ การย่อขนาดภาพจึงเป็นสิ่งสำคัญ เนื่องจากขนาดของตัวกรองมีความเท่าเดิมอยู่ตลอด ถ้าข้อมูลภาพที่เข้ามามีขนาดใหญ่เราจะได้ข้อมูลที่มีความละเอียดมาก ในขณะเดียวกันถ้าข้อมูลภาพมีขนาดเล็กลง ด้วยตัวกรองขนาดเท่าเดิม ตัวกรองจะสามารถครอบคลุมพื้นที่ของวัตถุเดิมได้มากขึ้น Pooling คือความสามารถในการย่อรูปแบบหนึ่งมี 2 ประเภทหลัก ๆ ที่เป็นที่นิยมคือ Max Pooling และ Mean Pooling โดย Max Pooling จะนำค่าที่มากที่สุดที่ตัวกรองทาบทับ อยู่มาเป็นผลลัพธ์ โดยตัวกรองของ Max Pooling จะทำงานในลักษณะเดียวกันกับ ตัวกรองในการทำคอนโวลูชัน โดยทั่วไปในการทำ Max Pooling จะใช้ตัวกรองขนาด 2 x 2 และมี Stride เท่ากับ 2



รูปที่ 2.17 แสดงการทำ Max Pooling ที่มีขนาดตัวกรองเท่ากับ 2×2 และ stride เท่ากับ 2

2.5 สถาปัตยกรรมต่าง ๆ ของโครงข่ายประสาทแบบคอนโวลูชันและการกำหนดองค์ประกอบ (CNN Architecture and Configuration)

2.6 เมตริกที่ใช้ในประเมินผลแบบจำลอง (Evaluation Metrics)

เมตริกที่ใช้ในการประเมินผลแบบจำลอง โดยปกติแล้วจะมีให้เลือกใช้หลายวิธีด้วยกัน แต่การจะหยิบเมตริกมาใช้ในการประเมินผลแบบจำลองนั้น จำเป็นจะต้องเลือกวิธีการหรือเมตริกที่เหมาะสมกับอัลกอริทึมของแบบจำลองนั้น ๆ เช่นการวัดความถูกต้องของการแยกประเภท (Classification Accuracy) [9] นั้น คืออัตราส่วนของจำนวนการคาดคะเนที่ถูกต้องต่อจำนวนตัวอย่างอินพุตทั้งหมด ซึ่งเรามักจะอ้างอิงถึงความหมายดังที่กล่าวมาข้างต้นตลอดเมื่อเราใช้คำว่า accuracy ในการประเมินผลแบบจำลองในการแยกประเภท และในบทนี้จะกล่าวถึงเมตริกต่าง ๆ ที่ใช้ในการประเมินผลแบบจำลองว่ามีความหมายอะไร ควรนำเมตริกต่าง ๆ ที่ว่า ไปใช้ตอนไหนและจะใช้งานมันอย่างไร ดังนี้

2.6.1 F1-Score

F1-Score คือค่าเฉลี่ยแบบ harmonic mean ระหว่าง precision และ recall โดยนักวิจัยสร้าง F1-Score ขึ้นมาเพื่อเป็น metric แบบตัวเดียวที่สามารถใช้วัดความสามารถของแบบจำลอง ได้โดยไม่ต้องเลือกตัวใดตัวหนึ่งระหว่าง precision recall และในทางปฏิบัติหากเราจะดูค่า F1-Score, precision หรือ recall เพื่อวัดประสิทธิภาพของแบบจำลอง เราควรจะดูค่าเหล่านี้ร่วมกับ ค่าความแม่นยำ (accuracy) เสมอโดยเฉพาะหากเราเจอปัญหากับข้อมูลที่มีสัดส่วนในการจัดกลุ่มไม่เท่ากัน (imbalanced classification) [10] โดย F1-Score สามารถคำนวณได้จากสมการที่ 2.8 และในส่วนของการคำนวณค่า precision และ recall สามารถคำนวณได้จากสมการ 2.9 และ 2.10 ตามลำดับ โดยที่

- **True Positive (TP)** คือ จำนวนของผลลัพธ์ที่แบบจำลองทำนายออกมาเป็นบวก (Positive) และผลลัพธ์ที่แท้จริงเป็นบวก (Positive)
- **True Negative (TN)** คือ จำนวนของผลลัพธ์ที่แบบจำลองทำนายออกมาเป็นลบ (Negative) แต่ผลลัพธ์ที่แท้จริงเป็นบวก (Positive)
- **False Positive (FP)** คือ จำนวนของผลลัพธ์ที่แบบจำลองทำนายออกมาเป็นบวก (Positive) แต่ผลลัพธ์ที่แท้จริงเป็นลบ (Negative)
- **False Negative (FN)** คือ จำนวนของผลลัพธ์ที่แบบจำลองทำนายออกมาเป็นลบ (Negative) และผลลัพธ์จริงเป็นลบ (Negative)

$$F1 = 2 \times \left(\frac{precision \times recall}{precision + recall} \right) \quad (2.8)$$

$$precision = \frac{TP}{(TP + FP)} \quad (2.9)$$

$$recall = \frac{TP}{(TP + FN)} \quad (2.10)$$

2.6.2 Receiver Operating Characteristic Curve (ROC Curve)

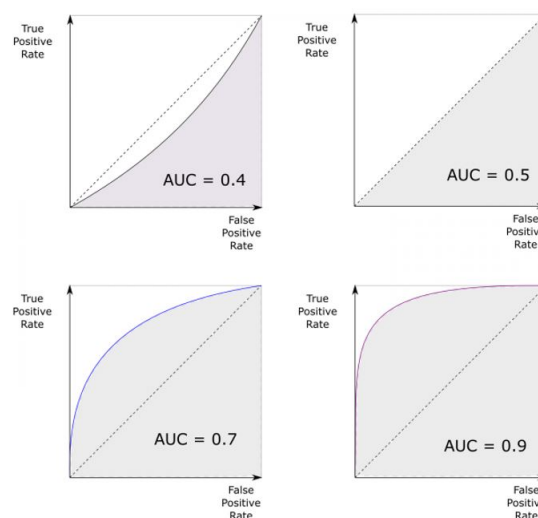
ROC Curve คือตัวกราฟแสดงความสัมพันธ์ระหว่างค่า True Positive Rate (TPR) หรือ recall และค่า False Positive Rate (FPR) เพื่อบ่งบอกประสิทธิภาพของการทดสอบแบบจำลองว่าสามารถแยกผลลัพธ์ที่เป็นบวก (Positive) และเป็นลบ (Negative) ออกจากกันได้ดีแค่ไหน โดย TPR หรือ recall คือค่าที่บอกว่าแบบจำลองของเราสามารถทำนายผลลัพธ์เป็น positive ได้เป็นอัตราส่วนเท่าไรของค่า positive ทั้งหมด และ FPR คือค่าที่บอกว่าแบบจำลองของเราสามารถทำนายผลลัพธ์เป็น positive ได้เป็นอัตราส่วนเท่าไรของค่า negative ทั้งหมดโดยที่ TPR และ FPR สามารถคำนวณได้จากสมการ 2.11 และ 2.12 ตามลำดับ

$$TPR = recall = \frac{TP}{TP + FN} \quad (2.11)$$

$$FPR = \frac{FP}{FP + TN} \quad (2.12)$$

2.6.3 Area under the ROC Curve (AUC)

AUC คือ metric ที่บ่งบอกว่าแบบทดสอบที่สร้างขึ้นมานั้น สามารถแบ่งแยกความแตกต่างระหว่างผลลัพธ์ที่เป็น positive และ negative ได้ดีแค่ไหน ยิ่ง AUC เข้าใกล้ 1 แสดงว่าแบบจำลองนั้นสามารถแยก positive และ negative ออกได้เป็นอย่างดี ในทางเทคนิค AUC คือพื้นที่ใต้กราฟของ ROC โดยมีตัวอย่างแสดงความสัมพันธ์ของ AUC และ ROC Curve ในรูปที่ 2.18



รูปที่ 2.18 แสดงความสัมพันธ์ของ AUC และ ROC Curve

2.6.4 Mean Average Precision (MAP)

ในการแข่งขัน BirdCLEF 2020 นี้ทางผู้จัดแข่งขันได้เลือกใช้ metric ตัวนี้เป็นตัววัดและประเมินผลโดย MAP คือค่าเฉลี่ยของค่าความแม่นยำเฉลี่ย (Average Precision (AveP)) สามารถคำนวณได้จากสมการ 2.13

$$MAP = \frac{\sum_{q=1}^Q AveP(q)}{Q} \quad (2.13)$$

โดย Q คือจำนวนของไฟล์ที่ใช้ในการทดสอบแบบจำลอง และ AveP(q) สำหรับไฟล์ q ที่เป็นไฟล์ที่ใช้ในการทดสอบแต่ละไฟล์ซึ่งสามารถคำนวณได้ตามสมการ 2.14

$$AveP = \frac{\sum_{k=1}^n (P(k) * rel(k))}{number\ of\ relevant\ document} \quad (2.14)$$

บทที่ 3

วิธีการดำเนินการวิจัย

ในส่วนของบทนี้จะพูดถึงลักษณะต่าง ๆ ของข้อมูลเสียงที่ได้ การดำเนินงานในการจัดการกับข้อมูลเสียง และสร้างแบบจำลองในแบบต่าง ๆ เพื่อคาดคะเนสายพันธุ์ของนกออกมาเป็นแบบหลาย ๆ ฉลาก (multi-label classification) ซึ่งลักษณะของข้อมูลเสียงที่ถูกนำมาใช้ในการทดลองและผลงานวิจัยที่เกี่ยวข้องจะถูกพูดถึงในหัวข้อที่ 3.1 ขั้นตอนการดำเนินงานทั้งหมดจะถูกกล่าวถึงในหัวข้อที่ 3.2 จนถึง หัวข้อที่ 3.4 และจะมีการสรุปผลการศึกษางานวิจัยที่เกี่ยวข้องและ ขั้นตอนการดำเนินงานทั้งหมดในหัวข้อที่ 3.5

3.1 บทนำและงานวิจัยที่เกี่ยวข้อง

3.1.1 การเก็บข้อมูลและลักษณะของข้อมูลเสียง ที่ถูกนำมาใช้ในการทดลอง

ข้อมูลในส่วนที่ถูกนำมาใช้ ให้เป็นข้อมูลสำหรับการฝึกในครั้งนี้ คือข้อมูลที่มาจากการแข่งขันในปี ค.ศ. 2019 [3] แต่ถูกปรับเปลี่ยน และพัฒนาคุณภาพขึ้นมาใหม่โดยผู้จัดทำและจัดทำข้อมูลรายใหม่อย่าง Xeno-canto [2]ที่เป็นเว็บไซต์ให้ผู้คนในสาขาต่าง ๆ ได้ร่วมกันแบ่งปันข้อมูลเสียงนกจากทุกที่ทั่วโลก และตัวเว็บไซต์เองไม่ได้เป็นแค่ชุดสะสมการบันทึกเสียงนกเพียงเท่านั้น แต่ยังเป็นโครงการความร่วมมือที่พร้อมจะให้ทุกคนที่ใช้งานเว็บไซต์มีส่วนร่วมในการช่วยกันแบ่งปันเสียงบันทึกของนก และระบุเสียงของนกชนิดต่างๆที่ปรากฏอยู่บนเว็บไซต์ด้วย

ข้อมูลจากการแข่งขันเมื่อปี ค.ศ. 2019 นั้น ประกอบไปด้วยทัศนียภาพของเสียงที่ถูกบันทึกด้วยมือกว่า 350 ชั่วโมง ซึ่งโดยส่วนใหญ่ถูกบันทึกไว้ด้วยเครื่องอัดเสียงภาคสนาม ช่วงเดือนมกราคม และเดือนมิถุนายน ในปี ค.ศ. 2017 ที่เมืองอิฐากา นครนิวยอร์ก ประเทศสหรัฐอเมริกา (Ithaca, NY, USA) และใช้เครื่องบันทึกเสียงรอบทิศทางที่ถูกจัดหาไว้โดยโครงการวิจัยด้านชีวเคมีประจำห้องปฏิบัติการด้านปักษีวิทยา แห่งมหาวิทยาลัยคอร์เนล (Bioacoustics Research Program of the Cornell Lab of Ornithology) เครื่องบันทึกเสียงเหล่านี้ สามารถบันทึกเสียงได้มากกว่า 30 หน่วยที่ขยายไปทั้งหมด 1 ตารางไมล์ผ่านแหล่งน้ำที่หลากหลาย และพืชพรรณหลายๆชนิด ซึ่งผู้ติดตั้งและผู้จัดการข้อมูลได้สุ่มเลือก 1 ไฟล์สำหรับแต่ละชั่วโมงในหนึ่งวัน ที่ถูกบันทึกจากเครื่องบันทึก 1 ใน 30 ตัวที่ถูกติดตั้งไว้ มารวบรวมและบันทึกไฟล์เข้าไว้ด้วยกันเป็นจำนวนทั้งหมด 15 วัน หลังจากนั้นก็นำไฟล์เสียงที่บันทึกไว้มาให้ผู้เชี่ยวชาญ อธิบายและทำเครื่องหมายไว้ว่าแต่ละช่วงเวลาในไฟล์ที่ถูกแบ่งไว้เป็นช่วงๆ ช่วงละ 5 วินาทีนั้น เป็นของนกสายพันธุ์ไหน

นอกจากนี้ในตัวชุดข้อมูลเมื่อปี ค.ศ. 2019 ได้มีการหยิบชุดข้อมูลจากการแข่งขัน BirdCLEF เมื่อปี ค.ศ. 2018 กลับมาใช้ซ้ำด้วย ซึ่งเป็นข้อมูลทัศนียภาพของเสียงที่มีความยาวทั้งหมดอยู่ประมาณ 4 ถึง 5 ชั่วโมง ซึ่งข้อมูลดังกล่าวถูกบันทึกที่ประเทศโคลอมเบีย โดยนักปักษีวิทยาจากมูลนิธิความหลากหลายทางชีวภาพแห่งโคลอมเบีย ที่มีชื่อว่า Paula Caycedo Rosales และสมาชิกของ Xeno-canto สามารถดูรายละเอียดเกี่ยวกับทัศนียภาพของเสียงโดยรวมได้ที่ หมายเหตุการทำงานของการแข่งขัน BirdCLEF เมื่อปี ค.ศ. 2018 [11]

ข้อมูลที่ใช้ในการฝึกแบบจำลอง สำหรับการแข่งขัน BirdCLEF ในปีนี้จะประกอบไปด้วยเสียงบันทึกของนกสายพันธุ์ต่างๆที่มาจากทั้งอเมริกาเหนือ อเมริกาใต้ และโซนยุโรป โดยที่ชุมชนของเว็บไซต์ Xeno-canto เป็นผู้จัดหาและสนับสนุนชุดข้อมูลการบันทึกเสียงคุณภาพสูงกว่า 70,000

การบันทึก ผ่านทางสายพันธุ์นกทั้งหมดกว่า 960 สายพันธุ์ การบันทึกในแต่ละครั้งจะมีเมตาดาต้า (metadata) ประกอบไปกับการบันทึกด้วย โดยในเมตาดาต้าจะประกอบไปด้วย สถานที่ในการเก็บบันทึก เวลา และคำอธิบายอื่นๆที่ถูกจัดไว้ให้โดยผู้บันทึกเสียงในการบันทึกครั้งนั้นๆ

ส่วนข้อมูลที่จะถูกนำไปใช้ในการทดสอบแบบจำลอง จะประกอบไปทัศนียภาพของเสียงโดยรอบกว่า 153 เสียงที่ถูกบันทึกไว้ในประเทศเปรู สหรัฐอเมริกา และประเทศเยอรมนี โดยในแต่ละเสียงที่ปรากฏ จะมีระยะเวลาทั้งหมด 10 นาทีและมีการเปล่งเสียงของนกที่ทับซ้อนกันภายในแต่ละช่วงเวลาในปริมาณที่ค่อนข้างสูง

3.1.2 พฤติกรรมการส่งเสียงร้องของนก

นกใช้เวลาส่วนใหญ่ไปกับการส่งเสียงร้อง แต่ในขณะเดียวกันเสียงร้องของนกในแต่ละฤดูกาลจะมีความแตกต่างกัน วัตถุประสงค์ในการส่งเสียงร้องของนกแบ่งออกเป็น 2 อย่างหลัก ๆ ประการแรกคือ นกเพศผู้จะส่งเสียงร้องเพื่อบ่งบอกถึงอาณาเขตของตนเอง และประกาศให้ศัตรูรับรู้ว่า จะปกป้องอาณาเขตของตนเองจากเผ่าพันธุ์อื่น ส่วนวัตถุประสงค์ที่สองคือ นกจะส่งเสียงร้องเพื่อดึงดูดคู่ครองและสร้างรัง นกเพศเมียมักจะเลือกคู่ครองโดยขึ้นอยู่กับการผสมผสานระหว่างการมองเห็นและเสียงร้อง แม้แต่นกเพศผู้ที่มีขนที่สวยงามในฤดูผสมพันธ์ ก็มีปัญหาในการหาคู่ครองเนื่องจากเสียงร้องที่ไม่สามารถเข้ากันได้ นกแต่ละสายพันธุ์จะมีเสียงร้องเป็นเอกลักษณ์ของตัวเอง ทำให้เมื่อนกได้ยินเสียงร้องจะสามารถแยกได้ว่าเสียงร้องนั้นมาจากสายพันธุ์ของตนเองหรือไม่ ถึงแม้ว่านกจะใช้เวลาและพลังงานส่วนใหญ่ไปกับการส่งเสียงร้อง แต่ก็ไม่สามารถทำได้ตลอดทุกฤดูกาล โดยส่วนใหญ่จะส่งเสียงร้องมากที่สุดในช่วงฤดูการทำรังและเมื่อฤดูทำรังสิ้นสุดลงนกจะส่งเสียงร้องน้อยลงและอาณาเขตของพวกมันจะพังทลายลงเนื่องจากมีนกหลายชนิดที่จะทำการอพยพตามฤดูกาล ในนกหลายสายพันธุ์จะมีเพียงเพศผู้ที่ส่งเสียงร้องเท่านั้น แต่นอกเหนือจากนั้น นกจะส่งเสียงร้องทั้งเพศผู้และเพศเมีย และในนกบางตัวจะไม่มีเสียงร้องเลย เช่นอีแร้งและนกกระสาที่แทบจะไม่สามารถสร้างเสียงใด ๆ ได้ [12] ในการระบุสายพันธุ์ของนกด้วยเสียงร้องสามารถสังเกตได้จากคุณลักษณะของเสียงร้องที่แตกต่างกัน 7 คุณลักษณะ [13] ดังต่อไปนี้

- ระดับเสียง (Pitch) เสียงร้องมีความสูงหรือต่ำแค่ไหน?
- คุณภาพของเสียง (Quality) มีเสียงที่แตกต่างกันในเสียงร้องหรือไม่?
- ความยาว (Length) เสียงร้องมีความยาวเท่าไร? ใช้เวลาเท่าไรในการร้องซ้ำ
- ความเร็วเพลง (Tempo) เสียงร้องมีกี่จังหวะ เร็วแค่ไหน? หรือช่วงไหนของเสียงร้องมีการหยุดชั่วคราว
- ความดังของเสียง (Volume) เสียงร้องมีการเปลี่ยนระดับความดังหรือไม่? ถ้ามีการเปลี่ยนระดับความดังของเสียง เปลี่ยนที่ตำแหน่งใดและเปลี่ยนอย่างไร? นกที่มีเสียงร้องคล้ายกันแต่ต่างชนิดกันจะมีความดังของเสียงที่ต่างกัน
- การทำซ้ำ (Repetition) มีพยางค์ของเสียงที่เหมือนกันซ้ำกันหลายครั้งหรือไม่ ซ้ำกันกี่ครั้ง และมีลำดับที่คล้ายกันเป็นส่วนหนึ่งของเสียงร้องเท่าไร?
- การจำลอง (Mimicry) มีเสียงที่ผิดปกติที่คล้ายสิ่งอื่นในเสียงร้องหรือไม่ เช่น เสียงสัญญาณกันขโมย เสียงเครื่องมือต่าง ๆ เนื่องจากอาจมีการเลียนแบบเสียงร้องของนก

3.1.3 กระบวนการขั้นพื้นฐาน (Baseline methods)

วิธีพื้นฐานที่ใช้กันทั่วไปในการสืบค้นเสียงมีพื้นฐานมาจากพลังงานอย่างใดอย่างหนึ่ง ไม่ว่าจะเป็นคลื่นสเปกตรัมข้ามสหสัมพันธ์ (spectrogram cross-correlation) แบบจำลองมาร์คอฟซ่อนเร้น (Hidden Markov model (HMM)) [14] ซึ่งกระบวนการพื้นฐานที่กล่าวมานี้มักเป็นที่รู้จักและใช้กันทั่วไปในการทำวิจัยเกี่ยวกับเสียงสะท้อนชีวภาพ (bioacoustics)

แต่ในบางครั้งวิธีที่ง่ายที่สุดมักเป็นการตั้งค่าจุดเปลี่ยนผ่านหาค่าผลลัพธ์ที่ได้ มีค่าสูงกว่าค่าที่ตั้งไว้ (thresholding) ซึ่งจะมีผลลัพธ์เป็นบวก (positive) หากค่าพลังงานในช่วงเวลาสั้นๆ ที่ตัดมานั้นมีค่าสูงกว่าค่าที่ตั้งไว้ และนอกจากนั้นจะเป็นค่าลบ (negative) และสำหรับการทำเสียงสะท้อนชีวภาพมักจะมีการจัดการกับเสียงรบกวน และการเพิ่มชุดข้อมูล (data augmentation) เพื่อประมาณค่าของเสียงรบกวนที่ได้มาเมื่อเวลาผ่านไป และเพื่อให้มั่นใจว่าสามารถรับมือกับเสียงรบกวนได้ตลอด

นอกเหนือจากนี้แล้วการแข่งขัน BirdClef ประจำปีค.ศ. 2018 ได้ให้ตัวอย่างลำดับการทำงานคร่าว ๆ ในการจัดการกับเสียงนกที่ใช้ในการแข่งขันมาเป็นระบบในการตรวจจับเสียงนกแบบพื้นฐาน (Baseline system) [?] เพื่อให้ผู้เข้าแข่งขันสามารถนำระบบพื้นฐานที่ได้ไปพัฒนาระบบขึ้นมาใหม่เพื่อให้ได้ผลลัพธ์ที่ดีขึ้น โดยระบบพื้นฐานในการตรวจจับเสียงนกที่ผู้จัดแข่งขันให้มานั้นได้ระบุถึงลำดับการทำงาน (Workflow) รายละเอียดชุดข้อมูลสถาปัตยกรรมโครงข่ายประสาทแบบคอนโวลูชัน (CNN Architecture) การประเมินผล และแนวทางในการพัฒนาระบบต่อไปด้วย

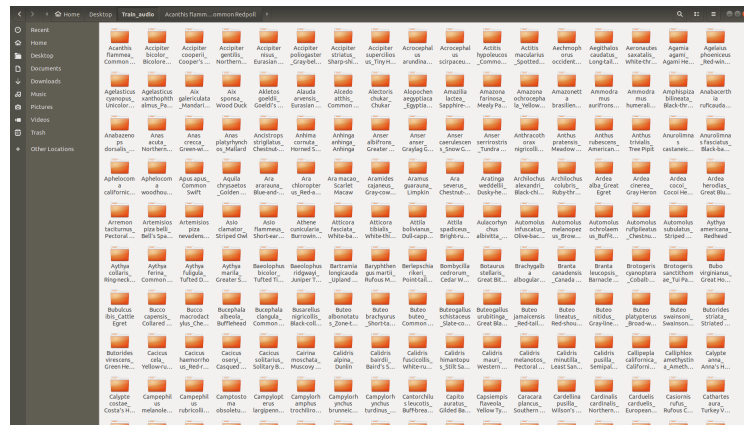
3.2 การจัดเตรียมการทดลอง

ข้อมูลที่ได้จากการแข่งขันในรอบปี ค.ศ. 2020 นี้ได้แบ่งไว้เป็นข้อมูล 3 ชุดหลักๆคือ ข้อมูลที่ใช้สำหรับการฝึกแบบจำลอง (training data) ข้อมูลที่ใช้สำหรับการทดสอบประสิทธิภาพของแบบจำลองเพื่อหา Hyperparameter ที่ดีที่สุด (validation data) และข้อมูลที่ไว้ใช้สำหรับการทดสอบแบบจำลอง (test data) ซึ่งก่อนที่จะสามารถสร้างแบบจำลอง หรือทำนายผลออกมาได้อย่างสมบูรณ์ก็ต้องมีการทำความเข้าใจตัวชุดข้อมูล และจัดเตรียมข้อมูลก่อนนำข้อมูลเข้าให้ดีกว่า

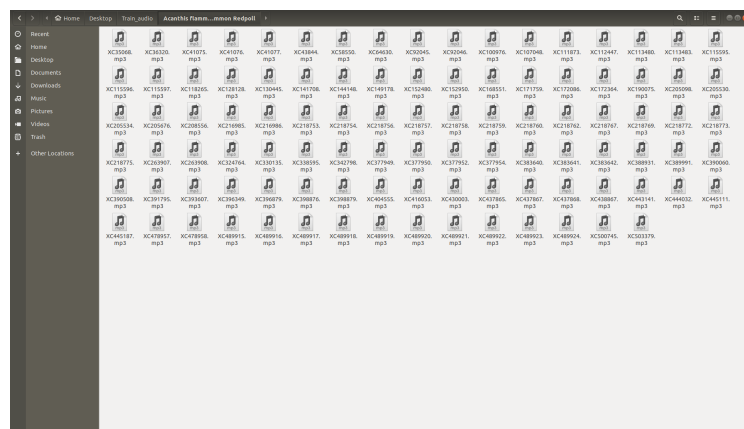
3.2.1 การสำรวจชุดข้อมูลเสียงนก (Data exploration)

จากการสำรวจข้อมูลในเบื้องต้นพบว่าข้อมูลถูกแบ่งออกเป็น 3 ส่วนใหญ่ ๆ โดยมีรายละเอียดดังต่อไปนี้

- **Audio Training Set** ข้อมูลชุดข้อมูลนี้เป็นชุดข้อมูลที่มีความไม่สมดุลกันของข้อมูลและข้อมูลชุดนี้ประกอบไปด้วยไฟล์เสียงของนกทั้งหมด 961 สายพันธุ์โดยมีไฟล์เสียงนกทั้งหมด 72324 ไฟล์โดยจัดเก็บนกแต่ละสายพันธุ์ไว้ด้วยโพลเดอร์ดังที่แสดงในรูป 3.1 และเก็บไฟล์เสียงของนกแต่ละสายพันธุ์จะถูกจัดเก็บไว้ในโพลเดอร์ที่เก็บไฟล์ .mp3 หลายๆไฟล์ดังที่แสดงในรูป 3.2



รูปที่ 3.1 โฟลเดอร์ที่ใช้ในการเก็บไฟล์เสียง



รูปที่ 3.2 ตัวอย่างไฟล์เสียงที่ถูกเก็บไว้ในโฟลเดอร์

- Audio Validation set** ข้อมูลชุดนี้คือข้อมูลที่ใช้ในการทดสอบแบบจำลองเพื่อหาค่า Hyperparameter ที่ดีที่สุด โดยข้อมูลในชุดนี้ประกอบไปด้วยข้อมูลเสียงที่เป็นไฟล์ .wav ทั้งหมด 12 ไฟล์ ซึ่งไฟล์เหล่านี้จะถูกนำไปประมวลผลในแบบจำลองและทำนายว่าเสียงของนก ณ ช่วงเวลานั้น เป็นเสียงของนกสายพันธุ์ใด และอีกส่วนคือไฟล์ข้อมูลที่เป็นไฟล์ .csv ซึ่งทั้งหมด 12 ไฟล์ซึ่งเป็นไฟล์เฉลยสำหรับการทำนายที่ได้จากแบบจำลอง ดังที่แสดงในรูป 3.3 โดยในช่องแรกจะเป็นช่องที่แสดงถึงช่วงเวลาของที่อยู่ภายในไฟล์เสียงนั้นๆ และช่องที่สองคือช่องที่บอกว่าในแต่ละช่วงเวลานั้นเสียงที่ได้ยินคือเสียงของนกสายพันธุ์อะไรแล้วจึงนำค่าที่ได้จากการทดสอบแบบจำลองไปปรับเปลี่ยนการตั้งค่า และหาค่า Hyperparameter ใหม่ให้กับแบบจำลองตามความเหมาะสม
- Audio Test set** คือชุดข้อมูลที่ไว้ใช้สำหรับทดสอบแบบจำลองที่ประกอบไปด้วยไฟล์เสียง .wav รวมอยู่ทั้งหมด 153 ไฟล์ที่เป็นไฟล์ทัศนียภาพของเสียง โดยแต่ละไฟล์จะมีความยาวอยู่ที่ไฟล์ละ 10 นาที ซึ่งเราต้องนำแบบจำลองของเราไปประมวลผลผ่านไฟล์ที่ได้เพื่อให้ได้คำตอบออกมาดังรูปที่ 3.3 และนำไฟล์คำตอบเหล่านั้นส่งไปยังเว็บไซต์ [21] ที่ผู้จัดแข่ง BirdCLEF 2020 เป็นคนกำหนด

00:00:00-00:00:05	cintin1
00:00:05-00:00:10	cintin1
00:00:05-00:00:10	citwoo1
00:00:10-00:00:15	butwoo1
00:00:10-00:00:15	cintin1

รูปที่ 3.3 ไฟล์แจกจ่ายการทำงานของแบบจำลอง

3.2.2 Metrics ที่ใช้สำหรับการประเมินประสิทธิภาพของแบบจำลอง

บทที่ 4

ผลการทดลองเบื้องต้นหรือระบบต้นแบบ

บทที่สี่

บทที่ 5

บทสรุป

บทที่ห้า

บรรณานุกรม

- [1] “BirdCLEF 2020 | ImageCLEF / LifeCLEF - Multimedia Retrieval in CLEF.” [Online]. Available: <https://www.imageclef.org/BirdCLEF2020>
- [2] “Xeno-canto Foundation and Naturalis Biodiversity Center (2020).” [Online]. Available: <https://www.xeno-canto.org/about/xeno-canto>
- [3] S. Kahl, F. R. Stöter, H. Goëau, H. Glotin, R. Planqué, W. P. Vellinga, and A. Joly, “Overview of BIRDCLEF 2019: Large-scale bird recognition in soundscapes,” 2019.
- [4] “BirdGenie Homepage (2020),” **BirdGenie**. [Online]. Available: <http://www.birdgenie.com/>
- [5] Jia Deng, Wei Dong, R. Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” 2009.
- [6] F. W. Pfeiffer, “Automatic differentiation in PyTorch Adam,” **ACM SIGNUM Newsletter**, 1987.
- [7] B. McFee, V. Lostanlen, M. McVicar, A. Metsai, S. Balke, C. Thomé, C. Raffel, A. Malek, D. Lee, F. Zalkow, Kyungyun Lee, O. Nieto, J. Mason, D. Ellis, R. Yamamoto, S. Seyfarth, E. Battenberg, [REDACTED], R. Bittner, Keunwoo Choi, J. Moore, Ziyao Wei, S. Hidaka, Nullmightybofo, P. Friesch, Fabian-Robert Stöter, D. Hereñú, Taewoon Kim, M. Vollrath, and A. Weiss, “librosa/librosa: 0.7.2,” Jan. 2020. [Online]. Available: <https://zenodo.org/record/3606573>
- [8] “มาทำความรู้จัก รูปแบบสัญญาณเสียง digital แบบ pcm กันเถอะ (ฉบับปรับปรุง) | re.v ->.” [Online]. Available: <https://rev.at1987.com/articles/pulse-code-modulation/>
- [9] H. M and S. M.N, “A Review on Evaluation Metrics for Data Classification Evaluations,” **International Journal of Data Mining & Knowledge Management Process**, 2015.
- [10] D. Soni, “Dealing with imbalanced classes in machine learning,” 2019. [Online]. Available: <https://towardsdatascience.com/dealing-with-imbalanced-classes-in-machine-learning-d43d6fa19d2>
- [11] H. Goëau, S. Kahl, H. Glotin, R. Planqué, W. P. Vellinga, and A. Joly, “Overview of BirdCLEF 2018: Monospecies vs. soundscape bird identification,” 2018.
- [12] “Why do birds sing? | Earth | EarthSky.” [Online]. Available: <https://earthsky.org/earth/why-do-birds-sing>

บรรณานุกรม (ต่อ)

- [13] “Identify Bird Calls - Easy Tips for Birding by Ear.” [Online]. Available: <https://www.thespruce.com/birding-by-ear-basics-387331>
- [14] B.-J. Yoon, “Hidden Markov Models and their Applications in Biological Sequence Analysis,” **Current Genomics**, 2009.

ภาคผนวก ก

เรื่องที่หนึ่ง

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi.

Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetur at, consectetur sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

ประวัติผู้เขียน

แก้ไขประวัติผู้เขียนในไฟล์ `author-bio.tex`