เว็บไซต์แปลภาษามือด้วย AI

AI-POWERED SIGN LANGUAGE TRANSLATION WEBSITE

โดย

นายอวิรุทธ์ จันทวรสุทธิ์

นายฉัตรชัย กาญจนวุฒิธรรม

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เว็บไซต์แปลภาษามือด้วย AI

AI-POWERED SIGN LANGUAGE TRANSLATION WEBSITE

โดย

นายอวิรุทธ์	**จันทวรสุทธิ์**	**62011042**
นายฉัตรชัย	**กาญจนวุฒิธรรม**	**62011311**

อาจารย์ที่ปรึกษา

ผศ.มนต์ชัย แช่มช้อย

ผศ.ดร.พิชญ สุพรรณกูล

ปริญญานิพนธ์เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2566

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง เว็บไซต์แปลภาษามือด้วย AI

AI-POWERED SIGN LANGUAGE TRANSLATION WEBSITE

ผู้จัดทำ

1. นายอวิรุทธ์ จันทวรสุทธิ์ 62011042
2. นายฉัตรชัย กาญจนวุฒิธรรม 62011311

..... **อาจารย์ที่ปรึกษา**

(ผศ.มนต์ชัย แช่มช้อย)

..... **อาจารย์ที่ปรึกษาร่วม**

(ผศ.ดร.พิชญ สุพรรณกูล)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การดำเนินโครงงานเรื่อง “เว็บไซต์แปลภาษามือด้วย AI” จะไม่สามารถสำเร็จลุล่วงไปได้ด้วยดี หากไม่ได้รับความช่วยเหลือ และความอนุเคราะห์อย่างดียิ่งจาก ผศ.มนต์ชัย แช่มช้อย และ ผศ.ดร.พิชญ สุพรรณกูล ที่กรุณาให้คำแนะนำ คำปรึกษา และแนวทางการแก้ไขปัญหาที่เป็นประโยชน์ต่อการศึกษา ค้นคว้าวิจัยให้โครงงานนี้สำเร็จสมบูรณ์ยิ่งขึ้น รวมถึงสนับสนุนสถานที่ เครื่องมือ และอุปกรณ์ต่างๆ ที่ใช้ระหว่างการจัดทำโครงงาน

ขอขอบคุณคณาจารย์และเจ้าหน้าที่ประจำภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ที่ได้อบรมสั่งสอนประสิทธิ์ประสาทวิชา ความรู้ และประสบการณ์ให้แก่ผู้จัดทำ

ขอกราบขอบพระคุณ บิดา มารดา และครอบครัว ที่ให้ความรัก ความห่วงใย และเป็นกำลังใจที่สำคัญเสมอมาและที่สำคัญคือสนับสนุนให้โอกาสทางด้านการศึกษาอันมีค่ายิ่งแก่ผู้จัดทำ

นายอวิรุทธ์ จันทวรสุทธิ์
นายฉัตรชัย กาญจนวุฒิธรรม
ผู้จัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เว็บไซต์แปลภาษามือด้วย AI

AI-POWERED SIGN LANGUAGE TRANSLATION WEBSITE

โดย นายอวิรุทธ์ จันทวรสุทธิ์ 62011042
นายฉัตรชัย กาญจนวุฒิธรรม 62011311

อาจารย์ที่ปรึกษา ผศ.มนต์ชัย แช่มช้อย
อาจารย์ที่ปรึกษาร่วม ผศ.ดร.พิชญ สุพรรณกูล

บทคัดย่อ

กลุ่มคนที่ใช้ภาษามือในการสื่อสารในพื้นที่สาธารณะมีอยู่ทั่วไป และหากคนเหล่านั้นต้องการความช่วยเหลือจากกลุ่มคนที่ไม่เข้าใจในภาษามือ อาจทำให้การสื่อสารให้เข้าใจล่าช้าหรืออาจเข้าใจความหมายไม่ตรงกับที่ต้องการสื่อสารได้ ปริญญานิพนธ์นี้จึงมีจุดประสงค์ที่ต้องการเพิ่มความรวดเร็วในการตรวจสอบภาษามือ และแปลภาษามือ โดยโครงงานนี้ได้นำเสนอเว็บไซต์แปลภาษามือแบบเรียลไทม์โดยอาศัยเทคโนโลยี AI สำหรับกลุ่มคนที่ไม่มีความเข้าใจในภาษามือและต้องการสื่อสารกับกลุ่มคนที่ใช้ภาษามือ ซึ่งใช้หลักการ COMPUTER VISION ในการเขียนโปรแกรมเพื่อจดจำภาษามือในท่าทางต่างๆ โดยระบบจะมีการทำงานคือเมื่อสแกนไปที่มือที่จะทำการแปลสำเร็จจะทำการแปลคำที่ได้จากภาษามือแสดงผ่านทางเว็บไซต์ และทำการจัดเรียงคำให้เป็นประโยคให้มีความหมายที่เข้าใจได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ABSTRACT

The group of people who use sign language to communicate in public areas faces challenges when they need assistance from individuals who do not understand sign language. This can lead to slower or misinterpreted communication. Recognizing the importance of this issue, the project aims to enhance the speed and accuracy of sign language interpretation. The project introduces a real-time sign language translation website that relies on AI technology, specifically Computer Vision. This system works by scanning the hand gestures used in sign language and then translating the signs into text on the website. It arranges the translated text into coherent sentences, making it easier for those who do not understand sign language to comprehend the intended message.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

		หน้า
กิตติกรรมประกาศ		I
บทคัดย่อ		II
สารบัญ		IV
สารบัญรูป		IX
บทที่ 1	**บทนำ**	
	1.1 ความเป็นมาและความสำคัญของปัญหา	1
	1.2 วัตถุประสงค์	1
	1.3 ขอบเขตของโครงงาน	1
บทที่ 2	**ทฤษฎีและหลักการที่เกี่ยวข้อง**	
	2.1 การเรียนรู้เชิงลึก (DEEP LEARNING)	2
	2.1.1 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (CONVOLUTION NEURAL NETWORKS: CNN)	2
	2.1.1.1 การดึงคุณลักษณะ (FEATURE EXTRACTION)	3
	2.1.1.2 การจำแนกประเภท (CLASSIFICATION)	6
	2.2 ภาษามือ (SIGN LANGUAGE)	8
	2.3 ชุดข้อมูลของรูปภาพภาษามือ	8
	2.4 MEDIAPIPE HOLISTIC	8
	2.4.1 การทำงานของ MEDIAPIPE HOLISTIC	9
	2.4.2 ประสิทธิภาพการทำงาน	10
	2.5 ภาษา PYTHON	11
	2.5.1 OPENCV-PYTHON	11

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

		หน้า
บทที่ 2	**ทฤษฎีและหลักการที่เกี่ยวข้อง**	
	2.5.2 KERAS	11
	2.6 ภาษา CSS	12
	2.7 ภาษา JAVASCRIPT	12
	2.7.1 การทำงาน	12
	2.8 AJAX	13
	2.9 DJANGO FRAMEWORK	14
	2.9.1 โครงสร้างโปรเจค DJANGO FRAMEWORK	15
	2.9.2 การทำงานของ DJANGO FRAMEWORK	16
	2.10 WEBSOCKET	18
	2.11 BOOTSTRAP	20
	2.11.1 การทำงานของ BOOTSTRAP	20
บทที่ 3	**การออกแบบและการจัดทำปริญญานิพนธ์**	
	3.1 การออกแบบ	22
	3.1.1 การออกแบบการเก็บ DATASET	22
	3.1.2 การออกแบบการฝึกสอนและทดสอบโมเดล	28
	3.1.3 การออกแบบการทำนายภาษามือจากรูปภาพ	30
	3.1.4 การออกแบบการแสดงผลภาษามือที่แปลได้ผ่านเว็บไซต์แบบเรียลไทม์	31
	3.1.5 การออกแบบเว็บแอปพลิเคชันที่ใช้แสดงผล	32
	3.2 เครื่องมือที่ใช้ในการทดลอง	33

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

บทที่ 3 การออกแบบและการจัดทำปริญญานิพนธ์

3.2.1 VISUAL STUDIO CODE 33

3.2.2 WEBCAM 34

3.3 การจัดเก็บผลการทดลอง 34

3.3.1 การทดสอบความแม่นยำในการฝึกฝนโมเดล 34

3.3.2 การทดสอบความแม่นยำในการตรวจจับมือ 34

3.3.3 การทดสอบความแม่นยำในการทำนายภาษามือบนเว็บไซต์ 34

บทที่ 4 ผลการทดลอง

4.1 การทดสอบความแม่นยำการฝึกฝนโมเดล 35

4.2 การทดสอบความแม่นยำของการตรวจจับมือ 36

4.3 การทดสอบความแม่นยำในการทำนายภาษามือบนเว็บไซต์ 37

4.3.1 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า GOOD LUCK 37

4.3.2 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า THANK YOU 38

4.3.3 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า UNDERSTAND 38

4.3.4 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า LIKE 39

4.3.5 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า SLOW 39

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

		หน้า
บทที่ 4	**ผลการทดลอง**	
	4.3.6 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า LINE	40
	4.3.7 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า GLOVES	40
	4.3.8 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า WALK	41
	4.3.9 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า COMPUTER	41
	4.3.10 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า PLAY	42
บทที่ 5	**สรุปผลและข้อเสนอแนะ**	
	5.1 สรุปผล	43
	5.2 ข้อเสนอแนะ	43
บรรณานุกรม		44
ภาคผนวก ก	โปรแกรมตรวจจับมือและบันทึกรูปภาพเก็บ DATASET	46
ภาคผนวก ข	โปรแกรมแปลภาษามือ	51
ภาคผนวก ค	โปรแกรมสร้าง ฝึกสอนและทดสอบโมเดล	57
ภาคผนวก ง	โปรแกรมสร้างเว็บแปลภาษามือ ส่วน SERVER (CONSUMERS)	63
ภาคผนวก จ	โปรแกรมสร้างเว็บแปลภาษามือ ส่วน SERVER (VIEW)	69
ภาคผนวก ฉ	โปรแกรมสร้างเว็บแปลภาษามือ ส่วน SERVER (URL)	71
ภาคผนวก ช	โปรแกรมสร้างเว็บแปลภาษามือ ส่วน CLIENT (ROUTER)	73

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

หน้า

ภาคผนวก ซ โปรแกรมสร้างเว็บแปลภาษามือ ส่วน CLIENT (BASE) 75
ภาคผนวก ฌ โปรแกรมสร้างเว็บแปลภาษามือ ส่วน CLIENT (VIDEOLIVE) 78
ภาคผนวก ญ โปรแกรมสร้างเว็บแปลภาษามือ ส่วน CLIENT (ABOUT) 85

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

หน้า

รูปที่

2.1 ตัวอย่างโครงสร้างของ CNN 2
2.2 ตัวอย่างของการทำการคอนโวลูชัน 3
2.3 การทำ FEATURE MAP จากการคอนโวลูชัน 3
2.4 ตัวกรองขนาด 3X3 4
2.5 ตัวอย่างในการเลื่อนตัวกรอง เมื่อกำหนดให้ STRIDE เท่ากับ 1 และ 2 4
2.6 ตัวอย่างของการทำ PADDING โดยให้ PADDING = 1 หรือ SAME MODE 4
2.7 ตัวอย่างการทำ MAX POOLING และ AVERAGE POOLING 5
2.8 ตัวอย่างการใช้งาน FLATTEN LAYER 6
2.9 การเชื่อมกันของแต่ละชั้นอย่างสมบูรณ์ใน FULLY CONNECTED LAYER 6
2.10 ค่าความน่าจะเป็นของ OUTPUT ที่ได้ โดยใช้สมการ SOFTMAX 7
2.11 รูปภาพภาษามือ 8
2.12 ขั้นตอนการทำงานของ MEDIAPIPE HOLISTIC 10
2.13 โครงสร้างโฟลเดอร์ใน DJANGO 15
2.14 ขั้นตอนการทำงานแบบ MTV ของ DJANGO 16
2.15 การทำงานเบื้องต้นของ DJANGO 16
2.16 CODE ของ PATH ต่าง ๆ ของ URL 17
2.17 CODE ในส่วนของการ REGISTER ของ VIEW 17
2.18 CODE ในส่วนของ MODEL 18
2.19 CODE การใช้ DJANGO TEMPLATE ในการทำไฟล์ HTML 18
2.20 การรับส่งข้อมูล 19
2.21 การเชื่อมต่อ WEBSOCKET 20
2.22 GRID SYSTEM 21
3.1 ท่าทางภาษามือที่ตรวจจับและบันทึก 22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

รูปที่		หน้า
3.2	โฟลเดอร์ที่บันทึกภาษามือแต่ละคำ	22
3.3	รูปภาษามือที่บันทึกในโฟลเดอร์	23
3.4	ท่าทางภาษามือคำว่า WORRY	23
3.5	ท่าทางภาษามือคำว่า THANK YOU	24
3.6	ท่าทางภาษามือคำว่า NEVER MIND	24
3.7	ท่าทางภาษามือคำว่า GOOD LUCK	24
3.8	ท่าทางภาษามือคำว่า MEET	25
3.9	ท่าทางภาษามือคำว่า SORRY	25
3.10	ท่าทางภาษามือคำว่า WORKING	26
3.11	ท่าทางภาษามือคำว่า DANGEROUS	26
3.12	ท่าทางภาษามือคำว่า APPROPRIATE	27
3.13	ท่าทางภาษามือคำว่า COMPUTER	27
3.14	บล็อกไดอะแกรมการทำ SIGN LANGUAGE RECOGNITION	28
3.15	LAYER ของ CNN MODEL สำหรับแยกรูปภาพภาษามือแต่ละคำ	29
3.16	พารามิเตอร์ของแต่ละ LAYER ของโมเดลที่ใช้ในการแยกรูปภาพภาษามือ	30
3.17	บล็อกไดอะแกรมการทำนายภาษามือจากรูปภาพ	31
3.18	บล็อกไดอะแกรมการแสดงผลบนเว็บไซต์	31
3.19	การแสดงผลของเว็บไซต์หน้าแรก	32
3.20	การแสดงผลของเว็บไซต์ในส่วนการทำนายภาษามือ	32
3.21	การแสดงผลของเว็บไซต์ในการเรียงคำจากปุ่ม ARRANGE WORDS	33
3.22	โปรแกรม VISUAL STUDIO CODE	33
3.23	WEBCAM รุ่น FIFINE K420	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

4.1 CNN MODEL 35

4.2 กราฟแสดงค่า LOSS, ACCURACY กับจำนวนรอบที่ฝึกสอน 35

4.3 MEDIAPIPE ตรวจจับมือทั้งสองข้างครบ 36

4.4 MEDIAPIPE ตรวจจับมือทั้งสองข้างไม่ครบ 36

4.5 หน้าเว็บไซต์ระหว่างทำนายภาษามือ 37

4.6 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า GOOD LUCK 37

4.7 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า THANK YOU 38

4.8 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า UNDERSTAND 38

4.9 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า LIKE 39

4.10 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า SLOW 39

4.11 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า LINE 40

4.12 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า GLOVES 40

4.13 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า WALK 41

4.14 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า COMPUTER 41

4.15 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า PLAY 42

4.16 แผนภูมิแท่งแสดงความถูกต้องในการทำนายของแต่ละคำ 42

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เนื่องจากในปัจจุบันมีกลุ่มคนที่จำเป็นจะต้องใช้ภาษามือในการสื่อสารในพื้นที่สาธารณะอยู่ทั่วไป และหากคนเหล่านั้นต้องการความช่วยเหลือจากกลุ่มคนที่ไม่เข้าใจในภาษามือ ทำให้การสื่อสารให้เข้าใจล่าช้าหรือเข้าใจความหมายไม่ตรงกับที่ต้องการสื่อสารได้ โดยทางคณะผู้จัดทำได้เห็นถึงความสำคัญของปัญหาที่เกิดขึ้น ปริญญานิพนธ์นี้จึงมีจุดประสงค์ที่ต้องการเพิ่มความรวดเร็วในการตรวจสอบภาษามือ และแปลภาษามือ โดยใช้เทคโนโลยีที่กำลังเป็นที่นิยมใจในปัจจุบัน และถูกใช้กันอย่างแพร่หลาย อย่างการเรียนรู้เชิงลึก (Deep Learning) รวมไปถึงเทคนิคอย่าง Information Extraction และเทคโนโลยี computer vision และเพิ่มความสะดวกรวดเร็วด้วยการสร้างเว็บไซต์ในการแปลภาษามือที่ใช้งานได้ผ่านโทรศัพท์ พร้อมกับจัดเรียงข้อความที่แปลออกมาเป็นประโยคให้เข้าใจยิ่งขึ้น

1.2 วัตถุประสงค์

1) เพื่อศึกษาทฤษฎีและการเขียนโปรแกรมด้าน Artificial Intelligence สำหรับการตรวจสอบภาษามือ

2) เพื่ออำนวยความสะดวกในการสื่อสารกับผู้ใช้ภาษามือได้อย่างรวดเร็ว

3) เพื่อนำเทคโนโลยี AI มาใช้ในชีวิตประจำวัน

4) เพื่อศึกษาและสร้างเว็บไซต์แปลภาษามือ

5) เพื่อออกแบบและสร้างระบบแปลภาษามือด้วย AI

1.3 ขอบเขตของปริญญานิพนธ์

ปริญญานิพนธ์นี้ได้ทำการออกแบบและสร้างเว็บไซต์แปลภาษามือด้วย AI โดยระบบประกอบไปด้วย 3 ส่วน คือ 1) ส่วนรับข้อมูล: เว็บแคมรับภาพของผู้ใช้ 2) ส่วนประมวลผล: โมเดลแปลภาษาแปลท่าทางภาษามือเป็นข้อความ 3) ส่วนแสดงผล: แสดงผลลัพธ์การแปลผ่านหน้าจอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและหลักการที่เกี่ยวข้อง

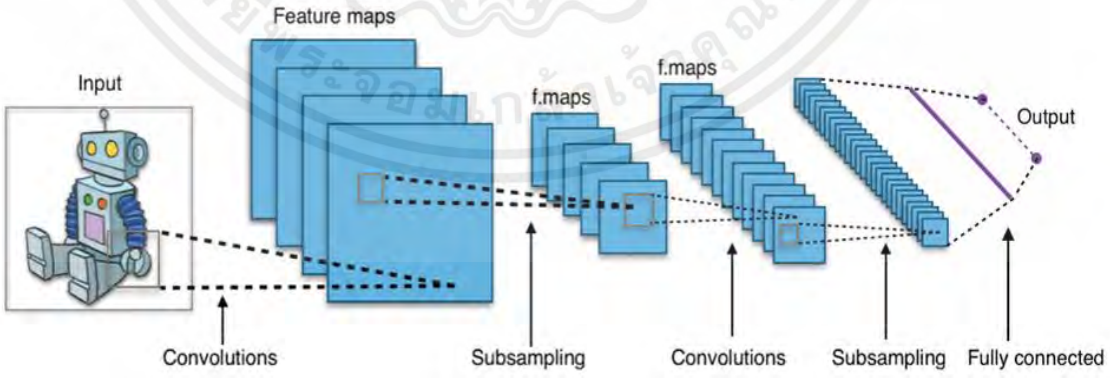
ปริญญานิพนธ์เรื่อง “เว็บไซต์แปลภาษามือด้วย AI” ได้ทำการออกแบบและพัฒนาโมเดลแปลภาษามือ ซึ่งอุปกรณ์นี้สามารถแปลภาษามือผ่านเว็บเบราเซอร์อัตโนมัติ ดังนั้นปริญญานิพนธ์ที่นำเสนอจึงมีหลักการที่เกี่ยวข้องดังต่อไปนี้

2.1 การเรียนรู้เชิงลึก (Deep Learning)

Deep Learning [1] คือวิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาท (Neural Network) มาซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูลดังกล่าวจะถูกนำไปใช้ในการตรวจจับรูปแบบ (Pattern) หรือจัดหมวดหมู่ข้อมูล (Classify the Data)

2.1.1 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Networks: CNN)

Convolutional Neural Network (CNN) [2] หรือ โครงข่ายประสาทเทียมแบบคอนโวลูชัน ดังรูปที่ 2.1 [2] เป็นโครงข่ายประสาทเทียมที่จำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นส่วนย่อย ๆ และนำกลุ่มของพื้นที่ย่อย ๆ เช่น ลายเส้น และการตัดกันของสี ซึ่งมนุษย์รู้ว่าพื้นที่ตรงนี้เป็นเส้นตรงหรือสีตัดกัน เพราะมนุษย์ดูทั้งจุดที่สนใจและบริเวณรอบ ๆ ประกอบกัน



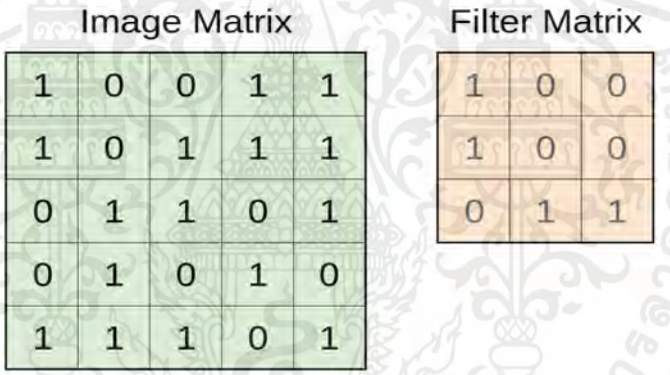
รูปที่ 2.1 ตัวอย่างโครงสร้างของ CNN [2]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

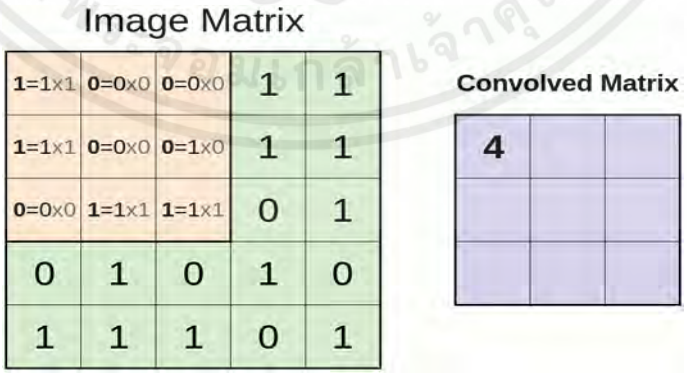
2.1.1.1 การดึงคุณลักษณะ (Feature extraction)

คือระบบการคำนวณที่สอดคล้องกับ Concept ของมันเองและต้องมีคณิตศาสตร์มารองรับ โดยการคำนวณตามแนวคิดนี้ใช้หลักการเดียวกันกับคอนโวลูชันเชิงพื้นที่ (Spatial Convolution) ในการทำงานด้าน Image Processing การคำนวณนี้จะเริ่มจากการกำหนดค่าในตัวกรอง (filter) หรือเคอร์เนล (kernel) ที่ช่วยดึงคุณลักษณะที่ใช้ในการรู้จำวัตถุออก โดยปกติตัวกรองเคอร์เนลอันหนึ่งจะดึงคุณลักษณะที่สนใจออกมาได้หนึ่งอย่าง จึงจำเป็นต้องใช้ตัวกรองหลายตัวกรองเพื่อหาคุณลักษณะทางพื้นที่หลายอย่างประกอบกัน โดยจะถูกแบ่งเป็นส่วนย่อยที่ทำงานร่วมกันอีก 3 ส่วน [3] คือ

1) Convolution layer เป็นกระบวนการที่ใช้สกัดเอาคุณลักษณะเด่นที่สำคัญของรูปภาพออกมา โดยยังคงความสัมพันธ์ของพิกเซลที่อยู่บริเวณพื้นที่ใกล้เคียงเอาไว้ [4] โดยการนำรูปภาพอินพุตมาคูณกับตัวกรองหรือเคอร์เนลที่ทำหน้าที่ดึงคุณลักษณะออกมา ดังรูปที่ 2.2 [5] โดยจะสกัดคุณลักษณะของอินพุตที่สนใจออกมาหรือที่เรียกว่า Feature map ดังตัวอย่างในรูปที่ 2.3 [5]



รูปที่ 2.2 ตัวอย่างของการทำการคอนโวลูชัน [5]



รูปที่ 2.3 การทำ Feature map จากการคอนโวลูชัน [5]

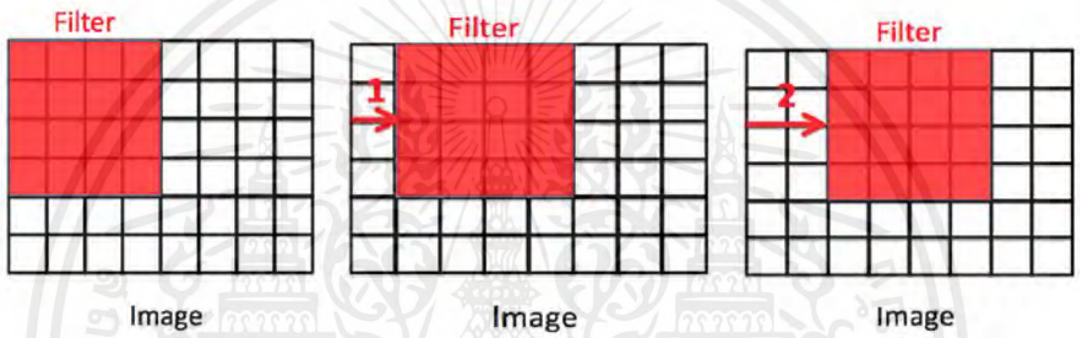
- ลักษณะของตัวกรอง (Filter) เป็นตัวกรองที่ใช้ในการดึงคุณลักษณะที่สนใจของข้อมูล ตัวอย่างของตัวกรองขนาด 3x3 แสดงดังรูปที่ 2.4 [5]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1	0	0
1	0	0
0	1	1

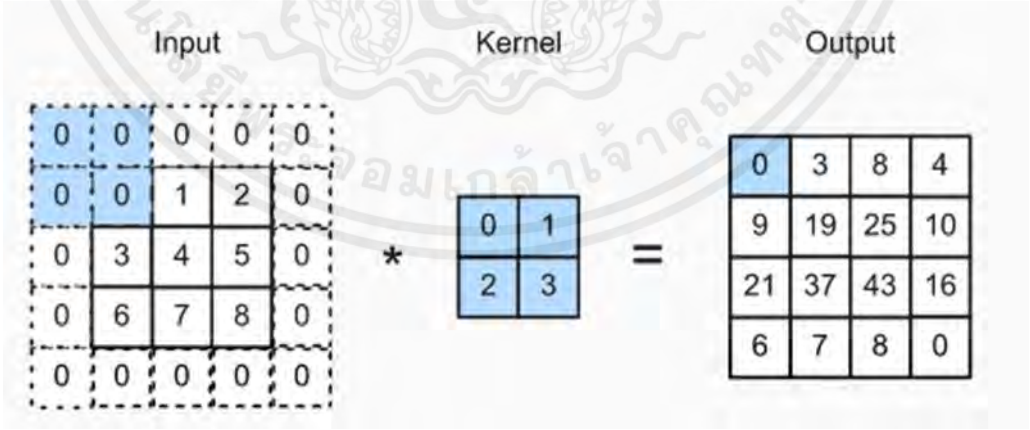
รูปที่ 2.4 ตัวกรองขนาด 3x3 [5]

- Stride เป็นค่าที่กำหนดในการเลื่อนตัวกรอง (Filter) ว่าต้องการให้ตัวกรองถูกเลื่อนไปด้วย Step เท่าไร แสดงดังตัวอย่างในรูปที่ 2.5 [6] เมื่อกำหนดให้ Stride เท่ากับ 1 และ 2



รูปที่ 2.5 ตัวอย่างในการเลื่อนตัวกรอง เมื่อกำหนดให้ Stride เท่ากับ 1 และ 2 [6]

- Padding เป็นส่วนที่ถูกเติมเข้าไป โดยทั่วไปจะเติมค่าที่เป็น 0 เพื่อให้การทำคอนโวลูชันแล้วมีผลลัพธ์ของ Feature map ยังคงมีขนาดเท่ากับอินพุต ดังรูปที่ 2.6 [6]



รูปที่ 2.6 ตัวอย่างของการทำ Padding โดยให้ Padding = 1 หรือ Same mode [6]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ฟังก์ชันเรคติไฟด์เชิงเส้น (Rectified Linear Unit: ReLU) เป็นฟังก์ชันที่เรียบง่ายและมีประสิทธิภาพที่ดี เนื่องจาก ถ้า Input เป็นบวก Slope จะเท่ากับ 1 ตลอดกาล ทำให้ Gradient ไม่หาย (ไม่เกิด Vanishing Gradient) ส่งผลให้เทรนโมเดลได้เร็วขึ้นมาก โดยสมการของ ReLU function [7] แสดงดังสมการที่ (2.1)

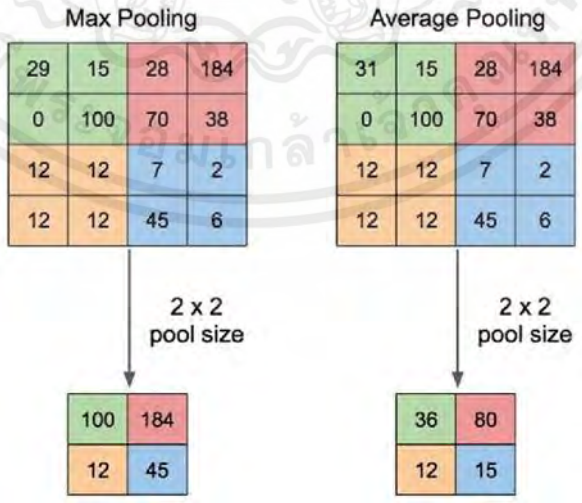
$$f(x) = \max(0,x) = \begin{cases} 0 \text{ for } x \leq 0 \\ x \text{ for } x > 0 \end{cases} \quad (2.1)$$

ถ้า x เป็นบวก Derivative ของ ReLU Function = 1 เป็นกุญแจสำคัญที่ช่วยให้การเทรน Deep Neural Network ทำได้รวดเร็วมากขึ้น แสดงดังสมการที่ (2.2)

$$f'(x) = \begin{cases} 0 \text{ for } x \leq 0 \\ 1 \text{ for } x > 0 \end{cases} \quad (2.2)$$

3) Pooling layer เป็นการช่วยลดมิติของ Feature map ให้มีขนาดเล็กลงแต่ยังคงรักษาข้อมูลที่สำคัญไว้ โดยทั่วไปการ Pooling มีสองแบบคือ Pooling ด้วยค่าสูงสุด (Max pooling) และ Pooling ด้วยค่าเฉลี่ย (Average pooling) นอกจากนี้การทำ Pooling ยังช่วยลดจำนวนของพารามิเตอร์และการคำนวณที่เกินจำเป็นในโครงข่าย [7]

ในกรณีที่ต้องการ Pooling ด้วยค่าสูงสุด (Max pooling) ดังรูปที่ 2.7 เป็นการกำหนดให้ตัวกรองมีขนาด 2x2 พิกเซล และทำการเลื่อนทีละ 2 พิกเซล โดยจะเลือกค่าที่สูงที่สุดในแต่ละ Pool ที่ถูกทาบด้วยตัวกรอง และการทำ Pooling ด้วยค่าเฉลี่ย (Average pooling) ดังรูปที่ 2.7 [7] โดยจะทำการหาค่าเฉลี่ยในแต่ละ Pool ที่ถูกทาบด้วยตัวกรอง

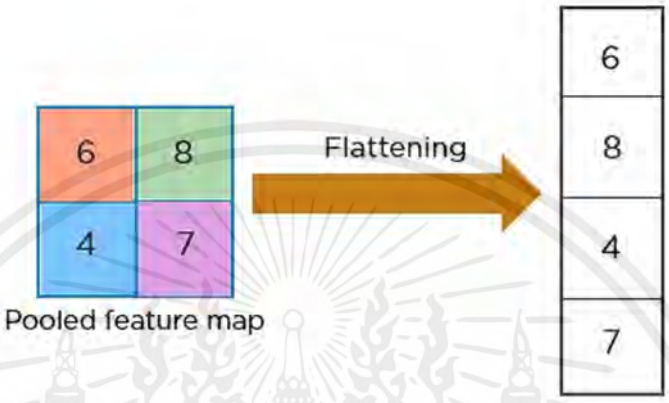


รูปที่ 2.7 ตัวอย่างการทำ Max pooling และ Average pooling [7]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

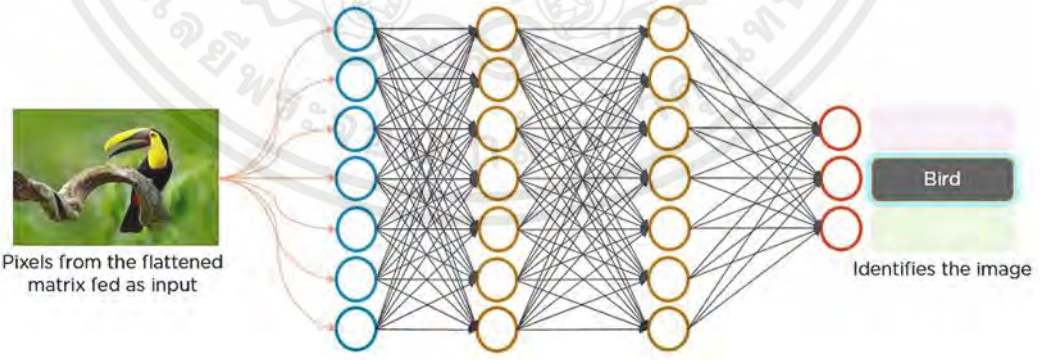
2.1.1.2 การจำแนกประเภท (Classification)

1) Flatten layer เป็นชั้นที่ทำการแปลงข้อมูล Output ที่มีหลายมิติ ให้เป็น 1 มิติ เพื่อเตรียมข้อมูลให้อยู่ใน format ที่พร้อมสำหรับเป็น Input ให้กับ Fully connected Layer ดังตัวอย่างในรูป 2.8 [8]



รูปที่ 2.8 ตัวอย่างการใช้งาน Flatten layer [8]

2) Fully connected layer เป็นชั้นที่เชื่อมโยงระหว่าง Feature Map และ Output แบบสมบูรณ์ นั่นหมายถึงทุก ๆ Neuron ที่อยู่ในชั้นสุดท้ายของ Feature Map จะถูกนำไปเปลี่ยนรูป (Reshape) หรือเรียกว่า Flatten เพื่อส่งไปคำนวณต่อในชั้นถัดไป ซึ่งก็คือชั้นของโครงข่ายประสาทเทียมแบบ MLP ดังนั้น Feature Map ในชั้นสุดท้ายจึงเปรียบเสมือนข้อมูล Input ไปยังโครงข่ายประสาทเทียม ดังรูปที่ 2.9 [8]



รูปที่ 2.9 การเชื่อมกันของแต่ละชั้นอย่างสมบูรณ์ใน Fully connected layer [8]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

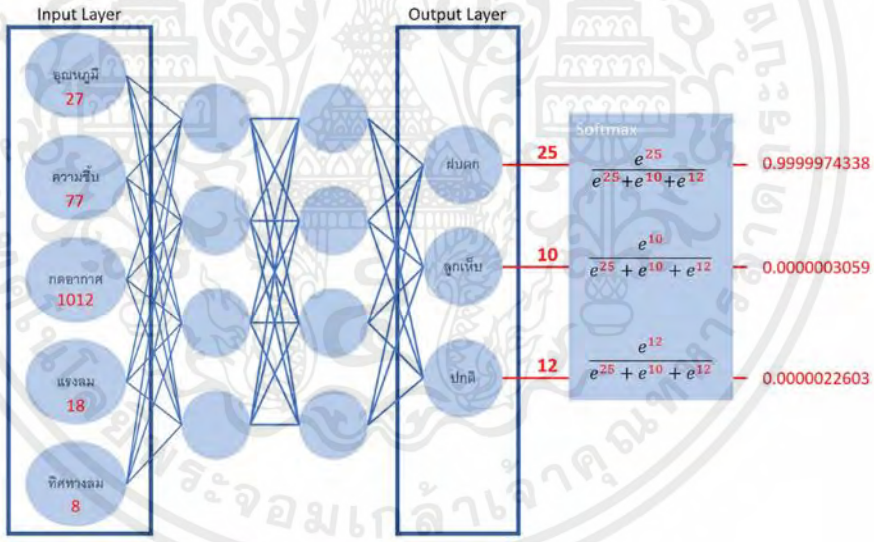
3) ฟังก์ชันซอฟต์แมกซ์ (Softmax) [9] เมื่อต้องการให้ Neural Network มี Output ออกเป็นค่า 0 หรือ 1 จึงต้องนำค่า Output จาก Neural Network เข้าสมการที่ปรับค่า Neural Network ให้อยู่ระหว่าง 0–1 ซึ่งสมการที่ใช้มีชื่อว่า Softmax หรือเรียกว่า Softmax function แสดงดังสมการ 2.3 โดยหลังจากนำ Output เข้า Softmax ผลที่ได้ออกมาจาก Softmax นั้นจะเป็นค่าที่บอกถึงความน่าจะเป็น

$$\sigma(z_j)=\frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \text{ for } j = 1,\ldots,K \qquad (2.3)$$

โดย z_j คือ ค่า Output ที่ออกจาก Neural Network

K คือ จำนวน Node ของ Output Layer

โดยสมมุติว่าค่าที่ออกมาจาก Neural Network คือค่า 25, 10 และ 12 ค่าที่ได้มาเข้าสมการ Softmax จะได้ผลลัพธ์ออกมาดังรูปที่ 2.10 [9]



รูปที่ 2.10 ค่าความน่าจะเป็นของ Output ที่ได้ โดยใช้สมการ Softmax [9]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2 ภาษามือ (Sign language)

ภาษามือ (Sign Language) เป็นระบบสื่อสารที่ใช้ส่งความหมายผ่านการใช้สัญลักษณ์ที่สร้างขึ้นด้วยมือและลำตัว ซึ่งใช้โดยคนที่มีการสูญเสียการได้ยินหรือผู้ที่มีความต้องการสื่อสารด้วยวิธีที่ต่างจากการใช้ภาษาพูดโดยปกติ ภาษามือมีความหลากหลายตามแต่ละประเทศและภูมิภาค และมีระบบสัญลักษณ์และกฎหมายที่แตกต่างกันไปในแต่ละภาษา

ในภาษามือ สัญลักษณ์ที่สร้างขึ้นด้วยมือและลำตัวถูกใช้เพื่อแสดงคำพูด, ความหมาย, และคำสั่งต่าง ๆ โดยมีการใช้เคลื่อนไหวของมือ, ลำตัว, และท่าทางต่าง ๆ เพื่อสร้างความหมาย นอกจากนี้ ภาษามือยังใช้มองเห็นส่วนของลำตัว, ใบหน้า, และท่าทางต่าง ๆ ในการสื่อสารเพิ่มเติม

การสร้างภาษามือและการใช้งานมันเป็นส่วนสำคัญของชุมชนคนหูหนวก และมีการพัฒนาและใช้งานอย่างกว้างขวางในหลายประเทศทั่วโลกเพื่อให้ผู้ที่มีการสูญเสียการได้ยินสามารถสื่อสารได้ได้อย่างมีประสิทธิภาพและเข้าใจกับผู้อื่นในชุมชนได้ดีขึ้น

2.3 ชุดข้อมูลของรูปภาพภาษามือ

โดยชุดข้อมูลรูปภาพภาษามือได้ทำการเก็บท่าภาษามือที่ใช้ในชีวิตประจำวัน จำนวน 100 ท่าทาง เพื่อสำหรับใช้ในการฝึกสอน และทดสอบโมเดล เพื่อให้โมเดลรู้และจดจำจนสามารถแยกรูปท่าทางภาษามือเป็นคำต่าง ๆ ได้ โดยตัวอย่างของรูปภาพภาษามือแสดงได้ดังรูปที่ 2.11

รูปที่ 2.11 รูปภาพภาษามือ

2.4 MediaPipe Holistic

MediaPipe Holistic คือโทโปโลยีล้ำสมัยที่สามารถตรวจจับท่าทาง มือ และใบหน้าของมนุษย์ในเวลาเดียวกัน และรองรับการใช้งานในแบบที่ไม่เคยมีแพลตฟอร์มไหนทำได้มาก่อน โซลูชันนี้จะใช้ Pipeline แบบใหม่ที่ประกอบด้วยการตรวจจับท่าทาง หน้า และมือที่ปรับแต่งให้ดีที่สุดเพื่อให้ทำงานได้เรียลไทม์ โดยใช้การโอนถ่ายหน่วยความจำระหว่าง Interference Backend ซึ่ง Pipeline จะรวมรูปแบบการปฏิบัติการและการประมวลผลที่แตกต่างกันตามการตรวจจับภาพแต่ละส่วนเข้าด้วยกัน และจะได้เป็นโซลูชันแบบครบวงจรที่ใช้งานได้แบบเรียลไทม์และสม่ำเสมอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

MediaPipe Holistic ใช้การทำงานแลกเปลี่ยนกันระหว่างการตรวจจับทั้งสามจุด โดยประสิทธิภาพของการทำงานจะขึ้นอยู่กับความรวดเร็วและคุณภาพของการแลกเปลี่ยนข้อมูล เมื่อรวมการตรวจจับทั้งสามเข้าด้วยกัน จะได้เป็นโทโปโลยีที่ทำงานร่วมกันเป็นหนึ่งเดียว โดยสามารถจับ Keypoints ของภาพเคลื่อนไหวได้ถึง 540+ จุด (ส่วนของท่าทาง 33 จุด มือข้างละ 21 จุด และส่วนใบหน้า 468 จุด) ซึ่งเป็นระดับที่ไม่เคยทำได้มาก่อน และสามารถประมวลผลได้เกือบจะเรียลไทม์ในการแสดงผลทางโทรศัพท์มือถือ

2.4.1 การทำงานของ MediaPipe Holistic

MediaPipe ใช้ข้อมูลจากกล้อง (camera) หรือวิดีโอ (video) เป็นแหล่งข้อมูลหลักในการทำงาน เพื่อประมวลผลภาพและวิดีโออย่างเชี่ยวชาญ เช่น การตรวจจับวัตถุ, การติดตามการเคลื่อนไหว, การแยกแยะมือและใบหน้า, การประมวลผลภาพเชิงลึก, การสร้างสัญญาณการทำงานเหนือภาพ (AR), การวิเคราะห์ท่าทาง (pose estimation)

นอกจากนี้ MediaPipe ยังสามารถใช้ข้อมูลจากแหล่งอื่นๆ ได้ เช่น ภาพจากไฟล์ภาพ, ข้อมูลเซ็นเซอร์, หรือข้อมูลที่รับมาจากเครือข่าย (network) และสามารถนำข้อมูลเหล่านี้มาประมวลผลและนำไปใช้งานได้ตามความต้องการของแอปพลิเคชันที่พัฒนาขึ้น

การทำงานของ MediaPipe จำเป็นต้องใช้ข้อมูลจากแหล่งต่างๆ ซึ่งสามารถประมวลผลเพื่อให้ได้ผลลัพธ์ที่ต้องการในแอปพลิเคชันได้อย่างถูกต้องและแม่นยำ

อันดับแรก MediaPipe Holistic จะระบุท่าทางของมนุษย์โดยใช้โมเดลตรวจจับท่าทางและโมเดลระบุ Keypoint หลังจากนั้นจึงนำ Keypoint ที่ระบุได้มาแบ่งออกเป็น 3 จุดสนใจ (Region of Interest: ROI) ครอปตัดส่วนที่เป็นแขน 2 ข้าง และส่วนหน้า แล้วจึงใช้ส่วนที่ครอปออกมาแทนเพื่อเพิ่มความละเอียดของจุดนั้น จากนั้น Pipeline จะทำการครอปเฟรมที่มีความละเอียดสูงสุดของจุด ROI ทั้งสองจุด แล้วจึงใช้กับโมเดลที่ตรวจจับส่วนของใบหน้าและมือเพื่อระบุตำแหน่ง Keypoints ตามส่วนต่าง ๆ และขั้นตอนสุดท้ายจะเป็นการนำ Keypoint ที่ได้มารวมกับ Keypoint ของโมเดลท่าทางในตอนต้น รวมกันเป็น 540 keypoint+ โดยรูปที่ 2.12 ขั้นตอนการทำงานของ MediaPipe Holistic

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.12 ขั้นตอนการทำงานของ MediaPipe Holistic [18]

MediaPipe Holistic จะใช้การคาดการณ์ท่าทางในทุก ๆ เฟรมล่วงหน้าไว้เป็นเสมือนจุด ROI เสริมไว้ก่อนตั้งแต่แรกเพื่อลดระยะเวลาในการตอบสนองของ Pipeline เวลาที่พบการเคลื่อนไหวที่รวดเร็วเกินไป นอกจากนี้วิธีนี้ยังช่วยให้โมเดลสามารถรักษาความสอดคล้องกันได้ทั่วทั้งรูปร่างและป้องกันไม่ให้เกิดความสับสนระหว่างมือซ้ายและมือขวา หรือส่วนที่ต่างกันของร่างกายในแต่ละเฟรม

2.4.2 ประสิทธิภาพการทำงาน

ประสิทธิภาพของ MediaPipe ได้ถูกปรับปรุงและปรับแต่งให้มีประสิทธิภาพสูงสุดโดยการใช้ข้อมูลจากหลายแหล่งต่างกัน ตัวอย่างข้อมูลที่ MediaPipe ใช้เพื่อเพิ่มประสิทธิภาพทำได้โดย

1) การใช้ GPU (Graphics Processing Unit) MediaPipe ใช้ GPU เพื่อประมวลผลข้อมูลที่ซับซ้อนและการทำงานที่ต้องการความเร็วสูง เช่น การประมวลผลภาพเชิงลึก (Deep Learning) และการประมวลผลภาพอื่นๆ ที่ต้องการความสามารถในการคำนวณที่สูง โดยการใช้ GPU ทำให้สามารถประมวลผลข้อมูลได้เร็วขึ้นมาก

2) การใช้ไลบรารีเฉพาะของเครื่องมือ (Specialized Libraries) MediaPipe ใช้ไลบรารีที่ถูกพัฒนาขึ้นมาเฉพาะเพื่อการประมวลผลภาพและวิดีโออย่างมีประสิทธิภาพ เช่น TensorFlow Lite, OpenCV, OpenGL เป็นต้น เมื่อใช้ไลบรารีเหล่านี้ MediaPipe สามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพและมีความแม่นยำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3)การใช้ระบบปฏิบัติการและไลบรารีขั้นสูง MediaPipe ใช้ฟังก์ชันและคลาสที่มีให้ในระบบปฏิบัติการและไลบรารีขั้นสูง เช่น การจัดการหน่วยความจำ, การจัดการกระแสข้อมูล, และการจัดการการเข้าถึงข้อมูลอื่นๆ ที่ช่วยเพิ่มประสิทธิภาพของการทำงาน

4) การใช้ขั้นตอนการประมวลผลที่ปรับแต่ง MediaPipe มีการปรับแต่งขั้นตอนการประมวลผลให้เหมาะสมกับงานและแอปพลิเคชันที่กำลังทำงานอยู่ เพื่อให้มีประสิทธิภาพที่สูงสุดและการทำงานได้ตามความต้องการของผู้ใช้

2.5 ภาษา Python

ภาษาโปรแกรม Python เป็นภาษาโปรแกรมคอมพิวเตอร์ระดับสูง โดยถูกออกแบบมาให้เป็นภาษาสคริปต์ที่อ่านง่าย โดยตัดความซับซ้อนของโครงสร้าง และไวยากรณ์ของภาษาออกไป ในส่วนของการแปลงชุดคำสั่งที่เราเขียนให้เป็นภาษาเครื่อง Python มีการทำงานแบบ Interpreter คือ เป็นการแปลชุดคำสั่งทีละบรรทัด เพื่อป้อนเข้าสู่หน่วยประมวลผลให้คอมพิวเตอร์ทำงานตามที่เราต้องการ นอกจากนั้นภาษาโปรแกรม Python ยังสามารถนำไปใช้ในการเขียนโปรแกรมได้หลากหลายประเภท โดยไม่ได้จำกัดอยู่ที่งานเฉพาะทางใดทางหนึ่ง (General-purpose language) จึงทำให้มีการนำไปใช้กันแพร่หลายในหลายองค์กรใหญ่ระดับโลก เช่น Google, YouTube, Instagram, Dropbox และ NASA เป็นต้น

2.5.1 OpenCV-Python

OpenCV ย่อมาจาก Opensource Computer Vision ที่สามารถใช้งานได้ฟรี โดยความสามารถ คือ ช่วยจัดการกับวิดีโอหรือรูปภาพ เพื่อให้เกิดการเปลี่ยนแปลงกับรูปภาพ อย่างเช่น ทำให้ภาพชัดขึ้น ทำให้ภาพเบลอ หรือช่วยในการลดสัญญาณรบกวน (Noise) ในรูปภาพ และมีเทคนิคที่ใช้ในการประมวลผลภาพที่เรียกว่าการแบ่งประเภทของวัตถุ

2.5.2 Keras

Keras เป็นไลบรารีที่เกิดจากการวิจัยของ Google Lab ที่ได้คิดค้น และพัฒนา Platform ใหม่ขึ้นมาเพื่อเรียกใช้งาน TensorFlow ในเบื้องหลังแทน อันเนื่องมาจากความซับซ้อนในการใช้งาน TensorFlow ทำให้สามารถเขียนโปรแกรมผ่าน Keras ได้โดยตรง Keras เป็น API สำหรับ Deep learning เพื่อให้ง่ายต่อการนำมาใช้ในการสร้าง neural network ซึ่งเขียนด้วยภาษา Python และรองรับการคำนวณ neural network จาก back end หลายประการ และเป็น API ที่ง่ายต่อการเรียนรู้และใช้งานเนื่องจากมี python front end ที่มีระดับของ abstraction สูง พร้อมทั้งยังมีตัวเลือกในการใช้ back end หลายประการเพื่อการคำนวณ ทำให้ Keras ช้ากว่า deep learning frameworks อื่น ๆ แต่ช่วยทำให้ผู้เริ่มต้นใช้งานได้ง่ายขึ้นอย่างมาก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 ภาษา CSS

CSS คือภาษาโปรแกรมมิ่งที่ย่อมาจาก Cascading Style Sheets ซึ่งเป็นภาษาสำหรับกำหนดรูปแบบและการแสดงผลของเว็บไซต์ โดย CSS ใช้สำหรับกำหนดสไตล์ต่างๆ เช่น สีพื้นหลัง (background color), ขนาดและรูปร่างของตัวอักษร, ขนาดและรูปร่างของกล่อง, การจัดวางเนื้อหา, และอื่นๆ เพื่อให้เว็บไซต์มีการแสดงผลที่สวยงามและเป็นระเบียบมากขึ้น

CSS กับ HTML นั้นทำหน้าที่ต่างกัน โดย HTML ใช้สำหรับกำหนดโครงสร้างของเว็บไซต์และเนื้อหา ในขณะที่ CSS ใช้สำหรับกำหนดรูปแบบและสไตล์ของเว็บไซต์ เพื่อให้มีลักษณะที่สวยงามและมีการจัดเรียงที่เหมาะสม การใช้ทั้งสองภาษาร่วมกันช่วยให้สร้างเว็บไซต์ที่มีประสิทธิภาพและมีความน่าสนใจมากยิ่งขึ้น

2.7 ภาษา JavaScript

JavaScript คือภาษาโปรแกรมคอมพิวเตอร์ที่ใช้ในการพัฒนาและปรับปรุงเว็บไซต์และแอปพลิเคชันเว็บ เป็นภาษาที่ถูกนำมาใช้เป็นหลักในการทำให้เว็บไซต์มีประสิทธิภาพและสร้างประสบการณ์ผู้ใช้ที่ดีขึ้น ซึ่ง JavaScript มีความสามารถในการปรับแต่งและควบคุมเนื้อหาบนเว็บไซต์ได้อย่างหลากหลาย เช่น เพิ่ม/ลบ/แก้ไขเนื้อหา HTML, การจัดการกับรูปแบบการแสดงผล (CSS), การโหลดข้อมูลแบบสด (live data loading), การตอบสนองต่อผู้ใช้ (user interaction) เป็นต้น

ข้อมูลที่ JavaScript ใช้ในการประมวลผลสามารถมาจากหลายแหล่ง เช่น

1) ข้อมูลจากผู้ใช้ (User Input) เช่น ข้อมูลที่ผู้ใช้ป้อนผ่านฟอร์ม, การคลิกปุ่ม, หรือการกระทำอื่นๆ บนหน้าเว็บ

2) ข้อมูลจากเซิร์ฟเวอร์ (Server Data) เช่น ข้อมูลที่ถูกส่งมาจากเซิร์ฟเวอร์ผ่าน AJAX หรือการโหลดข้อมูลจาก API

3) ข้อมูลจากฐานข้อมูล (Database) เช่น การดึงข้อมูลจากฐานข้อมูล MySQL, MongoDB, Firebase เป็นต้น

2.7.1 การทำงาน

JavaScript เป็นภาษาโปรแกรมที่ใช้ในการเขียนสคริปต์ฝั่งไคลเอ็นต์ (client-side scripts) บนเว็บไซต์ โดยภาษา JavaScript ใช้ข้อมูลจากหลายแหล่งต่าง ๆ เพื่อประมวลผลและปรับแต่งการทำงานของเว็บไซต์ได้อย่างหลากหลาย ตัวอย่างข้อมูลที่ JavaScript ใช้งานได้เช่น

1) DOM (Document Object Model): JavaScript สามารถเข้าถึงและปรับเปลี่ยนโครงสร้างและเนื้อหาของหน้าเว็บไซต์ผ่าน DOM ซึ่งเป็นโมเดลที่ใช้ในการแสดงโครงสร้างและเนื้อหาของเอกสาร HTML และ XML บนเว็บไซต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ข้อมูลจากฟอร์ม (Form Data): เมื่อผู้ใช้กรอกข้อมูลในฟอร์มบนเว็บไซต์ จะสามารถใช้ JavaScript เพื่อเข้าถึงและประมวลผลข้อมูลที่ผู้ใช้ป้อนเข้ามา

3) HTTP Requests: JavaScript สามารถทำการส่งคำขอไปยังเซิร์ฟเวอร์เพื่อรับหรือส่งข้อมูลจากแหล่งข้อมูลภายนอก เช่น การโหลดข้อมูลจาก API หรือการส่งข้อมูลผ่าน AJAX

4) Cookies และ Local Storage: JavaScript สามารถใช้งาน Cookies และ Local Storage เพื่อเก็บข้อมูลบนเครื่องผู้ใช้ และใช้ข้อมูลเหล่านี้ในการปรับแต่งประสบการณ์การใช้งานของผู้ใช้

5) การปรับแต่ง CSS: ผ่าน JavaScript, สามารถเข้าถึงและปรับแต่งสไตล์ CSS ของหน้าเว็บไซต์ได้ เช่น เปลี่ยนสีพื้นหลัง, ปรับขนาดและตำแหน่งขององค์ประกอบ HTML

JavaScript ใช้ข้อมูลจากหลายแหล่งต่าง ๆ เพื่อให้เว็บไซต์ทำงานได้ตามที่ต้องการและปรับปรุงประสบการณ์การใช้งานของผู้ใช้ได้อย่างมีประสิทธิภาพ

2.8 AJAX

Asynchronous JavaScript และ XML (AJAX) คือการผสมผสานกันของเทคโนโลยีการพัฒนาเว็บแอปพลิเคชันที่ทำให้เว็บแอปพลิเคชันตอบสนองต่อการโต้ตอบของผู้ใช้เร็วขึ้น เมื่อใดก็ตามที่ผู้ใช้ของคุณโต้ตอบกับเว็บแอปพลิเคชัน เช่น เมื่อพวกเขาคลิกปุ่มหรือกล่องกาเครื่องหมาย เบราว์เซอร์จะแลกเปลี่ยนข้อมูลกับเซิร์ฟเวอร์ทางไกล การแลกเปลี่ยนข้อมูลอาจเป็นสาเหตุทำให้หน้าเพจทำงานหนักและขัดจังหวะประสบการณ์ของผู้ใช้ได้ ด้วย AJAX เว็บแอปพลิเคชันจะสามารถส่งและรับข้อมูลในพื้นหลังเพื่อให้มีการรีเฟรชหน้าเว็บเฉพาะเพียงส่วนย่อย ๆ เท่าที่จำเป็นเท่านั้น

AJAX ใช้ JavaScript และ XML เพื่อเปิดใช้งานการเรียกใช้แบบ asynchronous เมื่อเบราว์เซอร์และเซิร์ฟเวอร์แลกเปลี่ยนข้อมูลกัน ต่อไปเราจะอธิบายวิธีที่เบราว์เซอร์แลกเปลี่ยนข้อมูลกันแบบดั้งเดิม และเปรียบเทียบกับการแลกเปลี่ยนข้อมูลด้วย AJAX

1) การแลกเปลี่ยนข้อมูลโดยไม่ใช้ AJAX ในโมเดลแบบดั้งเดิม เบราว์เซอร์ส่งคำขอ HTTP ไปยังฝั่งเซิร์ฟเวอร์เมื่อผู้ใช้ดำเนินการบางอย่าง เว็บเซิร์ฟเวอร์ได้รับและประมวลผลคำขอและส่งข้อมูลที่ปรับปรุงไปยังเบราว์เซอร์ จากนั้นเบราว์เซอร์รีเฟรชหน้าเว็บด้วยข้อมูลใหม่ในวิธีการนี้ เบราว์เซอร์จะโหลดอีกครั้งทั้งหน้าแม้ว่าข้อมูลที่ร้องขอประกอบด้วยการเปลี่ยนแปลงเล็กน้อย นอกจากนี้ เบราว์เซอร์อาจส่งคำขอบ่อยซึ่งทำให้ซอฟต์แวร์เว็บเซิร์ฟเวอร์ทำงานหนักขึ้น

2) การแลกเปลี่ยนข้อมูลกับ AJAX แทนที่จะอัพเดททั้งหน้า AJAX จะใช้ฟังก์ชัน JavaScript เพื่อสร้าง XMLHttpRequest object บนเบราว์เซอร์ จากนั้น AJAX จะรวบรวมข้อมูลหน้าในรูปแบบ XML ซึ่ง XMLHttpRequest object จะถูกส่งไปยังเว็บเซิร์ฟเวอร์ เว็บเซิร์ฟเวอร์ทำการประมวลผลตามการร้องขอและตอบสนองกับข้อมูลที่ร้องขอ สุดท้ายเบราว์เซอร์จะอัปเดตหน้าจอปัจจุบันด้วยข้อมูลล่าสุดโดยไม่ต้องรีเฟรชหน้าเว็บ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9 Django Framework

Django Framework เป็นเฟรมเวิร์ก (framework) สำหรับการพัฒนาเว็บแอปพลิเคชัน (web applications) โดยใช้ภาษา Python เป็นหลัก โดย Django มีความสามารถในการจัดการข้อมูลและสร้างเว็บแอปพลิเคชันขึ้นมาอย่างรวดเร็วและมีประสิทธิภาพ

ข้อมูลที่ Django ใช้ มาจากหลายแหล่งต่าง ๆ ได้แก่

1) ฐานข้อมูล (Database) Django สามารถเชื่อมต่อกับฐานข้อมูลที่หลากหลายได้ เช่น MySQL, PostgreSQL, SQLite, และ Oracle เพื่อจัดการข้อมูลที่เกี่ยวข้องกับแอปพลิเคชัน เช่น ข้อมูลผู้ใช้งาน, รายการสินค้า, หรือข้อมูลที่ใช้ในแอปพลิเคชันต่าง ๆ

2) HTTP Requests Django สามารถรับข้อมูลที่ส่งมาผ่าน HTTP requests เพื่อประมวลผล และส่งคำตอบกลับไปยังผู้ใช้ เช่น การสร้างและอัปเดตข้อมูลผ่านแบบฟอร์ม, การรับข้อมูลจากผู้ใช้ผ่าน AJAX requests

3) ไฟล์และระบบเก็บข้อมูล (Files and Data Storage) Django สามารถจัดการไฟล์และข้อมูลที่ส่งมาผ่านการอัปโหลดไฟล์ โดยสามารถเก็บข้อมูลไฟล์ได้ในระบบไฟล์ของเซิร์ฟเวอร์หรือบริการเก็บข้อมูลในคลาวด์ เช่น Amazon S3, Google Cloud Storage

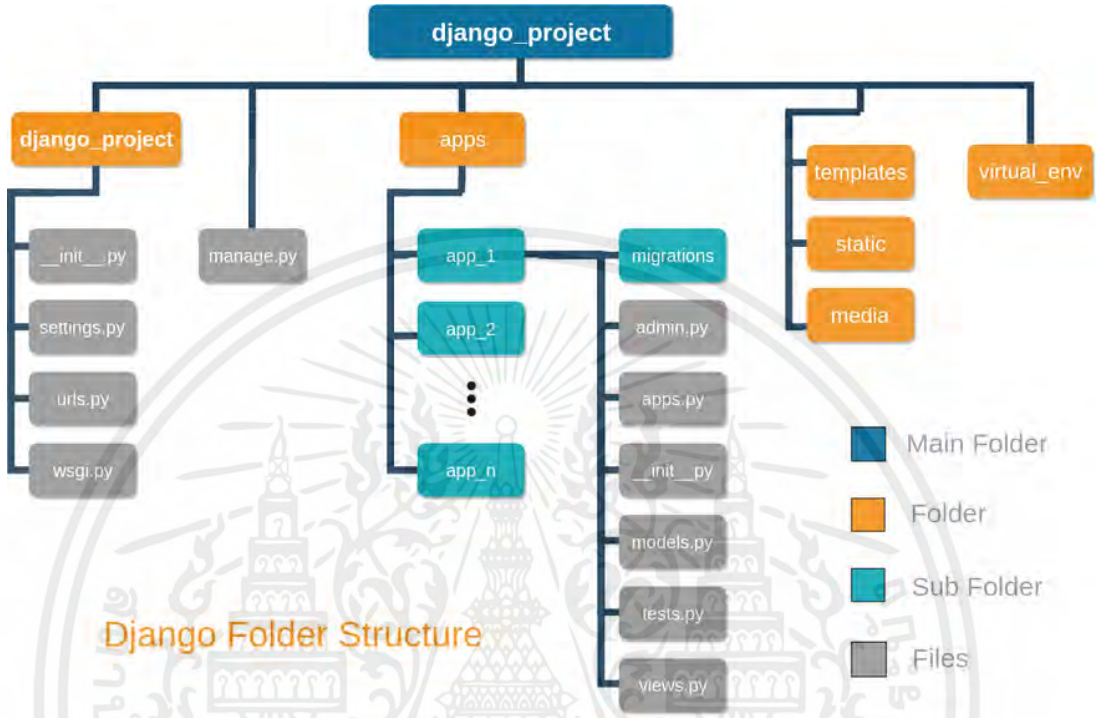
4) การตั้งค่าและการกำหนดค่า (Configuration and Settings) Django สามารถใช้ข้อมูลจากไฟล์การตั้งค่า (settings) เพื่อกำหนดค่าต่าง ๆ ของแอปพลิเคชัน เช่น การกำหนดค่าฐานข้อมูล, คีย์เริ่มต้น, การกำหนดค่าการตั้งค่าความปลอดภัย

Django Framework ใช้ข้อมูลจากหลายแหล่งต่าง ๆ เพื่อสร้างและจัดการเว็บแอปพลิเคชันอย่างมีประสิทธิภาพและเชื่อถือได้อย่างเหมาะสม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9.1 โครงสร้างโปรเจค Django Framework

โครงสร้างโฟลเดอร์ใน Django แสดงดังรูปที่ 2.13



รูปที่ 2.13 โครงสร้างโฟลเดอร์ใน Django

โดยแต่ละไฟล์ในโฟลเดอร์ Django_project มีหน้าที่ดังนี้

1) manage.py คือไฟล์ script สำหรับรันคำสั่งต่าง ๆ ที่เกี่ยวข้องกับ Django เช่น Run Server, Collectstatic, Model & Migration เป็นต้น

2) __init__.py คือ initial ไฟล์หรือไฟล์เปล่า ๆ มีไว้เก็บ Python Package โดยเราสามารถเพิ่ม Script การทำงานเข้าไปในไฟล์นี้ได้

3) settings.py คือไฟล์ที่ใช้สำหรับการตั้งค่าโปรเจคเช่น การตั้งค่าแอพ, เวลา, Path, ฐานข้อมูลที่ใช้ เป็นต้น

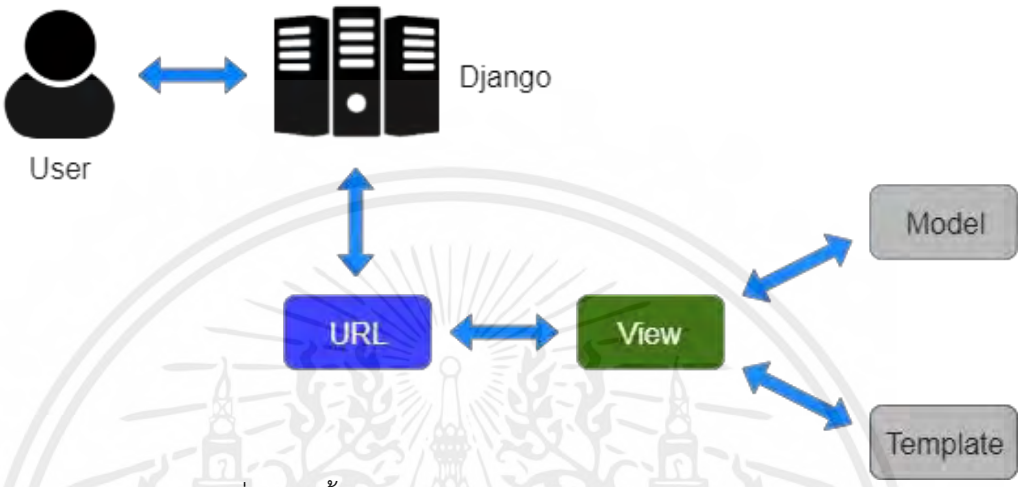
4) urls.py คือไฟล์ที่ใช้เก็บการ routing ของ HTTP request หรือเรียกอีกอย่างว่าการกำหนด url pattern ของ Django project

5) wsgi.py คือไฟล์ที่ใช้เก็บข้อมูลโปรเจคสำหรับการ Deployment (Production)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9.2 การทำงานของ Django Framework

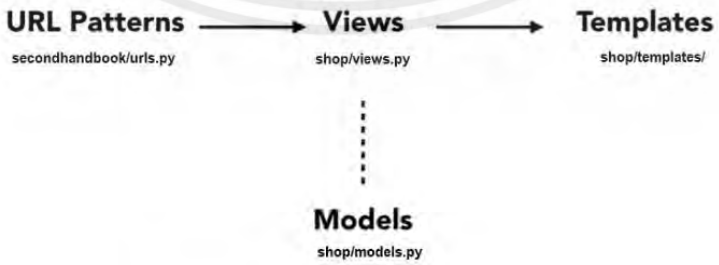
โดยการทำงานเป็นแบบ MTV (Model-View-Template) แสดงการทำงานดังรูปที่ 2.14



รูปที่ 2.14 ขั้นตอนการทำงานแบบ MTV ของ Django

จากรูปที่ 2.14 Model คือส่วนที่เก็บข้อมูลของ Application ส่วน View สำหรับประมวลผลคำสั่งหรือข้อมูลต่าง ๆ (เหมือนกับ Controller) แล้วโยนไปแสดงผลตรงส่วนของ Template และ Template คือหน้าตา Application เป็นส่วนที่ไว้ใช้แสดงผลข้อมูลผลลัพธ์จากการประมวลผลใน View มาแสดงผลในหน้าเว็บร่วมกับ HTML

โดยหลักการทำงานของ Django จะเริ่มจากการอ่าน url แล้วจึงนำไปหา url ที่ถูกต้องใน urls.py หลังจากนั้น จะทำการอ่าน code ใน views.py ในส่วนที่เกี่ยวข้อง โดยจะมีส่วนเสริมคือ models.py ที่มีข้อมูล Structure ของ Object ใน database ที่เราสร้างขึ้นแล้วจึงแสดงผลออกมาเป็นหน้าเว็บตามไฟล์ html ใน template ดังรูปที่ 2.15



รูปที่ 2.15 การทำงานเบื้องต้นของ Django

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Django URL คือส่วนที่เราเอาไว้สำหรับการวาง path ให้ django เพื่อให้ตัว django รับรู้ว่า url ที่ถูกส่งเข้ามาจะต้องนำไปทำงานที่โค้ดตรงส่วนไหนอย่างไร โดยจะถูกเก็บไว้ที่ไฟล์ urls.py โดยตัวอย่าง code ของ path ต่าง ๆ ของ url แสดงดังรูปที่ 2.16

```
urlpatterns = [
    path('admin/', admin.site.urls),
    path('', views.my_login, name='login'),
    path('login/', views.my_login, name='login'),
    path('logout/', views.my_logout, name='logout'),
    path('change_password/', views.change_password, name='change_password'),
    path('shop/', include('shop.urls')),
    path('register/', views.register, name='register')
]
```

รูปที่ 2.16 code ของ path ต่าง ๆ ของ url

Django View คือ ส่วนที่เราเอาไว้เก็บโค้ดสำหรับการทำงานในส่วนต่าง ๆ ตาม path ที่เราได้วางเอาไว้ใน url โดยจะถูกเก็บไว้ที่ไฟล์ views.py โดยตัวอย่าง code ในส่วนของการ register ของ view แสดงดังรูปที่ 2.17

```
def register(request):
    context = {}
    if request.method == 'POST':
        username = request.POST.get('username')
        if User.objects.filter(username = username) :
            context['error1'] = 'ชื่อซ้ำ'
        password = request.POST.get('password')
        email = request.POST.get('email')
        if User.objects.filter(email = email) :
            context['error2'] = 'emailซ้ำ'
        fname = request.POST.get('fname')
        lname = request.POST.get('lname')
        user = User.objects.create_user(username = username, email = email, password = password, first_name = fname, last_name = lname)
        if user:
            login(request, user)
            return redirect('index')

    return render(request, template_name='register.html', context=context)
```

รูปที่ 2.17 code ในส่วนของการ register ของ view

Django model คือ ส่วนที่เราเอาไว้ออกแบบ Structure ของ Object ใน database ที่เราจะนำมาใช้ในการทำงาน โดยจะถูกเก็บไว้ที่ไฟล์ model.py โดยตัวอย่าง code ในส่วนของ model แสดงดังรูปที่ 2.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```python
class Book(models.Model):
    title = models.CharField(max_length=100)
    user = models.ForeignKey(User, on_delete=models.CASCADE)
    price = models.CharField(max_length=100)
    pic = models.URLField(max_length = 500,null = True)
    date = models.DateTimeField(auto_now=True)
    buy = '01'
    sell = '02'
    TYPES = (
        (buy, 'Want to buy'),
        (sell, 'Want to sell')
    )
    open_status = models.BooleanField(default=True)
    status = models.CharField(max_length=2, choices=TYPES, default='01')

    def __str__(self):
        return self.title
```

รูปที่ 2.18 code ในส่วนของ model

Django Template คือการใช้ Django logic เข้าไปช่วยในการทำงานของหน้าเว็บ html ให้สะดวกขึ้นในบางส่วน และ ช่วยในการทำงานให้ง่ายต่อการส่งข้อมูลในหน้าเว็บไปให้ในส่วนการทำงานของ view โดยไฟล์ html จะถูกเก็บไว้ในโฟลเดอร์ templates โดยตัวอย่าง code ในส่วนของ template แสดงดังรูปที่ 2.19

```
<table class="table">
        {% for book in book_list %}
        <ul>
            {% if book.status == '01' %}
            {% if book.open_status == True %}
            <li>{{ book.title }}   
            <a href="{% url 'book_detail' book_id
            {% endif %}
            {% endif %}
        </ul>
        {% endfor %}
</table>
```

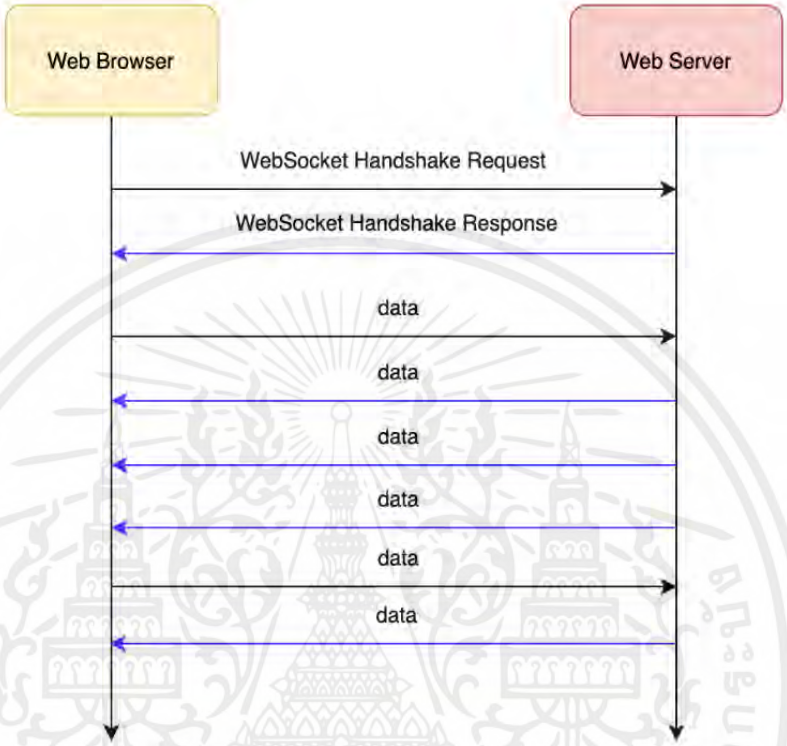
รูปที่ 2.19 Code การใช้ Django template ในการทำไฟล์ html

2.10 WebSocket

เป็น Protocol ที่ทำงานอยู่บน Socket ที่เป็นการเชื่อมต่อแบบ TCP (Transmission Control Protocol) รองรับ Full Duplex หรือ Bidirectional Communication (การสื่อสารแบบสองทิศทาง หมายถึง เป็นผู้รับและผู้ส่งได้ในเวลาเดียวกัน) เป็นมาตรฐานที่ถูกกำหนดโดย IETF (Internet Engineering Task Force) รหัส RFC6455 (The WebSocket Protocol) นิยมนำมาใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กับระบบที่ต้องการการอัพเดทข้อมูลแบบ Realtime เช่น ระบบ Chat, ระบบ Notification, ระบบหุ้น, Game, Developer Tools และอื่น ๆ



รูปที่ 2.20 การรับส่งข้อมูล

1) ถ้า Web Browser (Client) ต้องการข้อมูลจาก Web Server

2) Web Browser จะต้องทำการร้องขอเป็น WebSocket Handshake Request ไปยัง Web Server

3) Web Server ตอบกลับเป็น WebSocket Handshake Response (ไม่มีข้อมูลกลับมาด้วย) จากนั้นจึงเริ่มกระบวนการแลกเปลี่ยนข้อมูลกันระหว่าง Web Browser และ Web Server

4) การรับส่งข้อมูล สามารถทำได้ทั้ง 2 ทิศทาง และไม่ต้องรอให้ฝ่ายใดฝ่ายหนึ่งทำการร้องขอข้อมูลมาก่อน

5) แต่ละฝ่ายสามารถส่ง (Push) ข้อมูลไปมาหากันได้เลย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 2.21 การเชื่อมต่อ WebSocket

การเชื่อมต่อไปยัง WebSocket จะใช้ URI Scheme เป็น ws และ wss จะคล้าย ๆ กับ http และ https โดย ws เป็นการเชื่อมต่อแบบ Non-secure และ wss เป็นการเชื่อมต่อแบบ Secure คือ WebSocket ที่ทำงานอยู่บน TLS (Transport Layer Security) โดย TLS เป็นเทคโนโลยีการเข้ารหัสข้อมูล เพื่อเพิ่มความปลอดภัยในการสื่อสารหรือการส่งข้อมูลผ่าน Network

2.11 BootStrap

BootStrap เป็นเฟรมเวิร์ก (framework) ที่ใช้ในการพัฒนาเว็บไซต์และแอปพลิเคชัน เป็นที่รู้จักกันดีในวงการการพัฒนาเว็บไซต์ เนื่องจากมีความสามารถในการช่วยให้การออกแบบและพัฒนาเว็บไซต์เป็นเรื่องง่ายขึ้น โดย BootStrap มีชุดคำสั่ง (classes) และสไตล์ที่กำหนดมาก่อน (predefined styles) ที่ช่วยให้สร้างเว็บไซต์ที่มีการตอบสนองและลักษณะการแสดงผลที่สวยงามได้อย่างรวดเร็วและง่าย

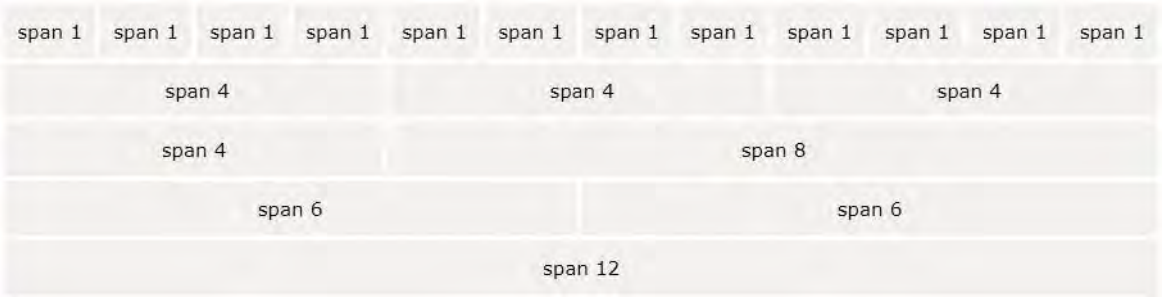
ข้อมูลที่ BootStrap ใช้มาจากการกำหนดคำสั่งและสไตล์ในฐานข้อมูลของ BootStrap เองโดยมีคำสั่งและสไตล์ที่ถูกสร้างขึ้นมาให้ใช้งานในการสร้างเว็บไซต์ ซึ่งประกอบไปด้วยองค์ประกอบต่างๆ เช่น grid system, typography, buttons, forms, modals, navigation bars, dropdowns, tooltips, และอื่นๆ อีกมากมาย

2.11.1 การทำงานของ BootStrap

การทำงาน BootStrap มีขั้นตอนดังนี้

1) BootStrap นั้นจะใช้โครงสร้างในลักษณะ Grid System ในการจัดวางเลย์เอาท์ของเว็บไซต์แสดงดังรูปที่ 2.22 โดยจะแบ่งช่องพื้นที่ในคอนเทนเนอร์ออกเป็น 12 ช่อง เท่า ๆ กัน ในกรณีที่เราต้องการ 3 คอลัมน์ เราก็อาจจะแบ่งพื้นที่ให้เป็น 4 ช่อง จำนวน 3 ชุด ก็จะได้ 12 ช่อง หรือ หากต้องการแบ่งพื้นที่เป็น 2 คอลัมน์โดยมีขนาดเท่า ๆ กัน โดยอาจจะแบ่งพื้นที่ให้เป็น 6 ช่อง จำนวน 2 ชุด ก็จะได้ 12 ช่องเช่นกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.22 Grid System

2) BootStrap ได้สร้างระบบให้เหมาะกับอุปกรณ์เคลื่อนที่ จึงมีการวางระบบจุดเปลี่ยนหน้าจอ (Responsive Breakpoints) โดยการใช้ media queries ไว้ตามจุดต่าง ๆ ให้ใช้งาน โดย Bootstrap จะใช้ความกว้างขั้นต่ำของหน้าจอ (minimum viewport widths) ในการกำหนดขอบเขตของหน้าจอ และขยายเพิ่มขึ้นตามขนาดหน้าจอที่เปลี่ยนไป เช่น มือถือจะมีขนาดหน้าจอมือถือจะอยู่ต่ำกว่า 576px หรือถ้าเป็นมือถือที่ใช้ดูในแนวนอนจะอยู่ระหว่าง 576px ถึง 767px หรือถ้าเป็นแท็บเลตความกว้างของหน้าจอจะอยู่ที่ 768px จนถึง 991px และถ้าเป็น 992px ขึ้นไปก็จะเป็นหน้าจอคอมพิวเตอร์

3) BootStrap ทำการสร้าง Class ของ CSS จำนวนมากมาให้ใช้งาน โดย Class แต่ละตัวนั้นมีการตั้งค่าไว้เรียบร้อยแล้ว ทำให้เราไม่จำเป็นต้องไปพิมพ์ค่าต่าง ๆ ในไฟล์ CSS อีก ซึ่งจะทำให้เราสามารถใช้งานได้ง่ายและรวดเร็วขึ้น ยกตัวอย่างเช่น

- col-3 เป็นการระบุว่าคอลัมน์นี้ใช้พื้นที่ 3 ช่อง จากทั้งหมด 12 ช่อง
- text-right เป็นการระบุว่าให้ข้อความอยู่ด้านขวา ซึ่ง BootStrap จะใส่ค่า CSS ให้เป็น text-align: right
- img-fluid เป็นการระบุให้ class นี้จะทำการใส่ค่า max-width: 100% และ height: auto ลงไปในรูปที่เราต้องการ และทำให้รูปนั้นเป็น responsive ในทุกอุปกรณ์

4) BootStrap มีเครื่องมือหรือองค์ประกอบที่ให้เราใช้งานได้ในทันที เช่น Media Object ที่เป็นชุดเครื่องมือที่เอาไว้วางรูปและเนื้อหาอยู่ในชุดเดียวกัน โดย BootStrap จะให้ตัวอย่างของ HTML และ CSS มาพร้อมกันเพื่อให้สามารถคัดลอกแล้วนำไปใช้ได้ทันที

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การออกแบบและการจัดทำปริญญานิพนธ์

3.1 การออกแบบ

3.1.1 การออกแบบการเก็บ Dataset

เขียนโปรแกรมตรวจจับมือทั้งสองข้างจากกล้องเว็บแคมโดยใช้ Hand Landmarks ของ mediapipe ใส่กรอบสี่เหลี่ยมรอบมือที่ตรวจจับได้และบันทึกภาพ ดังแสดงในรูปที่ 3.1

รูปที่ 3.1 ท่าทางภาษามือที่ตรวจจับและบันทึก

จัดเก็บภาพที่ถูกบันทึกในโฟลเดอร์ภาษามือในแต่ละคำ ดังแสดงในรูปที่ 3.2

รูปที่ 3.2 โฟลเดอร์ที่บันทึกภาษามือแต่ละคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาพภาษามือที่จัดเก็บในโฟลเดอร์ โดยจัดเก็บคำละ 200 ภาพ ดังแสดงในรูปที่ 3.3

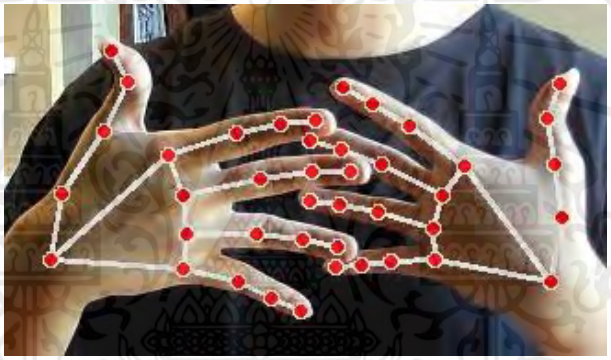
รูปที่ 3.3 รูปภาษามือที่บันทึกในโฟลเดอร์

ตัวอย่างท่าทางและความหมายของรูปที่นำมาฝึกสอนจำนวน 10 คำ

รูปที่ 3.4 ท่าทางภาษามือคำว่า Worry

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

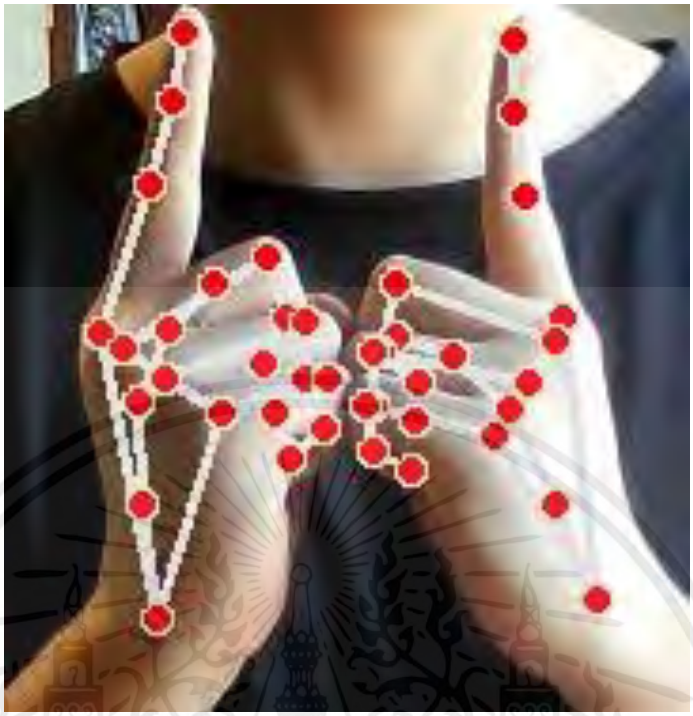
รูปที่ 3.5 ท่าทางภาษามือคำว่า Thank you



รูปที่ 3.6 ท่าทางภาษามือคำว่า Never mind

รูปที่ 3.7 ท่าทางภาษามือคำว่า Good luck

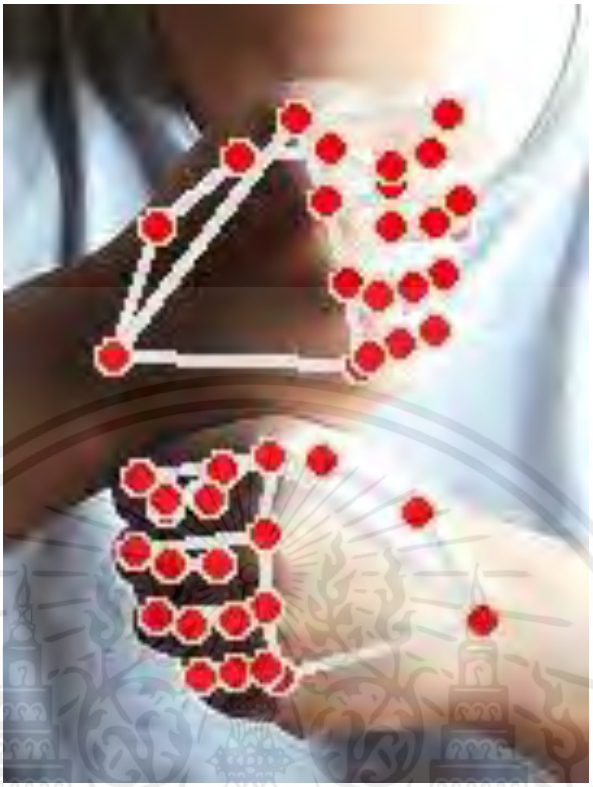
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.8 ท่าทางภาษามือคำว่า Meet

รูปที่ 3.9 ท่าทางภาษามือคำว่า Sorry

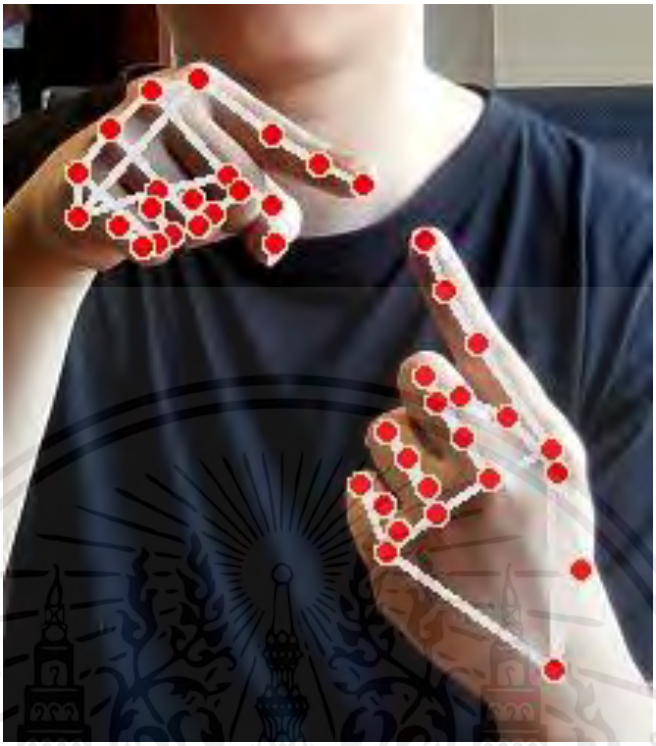
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.10 ท่าทางภาษามือคำว่า Working

รูปที่ 3.11 ท่าทางภาษามือคำว่า Dangerous

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.12 ท่าทางภาษามือคำว่า Appropriate

รูปที่ 3.13 ท่าทางภาษามือคำว่า Computer

จากนั้นจะนำรูปภาพ Dataset ที่อยู่ในโฟลเดอร์เข้าสู่โมเดลเพื่อทำฝึกสอนและทดสอบโมเดลต่อไป

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2 การออกแบบการฝึกสอนและทดสอบโมเดล

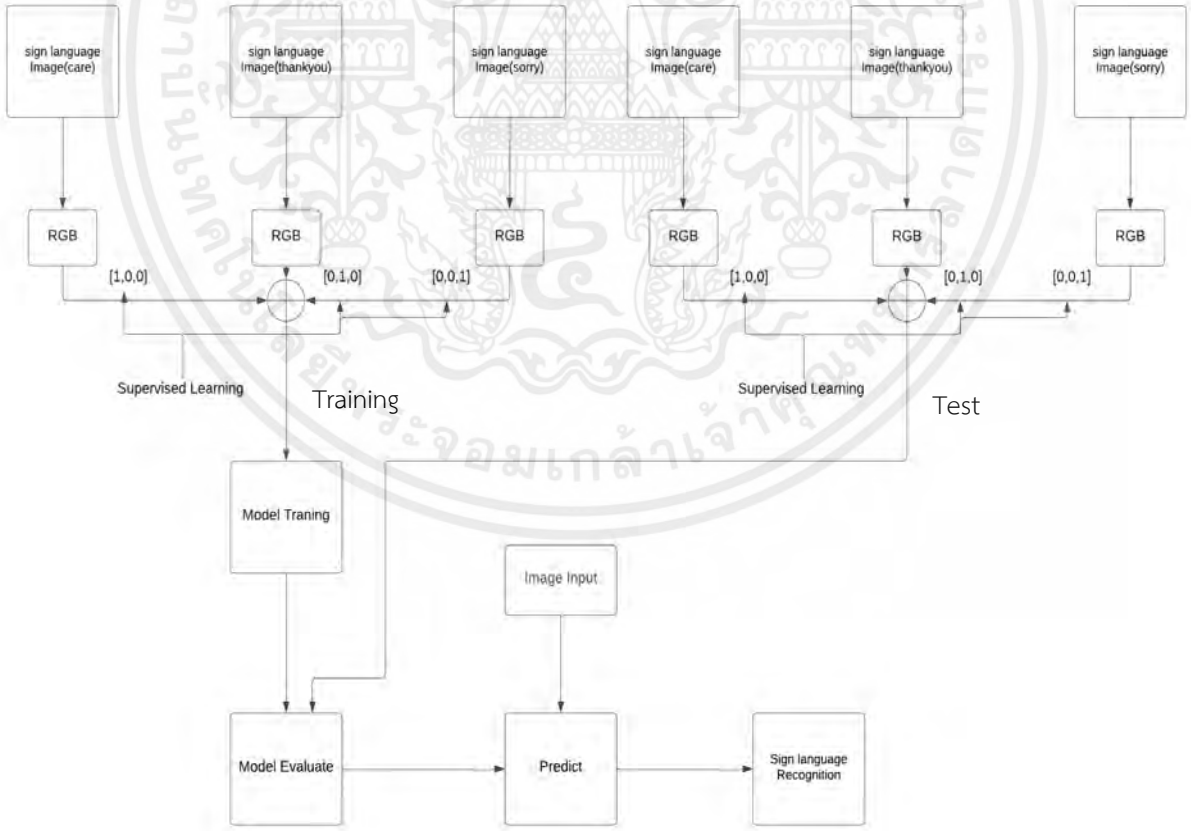
ในขั้นตอนการฝึกสอนและทดสอบโมเดลจำเป็นต้องทำการแบ่งชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดลเพื่อใช้ในการทดสอบความแม่นยำของโมเดล โดยทำการแบ่งชุดข้อมูลออกเป็น 3 ส่วน [1] ดังนี้

1) Training set จำนวน 160 ภาพ (80%)

2) Validation set จะทำการแบ่งมาจาก Training set 10%

3) Test set จำนวน 40 ภาพ (20%)

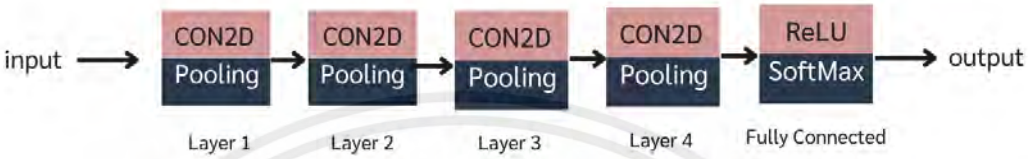
ขั้นตอนการทำ Sign Language Recognition แสดงได้ดังรูปที่ 3.14 โดยเริ่มจากการนำรูปที่เก็บ Dataset ไว้มาทำการระบุคลาสเพื่อให้โมเดลได้เรียนรู้ว่าท่าทางไหนคือรูปภาษามือในแต่ละคำ หลังจากผ่านขั้นตอนการฝึกสอนและทดสอบโมเดล จะทำการนำโมเดลไปใช้ในการทำนายรูปภาพเพื่อแยกรูปภาพภาษามือในแต่ละคำจากที่ฝึกสอนโมเดลไว้



รูปที่ 3.14 บล็อกไดอะแกรมการทำ Sign Language Recognition

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

CNN Model ที่ถูกฝึกสอนและทดสอบสำหรับแยกรูปภาพภาษามือในแต่ละคำ โดยทำการออกแบบ CNN Model เป็นจำนวน 6 Layer ประกอบไปด้วยส่วนที่เป็น Convolution มี ReLU จำนวน 4 Layer และส่วนที่เป็น Fully Connected มี ReLU กับ SoftMax อย่างละ 1 Layer แสดงดังรูปที่ 3.15



รูปที่ 3.15 Layer ของ CNN Model สำหรับแยกรูปภาพภาษามือแต่ละคำ

แต่ละ Layer ของ CNN Model ที่ออกแบบมีคุณสมบัติดังต่อไปนี้ โดยมีตัวอย่างของพารามิเตอร์ที่กำหนดให้ขนาดของอินพุตเข้ามีขนาด 32x32 ดังรูปที่ 3.16

1) Input: ข้อมูลรูปภาพภาษามือที่ได้จากการทดลองปรับขนาด 128x128, 56x56, 32x32, 28x28, 24x24 หน่วยเป็น pixels ที่เหมาะสมที่สุด โดยการทดลองนี้ใช้ขนาด 32x32

2) Layer 1-4: เป็นส่วนของ Convolution ที่ประกอบไปด้วย Kernel size ขนาด 3x3 pixels ที่มี Step ในการเลือกคือ 1 pixels โดยการเติม Padding 0 เข้าไป และใช้ฟังก์ชันของ ReLU ในการคำนวณก่อนส่งไปยังการทำ Max pooling เพื่อลดจำนวนพารามิเตอร์ที่ใช้ ซึ่งออกแบบขนาดการ Max pooling ให้มีขนาด 2x2 pixels และเลื่อนที่ละ 2-pixel แบบไม่เติม Padding ทำให้มีขนาดลดลง 4 เท่า

3) Fully Connected: เริ่มจากการทำ Flatten เพื่อจัดเรียงข้อมูลให้เป็นเวกเตอร์ 1 มิติจากนั้นนำเข้าสู่ฟังก์ชัน ReLU เพื่อคำนวณ และส่งไปยังฟังก์ชันของ SoftMax เพื่อทำการคำนวณหาค่าเพื่อใช้ในการทำนาย

4) Output: จะแสดงค่า 0-1 โดยผลรวมของความน่าจะเป็นของคลาสทั้งหมดต้องเท่ากับ ซึ่งค่าที่มากที่สุดจะถูกนำมาเป็นการระบุคลาสของข้อมูลที่ใช้ทำนาย

5) การเพิ่มประสิทธิภาพประกอบด้วย Optimizer ใช้ Adam ในส่วนของ Loss Function ใช้ Cross entropy

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 32, 32, 32)	896
max_pooling2d (MaxPooling2D)	(None, 16, 16, 32)	0
conv2d_1 (Conv2D)	(None, 16, 16, 128)	36992
max_pooling2d_1 (MaxPooling2D)	(None, 8, 8, 128)	0
conv2d_2 (Conv2D)	(None, 8, 8, 256)	295168
max_pooling2d_2 (MaxPooling2D)	(None, 4, 4, 256)	0
conv2d_3 (Conv2D)	(None, 4, 4, 512)	1180160
max_pooling2d_3 (MaxPooling2D)	(None, 2, 2, 512)	0
flatten (Flatten)	(None, 2048)	0
dense (Dense)	(None, 128)	262272
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 100)	12900

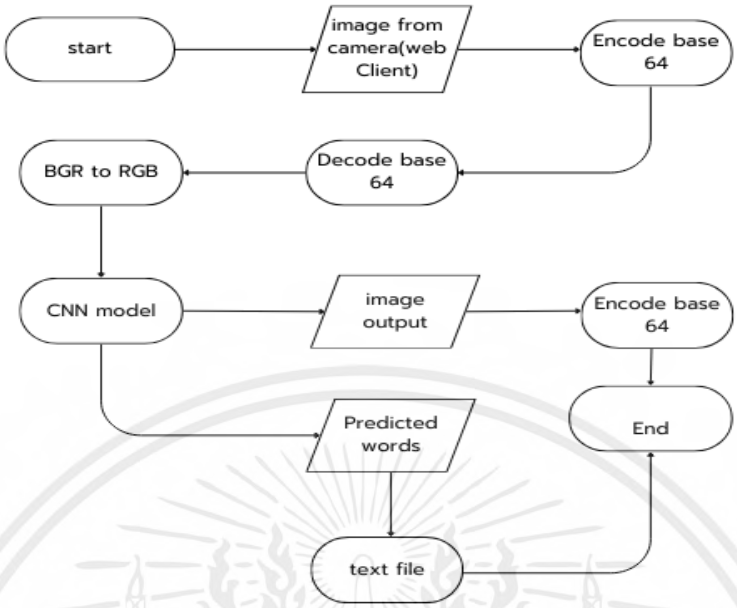
Total params: 1,788,388
Trainable params: 1,788,388
Non-trainable params: 0

รูปที่ 3.16 พารามิเตอร์ของแต่ละ Layer ของโมเดลที่ใช้ในการแยกรูปภาพภาษามือ

3.1.3 การออกแบบการทำนายภาษามือจากรูปภาพ

การทำนายภาษามือจากรูปภาพโดยมีอินพุตเป็นรูปภาพภาษามือแต่ละเฟรมจากกล้อง นำมาทำนายใน Model ในฝั่ง Server ที่ทำการฝึกสอนจะได้เอาต์พุตเป็นข้อความที่แปลภาษามือและส่งไปแสดงผลที่ Client โดยบล็อกไดอะแกรมการทำนายภาษามือจากรูปภาพ แสดงได้ดังรูปที่ 3.17

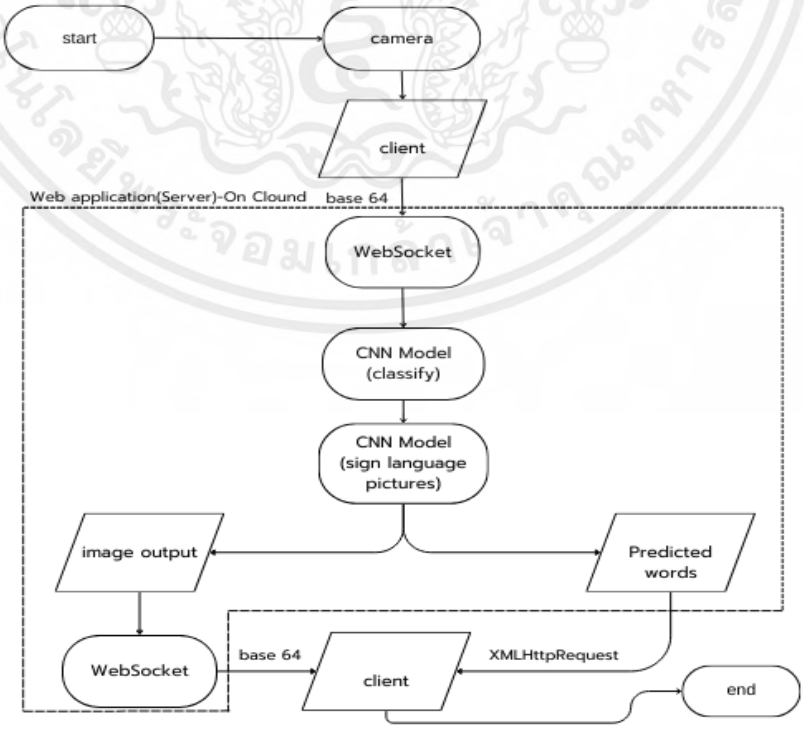
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.17 บล็อกไดอะแกรมการทำนายภาษามือจากรูปภาพ

3.1.4 การออกแบบการแสดงผลภาษามือที่แปลได้ผ่านเว็บไซต์แบบเรียลไทม์

ขั้นตอนการแสดงผลเว็บไซต์แปลภาษามือแบบเรียลไทล์ คือ การใช้ WebSocket protocol ในการรับส่งข้อมูลรูปภาพจากกล้องและไปประมวลผลแบบเรียลไทม์ โดยแสดงคำที่แปลได้จากภาษามือในแต่ละเฟรมและส่งคำที่ทำนายได้ในแต่ละเฟรมไปแสดงผลที่หน้าเว็บแบบเรียลไทม์ แสดงแผนผังการ แสดงผลของเว็บไซต์ได้ดังรูปที่ 3.18

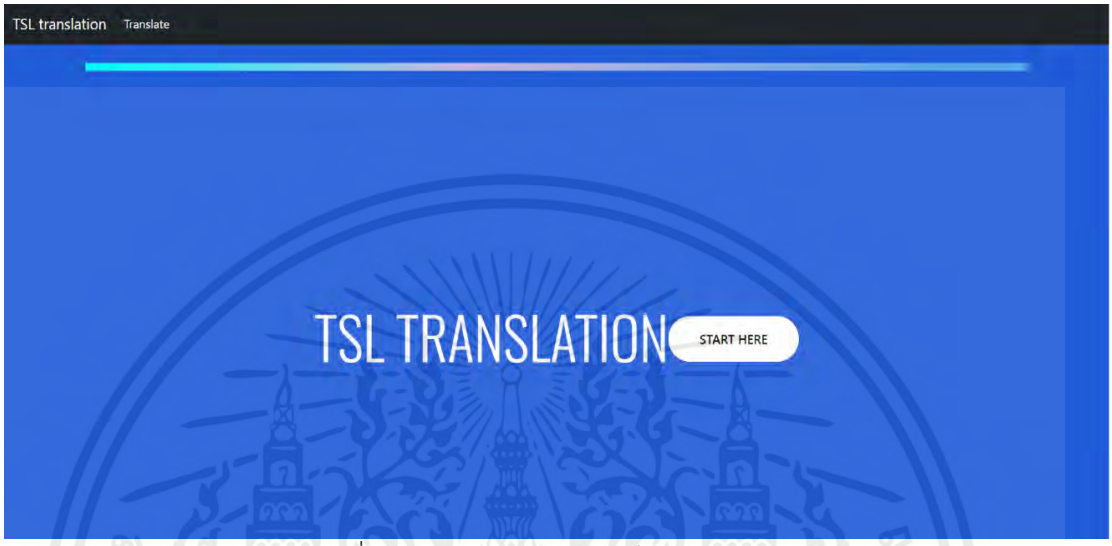


รูปที่ 3.18 บล็อกไดอะแกรมการแสดงผลบนเว็บไซต์

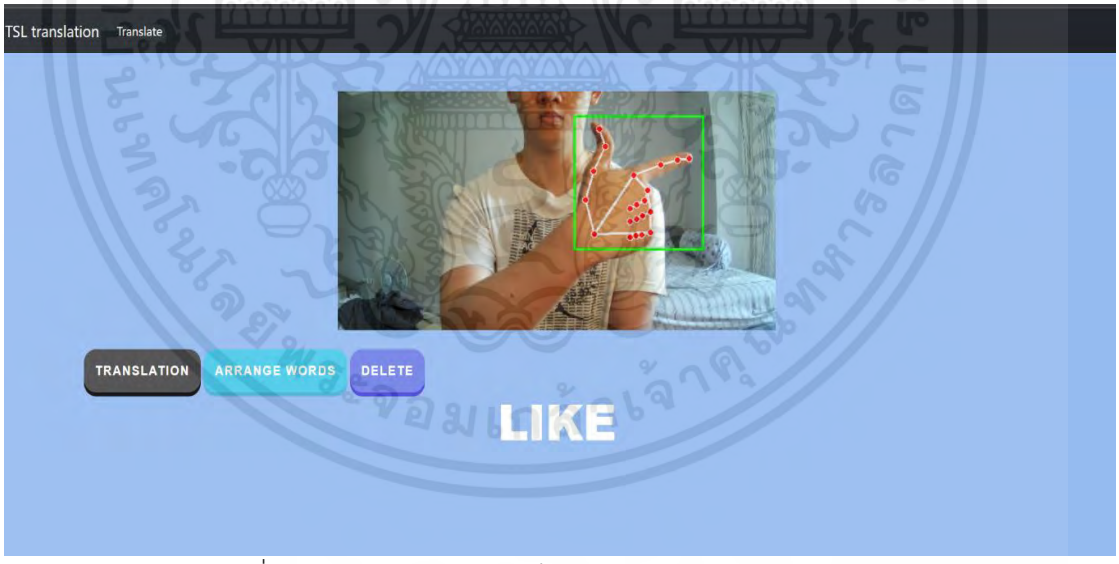
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.5 การออกแบบเว็บแอปพลิเคชันที่ใช้แสดงผล

ใช้ Django framework ในการสร้างเว็บแอปพลิเคชันเพื่อแสดงผลคำแปลภาษามือและภาพจากกล้องแบบแบบเรียลไทม์ แสดงได้ดังรูปที่ 3.19 และ 3.20



รูปที่ 3.19 การแสดงผลของเว็บไซต์หน้าแรก



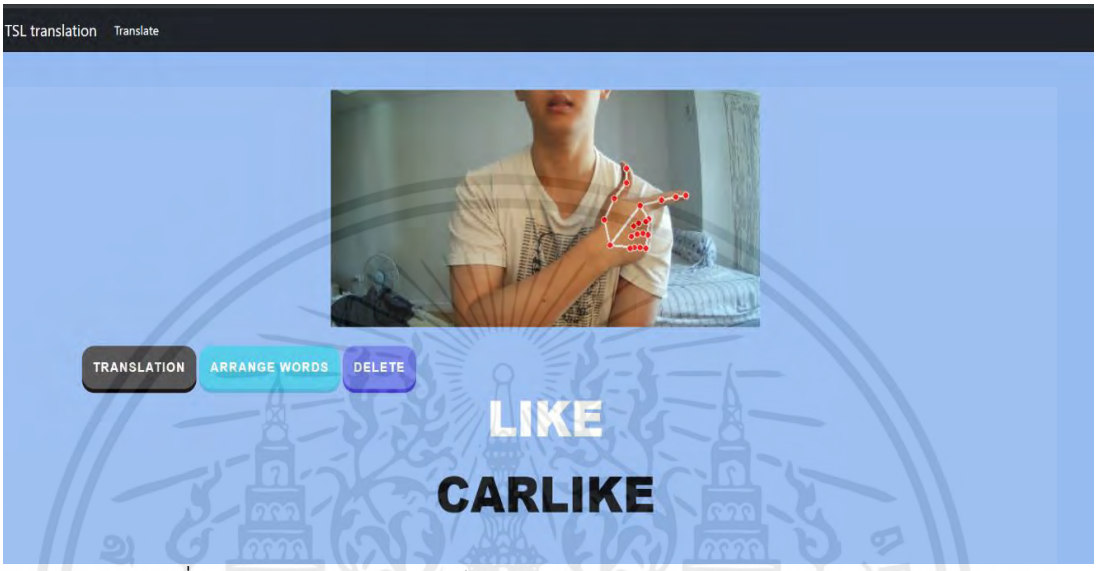
รูปที่ 3.20 การแสดงผลของเว็บไซต์ในส่วนการทำนายภาษามือ

เมื่อต้องการใช้งานเว็บไซต์แปลภาษามือให้กดปุ่ม START HERE เพื่อเข้าสู่หน้าแรกของเว็บไซต์ เว็บไซต์จะขออนุญาตเข้าถึงกล้อง เมื่ออนุญาตให้เข้าถึงกล้องแล้วจึงจะเข้าสู่ส่วนของการแปลภาษามือ โดยการแปลภาษามือมี 3 ฟังก์ชันหลักดังนี้

1.) เมื่อกดปุ่ม Translation จะเริ่มแปลภาษามือแบบเรียลไทม์ โดยท่าทางที่ทำนายได้จะแสดงผลเป็นข้อความสีขาว ดังที่แสดงในรูป 3.20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.) เมื่อกดปุ่ม Arrange words จะนำข้อความสีขาวมาเก็บเป็นข้อความสีดำ โดยข้อความสีดำจะไม่หายหรือเปลี่ยนตามท่าทางมือที่กล้องตรวจจับได้ และสามารถเรียงคำต่อกันได้เพื่อให้ง่ายต่อการเข้าใจ ดังแสดงในรูปที่ 3.21



รูปที่ 3.21 การแสดงผลของเว็บไซต์ในการเรียงคำจากปุ่ม Arrange words

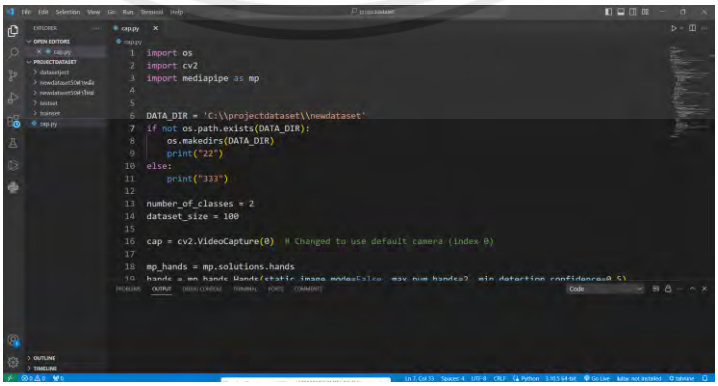
3.) เมื่อกดปุ่ม Delete จะทำการลบข้อความสีดำทั้งหมด

3.2 เครื่องมือที่ใช้ในการทดลอง

ในปริญญานิพนธ์นี้ มีอุปกรณ์และเครื่องมือที่ใช้ในการทดลอง ดังนี้

3.2.1 Visual Studio Code

Visual Studio Code ใช้สำหรับการเขียนโปรแกรมโดยใช้ภาษา Python ในการตรวจจับภาษามือ แปลภาษามือโดยใช้การเรียนรู้เชิงลึก และสร้างเว็บไซต์ ดังแสดงในรูปที่ 3.22



รูปที่ 3.22 โปรแกรม Visual Studio Code

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2.2 Webcam

Webcam ใช้สำหรับการเก็บ Dataset เพื่อใช้ฝึกฝนและทดสอบโมเดลแปลภาษามือ ดังแสดงในรูปที่ 3.23

- รุ่น Fifine K420
- Frame rate: 30 FPS
- maximum dynamic resolution: 1440P/2k

รูปที่ 3.23 Webcam รุ่น Fifine K420

3.3 การจัดเก็บผลการทดลอง

3.3.1 การทดสอบความแม่นยำในการฝึกฝนโมเดล

ทดสอบค่าพารามิเตอร์ต่าง ๆ จนกว่าจะพบค่าที่เหมาะสม เพื่อให้โมเดลทำงานได้อย่างมีประสิทธิภาพ

3.3.2 การทดสอบความแม่นยำในการตรวจจับมือ

ทดสอบความแม่นยำในการตรวจจับมือของ MediaPipe

3.3.3 การทดสอบความแม่นยำในการทำนายภาษามือบนเว็บไซต์

นำโมเดลมาทำนายจริงเพื่อทดสอบโมเดลที่ผ่านการฝึกฝนมาแล้ว โดยวัดความแม่นยำและประสิทธิภาพของโมเดลเมื่อเผชิญกับข้อมูลใหม่ที่ไม่เคยเห็นมาก่อน เพื่อประเมินความถูกต้องของโมเดลเมื่อใช้งานในสถานการณ์จริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

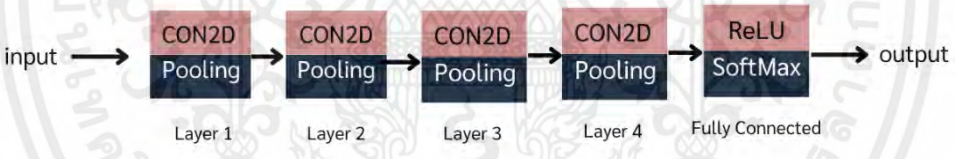
บทที่ 4

ผลการทดลอง

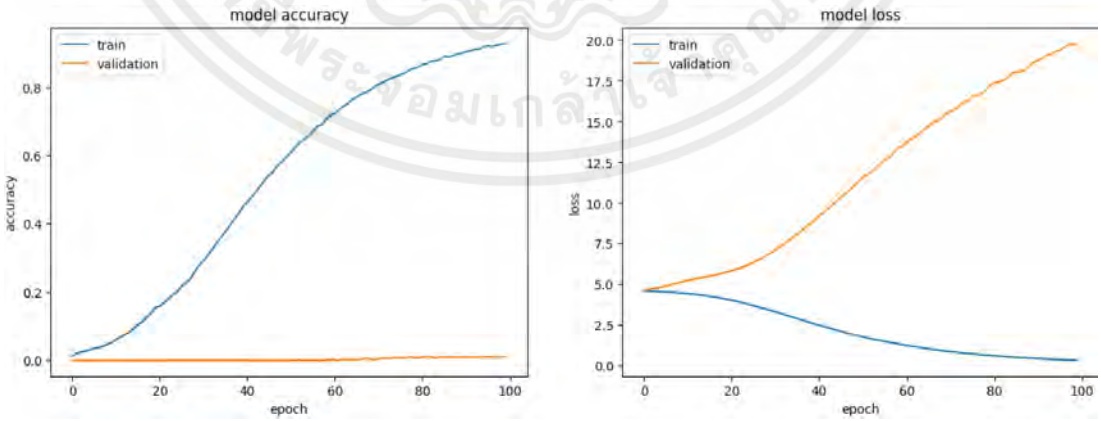
ผู้จัดทำได้ทำการเก็บผลการทำงานของระบบ โดยแบ่งการทดลองและจัดเก็บผลการทดลองเป็นส่วนๆ ดังต่อไปนี้

4.1 การทดสอบความแม่นยำการฝึกฝนโมเดล

ในการฝึกสอนและทดสอบโมเดลได้ทำการสร้าง CNN Model เป็นจำนวน 6 Layer แสดงดังรูปที่ 4.1 โดยกำหนดอินพุตที่ใช้ฝึกสอนให้กับโมเดลเป็นรูปภาพภาษามือ 200 รูป และรูปภาพที่ใช้ฝึกสอนจำนวน 160 รูป โดยก่อนทำการฝึกสอนโมเดลได้ทำการทดลองปรับขนาดของรูปภาพที่เป็นอินพุตเท่ากับ 32x32 pixel ก่อนเข้าสู่โมเดล เพื่อให้ขนาดที่เหมาะสมและปรับให้มีขนาดที่เท่ากันทุกรูป โดยได้ทำการฝึกสอนโมเดลเป็นจำนวน Epoch = 100 รอบได้กราฟแสดง Accuracy และ Loss กับจำนวนรอบที่ฝึกสอน ดังรูปที่ 4.2



รูปที่ 4.1 CNN Model

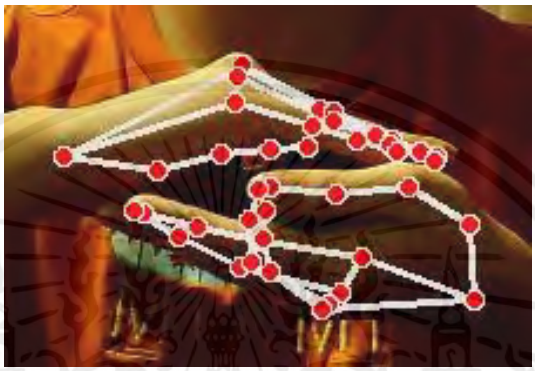


รูปที่ 4.2 กราฟแสดงค่า Loss, Accuracy กับจำนวนรอบที่ฝึกสอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 การทดสอบความแม่นยำของการตรวจจับมือ

ทดสอบความแม่นยำในการตรวจจับมือ โดยทำการบันทึกรูปที่ตรวจจับมือลงในโฟลเดอร์จำนวน 500 รูป จากนั้นตรวจรูปมือที่บันทึกออกมาได้ ได้รูปที่ตรวจจับมือทั้งสองข้างครบ ดังแสดงในรูปที่ 4.3 จำนวน 485 รูป และรูปที่ตรวจจับมือได้ไม่ครบ ดังแสดงในรูปที่ 4.4 จำนวน 15 รูป



รูปที่ 4.3 MediaPipe ตรวจจับมือทั้งสองข้างครบ

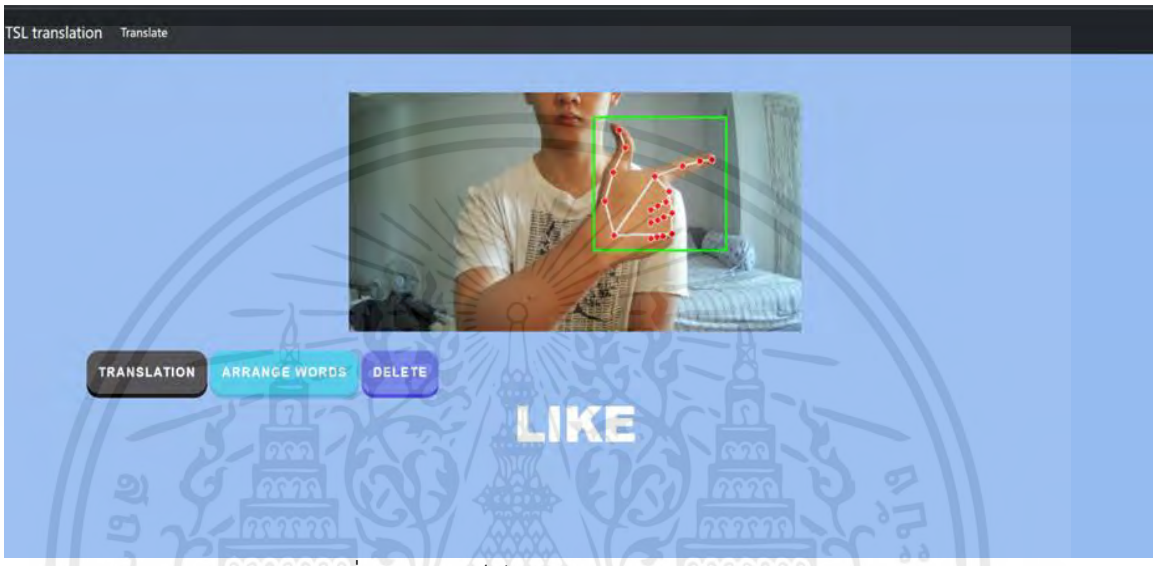
รูปที่ 4.4 MediaPipe ตรวจจับมือทั้งสองข้างไม่ครบ

จากการทดสอบพบว่าค่าความแม่นยำในการตรวจจับมือของ MediaPipe มีค่าเท่ากับ 97 เปอร์เซ็นต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การทดสอบความแม่นยำในการทำนายภาษามือบนเว็บไซต์

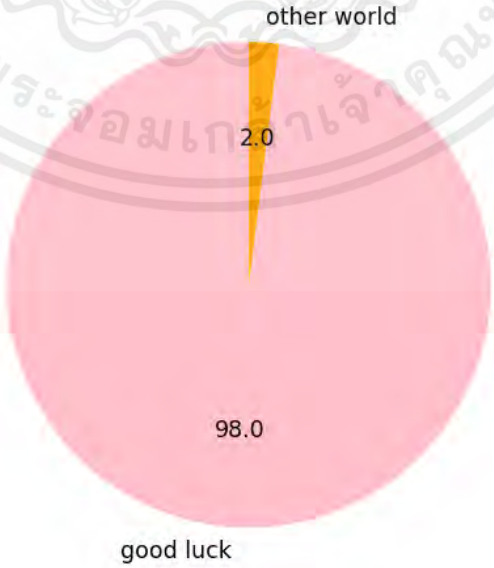
ทดสอบการทำนายภาษามือผ่านเว็บไซต์ที่สร้างขึ้นผ่านกล้องแบบเรียลไทม์ ดังแสดงในรูปที่ 4.5 จะทำทดสอบภาษามือผ่านกล้องเป็นจำนวน 10 คำ จากภาษามือทั้งหมด 100 คำ



รูปที่ 4.5 หน้าเว็บไซต์ระหว่างทำนายภาษามือ

4.3.1 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Good luck

ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.6 โดยทำนายถูกทั้งหมด 49 ครั้ง

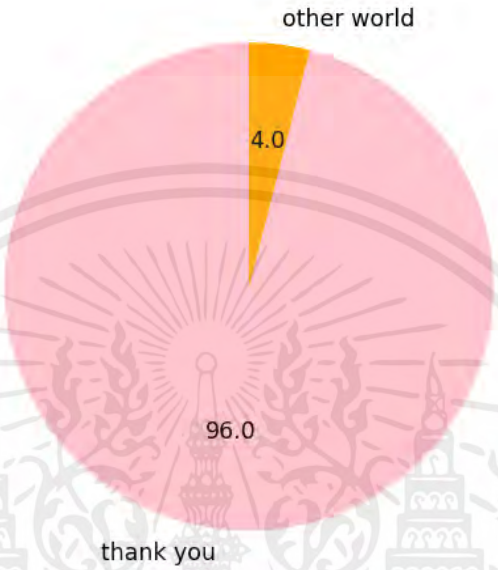


รูปที่ 4.6 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Good luck

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.2 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Thank you

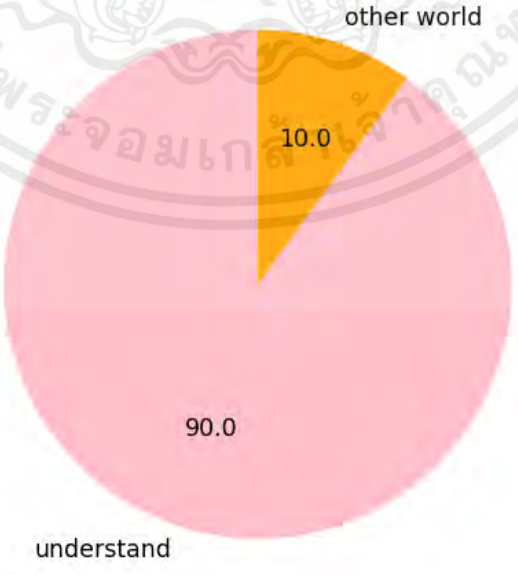
ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.7 โดยทำนายถูกทั้งหมด 48 ครั้ง



รูปที่ 4.7 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Thank you

4.3.3 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Understand

ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.8 โดยทำนายถูกทั้งหมด 45 ครั้ง

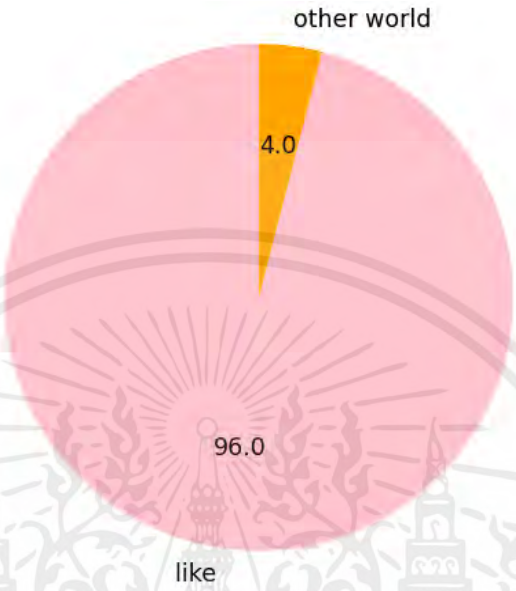


รูปที่ 4.8 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Understand

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.4 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Like

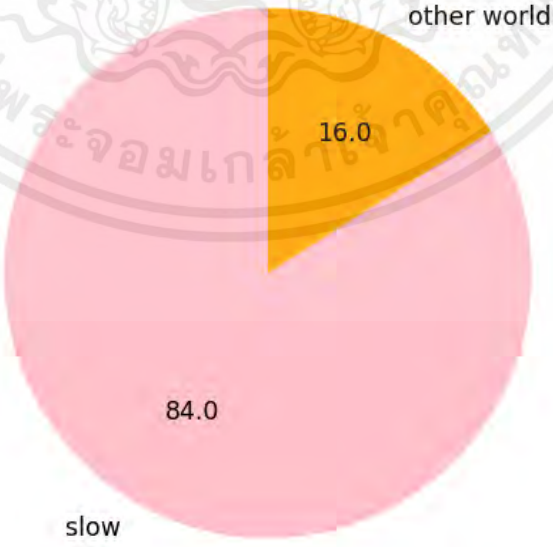
ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.9 โดยทำนายถูกทั้งหมด 48 ครั้ง



รูปที่ 4.9 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Like

4.3.5 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Slow

ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.10 โดยทำนายถูกทั้งหมด 42 ครั้ง

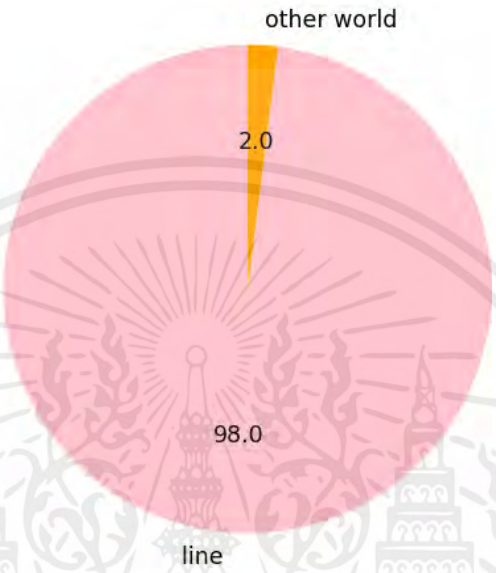


รูปที่ 4.10 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Slow

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.6 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Line

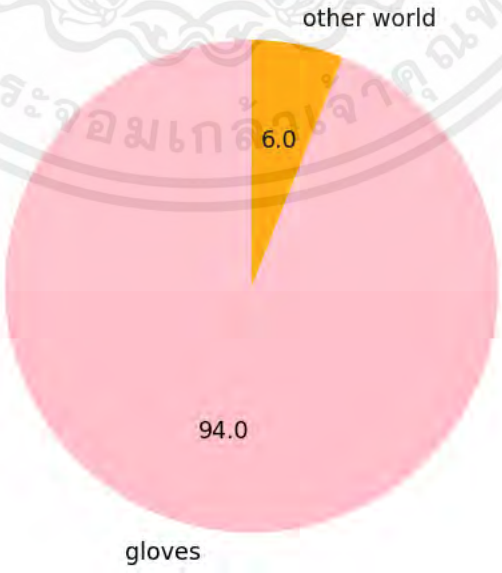
ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.11 โดยทำนายถูกทั้งหมด 49 ครั้ง



รูปที่ 4.11 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Line

4.3.7 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Gloves

ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.12 โดยทำนายถูกทั้งหมด 47 ครั้ง

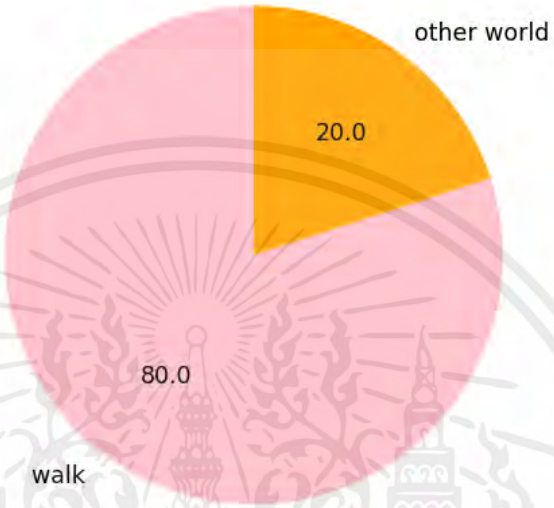


รูปที่ 4.12 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Gloves

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3.8 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Walk

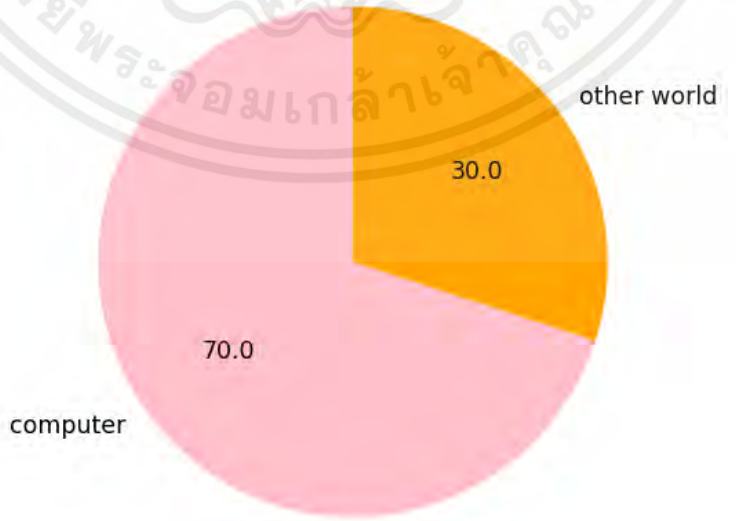
ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.13 โดยทำนายถูกทั้งหมด 40 ครั้ง



รูปที่ 4.13 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Walk

4.3.9 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Computer

ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.14 โดยทำนายถูกทั้งหมด 35 ครั้ง

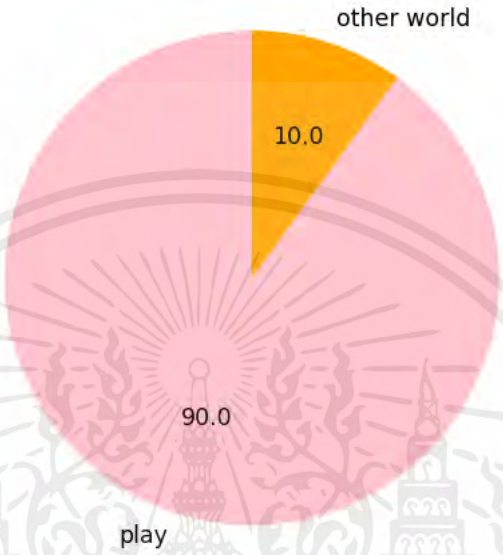


รูปที่ 4.14 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Computer

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

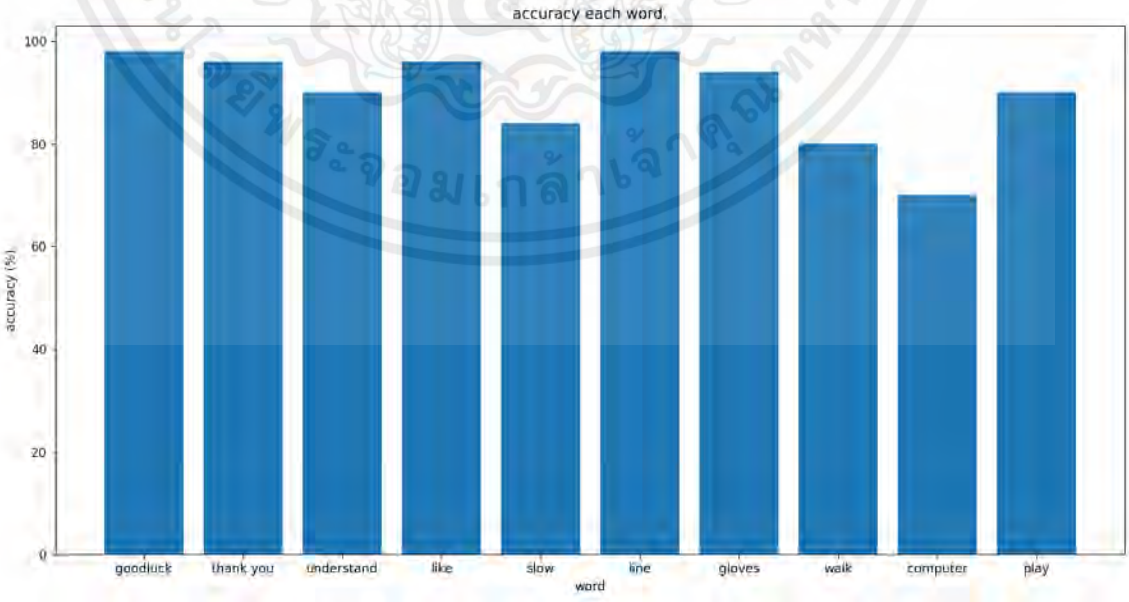
4.3.10 ทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ของคำว่า Play

ทำการทำนายจำนวน 50 ครั้ง แสดงผลลัพธ์ดังรูปที่ 4.15 โดยทำนายถูกทั้งหมด 45 ครั้ง



รูปที่ 4.15 แผนภูมิวงกลมแสดงความถูกต้องในการทำนายคำว่า Play

จากการทดสอบความแม่นยำในการทำนายภาษามือผ่านเว็บไซต์ทั้งหมด 10 คำ ได้แผนภูมิแท่งแสดงความถูกต้องในการทำนายของแต่ละคำ ดังแสดงในรูปที่ 4.16



รูปที่ 4.16 แผนภูมิแท่งแสดงความถูกต้องในการทำนายของแต่ละคำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผล

ปริญญานิพนธ์นี้นำเสนอการศึกษาและการใช้งานการเรียนรู้เชิงลึกร่วมกับเทคโนโลยี Computer Vision เพื่อแก้ไขปัญหาการสื่อสารกับผู้ใช้ภาษามือ ผ่านการตรวจจับมือทั้งสองข้างและบันทึกภาษามือในแต่ละคำ เพื่อนำรูปที่ได้ไปฝึกสอนและใช้โมเดลที่ได้ในการแปลภาษามือผ่านเว็บไซต์ เพื่อเพิ่มความสะดวกรวดเร็วในการสื่อสารกับผู้ใช้ภาษามือ

จากการทดสอบความแม่นยำในการตรวจจับมือของโมเดล MediaPipe พบว่ามีค่าความแม่นยำอยู่ที่ 97 เปอเซ็นต์ โดยใช้โมเดลตรวจจับภาพมือทั้งหมด 500 รูป พบว่าตรวจจับมือทั้งสองข้างครบ 485 รูป และตรวจจับมือทั้งสองข้างไม่ครบ 15 รูป

ในการทดลองทำนายภาษามือแบบเรียลไทม์ผ่านเว็บไซต์เป็นจำนวน 10 คำ โดยทดสอบการทำนายแต่ละคำจำนวน 50 เฟรม และคำนวณค่าความถูกต้องเป็นเปอร์เซ็นต์ คำที่ได้เปอร์เซ็นต์ความถูกต้องมากที่สุดที่ 98 เปอร์เซ็นต์ คือคำว่า Good luck และ Line ส่วนคำที่ได้เปอร์เซ็นต์ความถูกต้องน้อยที่สุดที่ 70 เปอร์เซ็นต์ คือคำว่า Computer

สรุปคือค่าความแม่นยำที่ได้น้อยเกิดจากการที่ภาษามือบางคำมีท่าทางที่คล้ายกัน เมื่อนำโมเดลที่ผ่านการฝึกฝนมาใช้งานจริงทำให้ประสิทธิภาพลดลงกว่าตอนฝึกฝนโมเดล

5.2 ข้อเสนอแนะ

จากการใช้งานจริงโมเดลยังมีการทำนายที่ผิดพลาด สามารถปรับปรุงประสิทธิภาพให้ดียิ่งขึ้นโดยการเพิ่มจำนวนรูปภาพที่ใช้ฝึกฝน เพิ่มความหลากหลายของความเข้มแสงที่หลากหลาย การลองผิดลองถูกเลือกใช้ค่าพารามิเตอร์ต่าง ๆ จนกว่าจะพบค่าที่เหมาะสม เพื่อเพิ่มความแม่นยำในการทำนายมากขึ้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Mihir Garimella. “Sign Language Recognition.”
https://towardsdatascience.com/sign-language-recognition-with-advanced-computer-vision-7b74f20f3442

[2] Divya Sheel. “Deep Learning.”
https://new.abb.com/news/detail/58004

[3] Satsawat Natakarnkitkul. “convolutional neural network.”
https://towardsai.net/p/machine-learning/beginner-guides-to-convolutional-neural-network-from-scratch-kuzushiji-mnist-75f42c175b21

[4] Natthawat Phongchit. “convolutional neural network.”
https://medium.com/@natthawatphongchit

[5] Surapong Kanoktipsatharporn. “Convolutional Neural Network.”
https://www.bualabs.com/archives/2461

[6] Tashmit. “Convolution Layers.”
https://www.codingninjas.com/studio/library/convolution-layer-padding-stride-and-pooling-in-cnn

[7] Vishal Rajput. “Pooling layers.”
https://medium.com/aiguys/pooling-layers-in-neural-nets-and-their-variants-f6129fc4628b

[8] Avijeet Biswal. “Flatten layer.”
https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network

[9] ECoding. “Softmax Function.”
https://medium.com/super-ai-engineer/softmax-function-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-eae1f1bbef63.

[10] สารานุกรมเสรี. “ภาษามือ.”
https://th.wikipedia.org/wiki/%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2%E0%B8%A1%E0%B8%B7%E0%B8%AD

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[11] Ivan Grishchenko and Valentin Bazarevsky. “MediaPipe”
https://blog.research.google/2020/12/mediapipe-holistic-simultaneous-face.html

[12] Kasitin Y. “Django.”
https://medium.com/@lunaticabeexe/%E0%B8%97%E0%B8%B3%E0%B8%84%E0%B8%A7%E0%B8%B2%E0%B8%A1%E0%B8%A3%E0%B8%B9%E0%B9%89%E0%B8%88%E0%B8%B1%E0%B8%81%E0%B8%81%E0%B8%B1%E0%B8%9A-django-1b219d5ef06e

[13] KongRuksiam Studio “Django.”
https://kongruksiam.medium.com/%E0%B8%9B%E0%B8%B9%E0%B8%9E%E0%B8%B7%E0%B9%89%E0%B8%99%E0%B8%90%E0%B8%B2%E0%B8%99%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%9E%E0%B8%B1%E0%B8%92%E0%B8%99%E0%B8%B2webapplication%E0%B8%81%E0%B8%B1%E0%B8%9A%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2python%E0%B8%94%E0%B9%89%E0%B8%A7%E0%B8%A2-django-framework-9d3b7f48718a

[14] Withoutcoffee Icantbedev “Django.”
https://devhub.in.th/blog/django-model-and-orm.

[15] AWS. “AJAX”
https://aws.amazon.com/th/what-is/ajax/

[16] AWS. “javascript”
https://aws.amazon.com/th/what-is/javascript/

[17] จิตกร พิทักษ์เมธากุล “websocket”
https://www.jittagornp.me/blog/what-is-websocket/

[18] Medium “Mediapipe Holistic”
https://sertiscorp.medium.com/mediapipe-holistic/

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

โปรแกรมตรวจจับมือและบันทึกรูปภาพเก็บ dataset

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
import os
import cv2
import mediapipe as mp

DATA_DIR = 'D:\jectsun\data41-50'
if not os.path.exists(DATA_DIR):
    os.makedirs(DATA_DIR)

number_of_classes = 15
dataset_size = 250

cap = cv2.VideoCapture(0)  # Changed to use default camera (index 0)

mp_hands = mp.solutions.hands
hands = mp_hands.Hands(static_image_mode=False, max_num_hands=2,
min_detection_confidence=0.5)

for j in range(number_of_classes):
    if not os.path.exists(os.path.join(DATA_DIR, str(j))):
        os.makedirs(os.path.join(DATA_DIR, str(j)))

    print('Collecting data for class {}'.format(j))

    done = False
    while not done:
        ret, frame = cap.read()

        # Detect hands
        results = hands.process(cv2.cvtColor(frame, cv2.COLOR_BGR2RGB))

        # Draw hand landmarks (optional, for visualization)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        if results.multi_hand_landmarks:
            for hand_landmarks in results.multi_hand_landmarks:
                mp.solutions.drawing_utils.draw_landmarks(frame, hand_landmarks,
mp_hands.HAND_CONNECTIONS)

            # Get bounding box around hands
            x_min, y_min, x_max, y_max = float('inf'), float('inf'), float('-inf'), float('-inf')
            for hand_landmarks in results.multi_hand_landmarks:
                for point in hand_landmarks.landmark:
                    x, y = int(point.x * frame.shape[1]), int(point.y * frame.shape[0])
                    x_min = min(x_min, x)
                    y_min = min(y_min, y)
                    x_max = max(x_max, x)
                    y_max = max(y_max, y)

            # Increase the size of the bounding box
            x_min -= 25
            y_min -= 25
            x_max += 25
            y_max += 25

            # Ensure the bounding box stays within the frame boundaries
            x_min = max(x_min, 0)
            y_min = max(y_min, 0)
            x_max = min(x_max, frame.shape[1])
            y_max = min(y_max, frame.shape[0])

            # Draw a frame around the hands
            cv2.rectangle(frame, (x_min, y_min), (x_max, y_max), (0, 255, 0), 5)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        cv2.putText(frame, 'Ready? Press "Q" to Start Capture, "Esc" to Quit', (50, 50),
cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255, 0), 2, cv2.LINE_AA)
        cv2.imshow('frame', frame)

        key = cv2.waitKey(1)
        if key == ord('q'):
            done = True
        elif key == 27:  # 27 is the ASCII code for 'Esc'
            done = True
            break

    if done:  # Start capturing images
        counter = 0
        while counter < dataset_size:
            ret, frame = cap.read()

            # Detect hands
            results = hands.process(cv2.cvtColor(frame, cv2.COLOR_BGR2RGB))

            # Draw hand landmarks (optional, for visualization)
            if results.multi_hand_landmarks:
                for hand_landmarks in results.multi_hand_landmarks:
                    mp.solutions.drawing_utils.draw_landmarks(frame, hand_landmarks,
mp_hands.HAND_CONNECTIONS)

                # Get bounding box around hands
                x_min, y_min, x_max, y_max = float('inf'), float('inf'), float('-inf'), float('-inf')
                for hand_landmarks in results.multi_hand_landmarks:
                    for point in hand_landmarks.landmark:
                        x, y = int(point.x * frame.shape[1]), int(point.y * frame.shape[0])
                        x_min = min(x_min, x)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                y_min = min(y_min, y)
                x_max = max(x_max, x)
                y_max = max(y_max, y)

            # Increase the size of the bounding box
            x_min -= 25
            y_min -= 25
            x_max += 25
            y_max += 25

            # Ensure the bounding box stays within the frame boundaries
            x_min = max(x_min, 0)
            y_min = max(y_min, 0)
            x_max = min(x_max, frame.shape[1])
            y_max = min(y_max, frame.shape[0])

            # Draw a frame around the hands
            #cv2.rectangle(frame, (x_min, y_min), (x_max, y_max), (0, 255, 0), 5)

        cv2.imshow('frame', frame)

        key = cv2.waitKey(1)
        if key == 27:  # Press 'Esc' to stop capturing
            break

        cv2.imwrite(os.path.join(DATA_DIR, str(j), '{}.jpg'.format(counter)),
frame[y_min:y_max, x_min:x_max])
        counter += 1
cap.release()
cv2.destroyAllWindows()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

โปรแกรมแปลภาษามือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
import os
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'
import tensorflow as tf
import cv2
import mediapipe as mp
from keras.models import load_model
import numpy as np
import time
import pandas as pd

# Load the pre-trained model
model = load_model('D:/jectsun/model/CNN_1.h5')
capz = 0
mphands = mp.solutions.hands
hands = mphands.Hands(max_num_hands=2)  # Set max_num_hands to 2
mp_drawing = mp.solutions.drawing_utils
cap = cv2.VideoCapture(0)

_, frame = cap.read()

h, w, c = frame.shape

img_counter = 0
analysisframe = ''
letterpred = ['Care', 'Thank you', 'Sorry!']
last_prediction_time = time.time()
result_P = ""
confidence_P = ""
while True:
    _, frame = cap.read()
    k = cv2.waitKey(1)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    if k%256 == 27:
        # ESC pressed
        print("Escape hit, closing...")
        break
    current_time = time.time()
    if current_time - last_prediction_time >= 3:
        analysisframe = frame
        showframe = analysisframe
        cv2.imshow("Frame", showframe)
        framergbanalysis = cv2.cvtColor(analysisframe, cv2.COLOR_BGR2RGB)
        resultanalysis = hands.process(framergbanalysis)
        hand_landmarksanalysis = resultanalysis.multi_hand_landmarks
        if hand_landmarksanalysis and (len(hand_landmarksanalysis) == 1 or
len(hand_landmarksanalysis) == 2):
            x_max = 0
            y_max = 0
            x_min = w
            y_min = h
            for handLMsanalysis in hand_landmarksanalysis:
                for lmanalysis in handLMsanalysis.landmark:
                    x, y = int(lmanalysis.x * w), int(lmanalysis.y * h)
                    if x > x_max:
                        x_max = x
                    if x < x_min:
                        x_min = x
                    if y > y_max:
                        y_max = y
                    if y < y_min:
                        y_min = y
                    # Draw dots on hand landmarks
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                mp_drawing.draw_landmarks(analysisframe, handLMsanalysis,
mphands.HAND_CONNECTIONS)
            y_min -= 20
            y_max += 20
            x_min -= 20
            x_max += 20

            analysisframe = cv2.cvtColor(analysisframe, cv2.COLOR_BGR2RGB)
            analysisframe = analysisframe[y_min:y_max, x_min:x_max]
            cv2.imwrite('D:/jectsun/data/test/23.jpg', analysisframe)
            analysisframe = cv2.resize(analysisframe,(32,32))

            x_pre = np.array(analysisframe)
            x_pre = x_pre.astype('float32')
            x_pre = x_pre/255
            x_pre = np.reshape(x_pre ,(1,32,32,3))
            prediction = model.predict(x_pre)
            print(prediction)


            label = ['care', 'thank you', 'Sorry!']
            result = label[np.argmax(prediction)]
            confidence = round(np.max(prediction) * 100, 2)
            print("Predicted real ", result)
            print('Confidence : ', confidence)

            last_prediction_time = current_time
            result_P = result
            confidence_P = confidence

            # Draw result and confidence on the frame
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    cv2.putText(frame, f"Result: {result_P}", (10, 30), cv2.FONT_HERSHEY_SIMPLEX, 1, (0,
255, 0), 2, cv2.LINE_AA)
    cv2.putText(frame, f"Confidence: {confidence_P}%", (10, 70),
cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255, 0), 2, cv2.LINE_AA)
```

```
framergb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
result = hands.process(framergb)
hand_landmarks = result.multi_hand_landmarks
if hand_landmarks:
    for handLMs in hand_landmarks:
        x_max = 0
        y_max = 0
        x_min = w
        y_min = h
        for lm in handLMs.landmark:
            x, y = int(lm.x * w), int(lm.y * h)
            if x > x_max:
                x_max = x
            if x < x_min:
                x_min = x
            if y > y_max:
                y_max = y
            if y < y_min:
                y_min = y
        y_min -= 20
        y_max += 20
        x_min -= 20
        x_max += 20
        cv2.rectangle(frame, (x_min, y_min), (x_max, y_max), (0, 255, 0), 2)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            capz=capz+1

    cv2.imshow("Frame", frame)


cap.release()
cv2.destroyAllWindows()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

โปรแกรมสร้าง ฝึกสอนและทดสอบโมเดล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras import layers, optimizers, datasets
from matplotlib import pyplot as plt
import numpy as np
import cv2
import imutils
from imutils import paths
import os
import random
from os import listdir
from tqdm import tqdm
from os.path import isfile, join

width = 32
trainpath = 'train/'
testpath = 'test/'
trainImg = [trainpath+f for f in listdir(trainpath)]
testImg = [testpath+f for f in listdir(testpath)]
trainImg,testImg

def img2data(path):
    rawImgs = []
    labels = []
    h=0
    width = 32
    for imagePath in (path):
        img_train = list(paths.list_images(imagePath))
        for i in img_train:
            img= cv2.imread(img_train[h])
            img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            img = cv2.resize(img ,(width,width))
            rawImgs.append(img)
            h+=1
            l = imagePath.split('/')[1]
            if l == '0':
                labels.append([1,0,0])
            elif l == '1':
                labels.append([0,1,0])
            elif l == '2':
                labels.append([0,0,1])


        print(str(h)+"/"+str(len(img_train)))
        print(imagePath)
        h=0
    print("complete")
    return rawImgs, labels
x_train = []
y_train = []
x_train, y_train = img2data(trainImg)
print(len(x_train))
print(len(y_train))
x_test = []
y_test = []
x_test, y_test = img2data(testImg)
print(len(x_test))
print(len(y_test))
x_train = np.array(x_train)
y_train = np.array(y_train)
x_test = np.array(x_test)
y_test = np.array(y_test)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
x_train = x_train.astype('float32')
x_test = x_test.astype('float32')
x_train = x_train/255
x_test= x_test/255

print(x_train.shape)
print(y_train.shape)
print(x_test.shape)
print(y_test.shape)
x_train
model=keras.Sequential()
model.add(keras.layers.Conv2D(filters=32,kernel_size = 3,strides = (1,1),padding =
'same', activation='relu',input_shape = (width,width,3)))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2), padding = 'valid'))


model.add(keras.layers.Conv2D(filters=128,kernel_size = 3,strides = (1,1),padding =
'same', activation='relu'))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2), padding = 'valid'))


model.add(keras.layers.Conv2D(filters=256,kernel_size = 3,strides = (1,1),padding =
'same', activation='relu'))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2), padding = 'valid'))

model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Flatten())

model.add(keras.layers.Dense(units=128, activation='relu'))
model.add(keras.layers.Dropout(0.2))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
model.add(keras.layers.Dense(3,activation='softmax'))
model.summary()
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=0.0001),loss="categorical_crossentropy",metrics=['accuracy'])
history=model.fit(x=x_train,y=y_train,epochs=20,batch_size=32,validation_split=0.1)

plt.figure(figsize=(15,5))
plt.subplot(121),
plt.plot(history.history['accuracy'])
plt.plot(history.history['val_accuracy'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')

plt.subplot(122)
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')

test_loss,test_acc=model.evaluate(x=x_test,y=y_test)
print("Test Accuracy %.2f"%test_acc)

y_pred = model.predict(x_test)
#y_pred = [np.argmax(y_pred[i]) for i in range(len(y_pred))]
#y_test_class = [np.argmax(y_test[i]) for i in range(len(y_test))]
y_pred_class = np.argmax(y_pred,axis=1).tolist()
y_test_class = np.argmax(y_test,axis=1).tolist()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
from sklearn.metrics import confusion_matrix,f1_score
import pandas as pd
import seaborn as sn
cfArr = confusion_matrix(y_test_class, y_pred_class)
dfCm = pd.DataFrame(cfArr, index = ['0','1','2'], columns = ['0','1','2'])
plt.figure()
sn.heatmap(dfCm, annot=True, cmap="Blues")
plt.title('Confusion Matrix',fontsize = 15)
plt.show()

img_pre = list(paths.list_images('data/'))
img= cv2.imread(img_pre[1])
new_image_4 = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
plt.figure()
plt.subplot(122),plt.imshow(new_image_4)


img_re = cv2.resize(new_image_4 ,(32,32))
x_pre = np.array(img_re)
x_pre = x_pre.astype('float32')
x_pre = x_pre/255
x_pre = np.reshape(x_pre ,(1,32,32,3))
predict = model.predict(x_pre)
label = ['0','1','2']
result = label[np.argmax(predict)]

predict,result
model.save('./model/CNN_1.h5')
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ง

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Server (Consumers)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
import json
from channels.generic.websocket import WebsocketConsumer
import base64
import cv2
import numpy as np
from keras.models import load_model
import cv2,os
import numpy as np
import mediapipe as mp

os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'
current_time = 0
model = load_model('./CNN_M0-100.h5')
mphands = mp.solutions.hands
hands = mphands.Hands(max_num_hands=2)  # Set max_num_hands to 2
mp_drawing = mp.solutions.drawing_utils
result_P = "null"
confidence_P = ""
analysisframe = ''
b = os.path.dirname(os.path.abspath(__file__))

list1 =[]


class ChatConsumer(WebsocketConsumer):
    def connect(self):
        self.accept()

        self.send(text_data=json.dumps({
            'type' : 'connection_established',
            'massage':'You are now connected'
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        }))


    def receive(self, text_data):

            try:
                encoded_data = text_data.split(',')[1]
                data = base64.b64decode(encoded_data)
                np_arr = np.fromstring(data, np.uint8)
                img1 = cv2.imdecode(np_arr, cv2.IMREAD_COLOR)
                #print(img1)
                frame=img1


                #print(frame)
                h, w, c = frame.shape


                analysisframe = frame
                framergbanalysis = cv2.cvtColor(analysisframe, cv2.COLOR_BGR2RGB)
                resultanalysis = hands.process(framergbanalysis)
                hand_landmarksanalysis = resultanalysis.multi_hand_landmarks
                if hand_landmarksanalysis and (len(hand_landmarksanalysis) == 1 or
len(hand_landmarksanalysis) == 2):
                    x_max = 0
                    y_max = 0
                    x_min = w
                    y_min = h
                    for handLMsanalysis in hand_landmarksanalysis:
                        for lmanalysis in handLMsanalysis.landmark:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                    x, y = int(lmanalysis.x * w), int(lmanalysis.y * h)
                    if x > x_max:
                        x_max = x
                    if x < x_min:
                        x_min = x
                    if y > y_max:
                        y_max = y
                    if y < y_min:
                        y_min = y
                    # Draw dots on hand landmarks
                    mp_drawing.draw_landmarks(analysisframe,
handLMsanalysis, mphands.HAND_CONNECTIONS)
                y_min -= 20
                y_max += 20
                x_min -= 20
                x_max += 20

                analysisframe = cv2.cvtColor(analysisframe, cv2.COLOR_BGR2RGB)
                analysisframe = analysisframe[y_min:y_max, x_min:x_max]

                analysisframe = cv2.resize(analysisframe,(32,32))
                x_pre = np.array(analysisframe)
                x_pre = x_pre.astype('float32')
                x_pre = x_pre/255
                x_pre = np.reshape(x_pre ,(1,32,32,3))
                prediction = model.predict(x_pre)


                label = ['shower','Reading
documents','Dangerous','Anything','What','appropriate','search','deception','cold','Book','H
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
eavy','Tiger','Sorry','fun','Run','Swimming','drawing','already','Play','discount','Hurry','Full','W
aiting','Car','subway','electric train',
'nevermind','cat','Happy','Zebra','Look','expensive','Friend','poison','Meet','Practice','Rain','
bound','Butterfly','go','waste of
money','worried','indebted','Prevention','Sick','Save','nervous','insurance','Light','water','fre
sh milk','You','Business','Phone','television','Travel','How much','equal','Working','Cooking'
,'soldiers','gloves','copy','removing
headphones','Road','decision','Afternoon','Morning','fall','Clock
ticking','Walk','Children','Take care','Good luck','Slow','like','Help','collision','me','cash'
,'Who','plane','you','computer','alone','Understand','Thank
you','Sorry','mine','yours','his','Shuman','Kanom
Jeen','Glass','chair','considerate','fear','binoculars','TICTOK','Line']

                result = label[np.argmax(prediction)]
                confidence = round(np.max(prediction) * 100, 2)

                result_P = result
                confidence_P = confidence

                list1.append(result_P)
                framergb = cv2.cvtColor(frame, cv2.COLOR_BGR2RGB)
                result = hands.process(framergb)
                hand_landmarks = result.multi_hand_landmarks
                if hand_landmarks:
                    for handLMs in hand_landmarks:
                        x_max = 0
                        y_max = 0
                        x_min = w
                        y_min = h
                        for lm in handLMs.landmark:
                            x, y = int(lm.x * w), int(lm.y * h)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                    if x > x_max:
                        x_max = x
                    if x < x_min:
                        x_min = x
                    if y > y_max:
                        y_max = y
                    if y < y_min:
                        y_min = y
                y_min -= 20
                y_max += 20
                x_min -= 20
                x_max += 20
                cv2.rectangle(frame, (x_min, y_min), (x_max, y_max), (0, 255,
0), 2)



            with open(b +"\\templates\\sun.txt", 'w',encoding='utf-8') as file:
                file.write(result_P)

            _, buffer = cv2.imencode('.jpg', frame)
            uri = base64.b64encode(buffer).decode('utf-8')

            self.send(f'{uri}')
        except:
            try:
                self.send(text_data)
            except Exception as e:
                print(e)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก จ

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Server (View)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
from django.shortcuts import render,redirect
from django.http import HttpResponse
from databaseAI.models import Person,AImodel
from django.contrib import messages
from django.http.response import StreamingHttpResponse
import time
from .models import Todo
import os
import requests
# Create your views here.


def text(request):

    return render(request,"sun.txt")

def videolive(request):

    return render(request, 'chat/videolive.html')
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ฉ

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Server (URL)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
from django.urls import path,include
from databaseAI import views
#from databaseAI import sun



urlpatterns = [

    path('about',views.about),
    path('text',views.text),
    path('videolive',views.videolive),

]
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ช

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Client (Router)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
from django.urls import re_path
from . import consumers

websocket_urlpatterns = [
    re_path(r'ws/socket-server/', consumers.ChatConsumer.as_asgi())

]
```


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ซ

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Client (Base)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

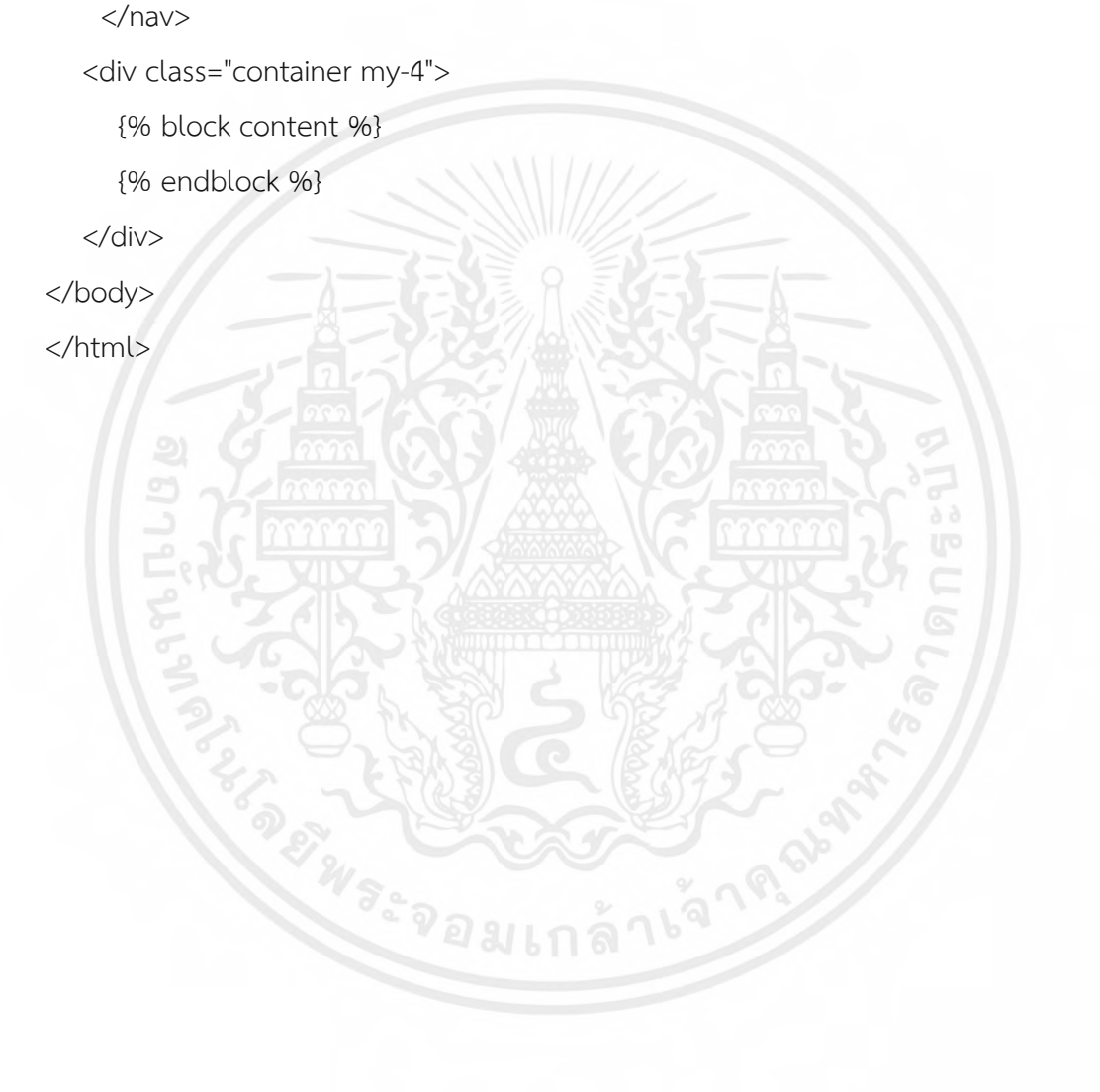
```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <link
href="https://cdn.jsdelivr.net/npm/bootstrap@5.0.2/dist/css/bootstrap.min.css"
rel="stylesheet" integrity="sha384-
EVSTQN3/azprG1Anm3QDgpJLIm9Nao0Yz1ztcQTwFspd3yD65VohhpuuCOmLASjC"
crossorigin="anonymous">
    {% block title %}
    {% endblock %}
</head>
<body>
    <!--<a href="/" class="btn btn-primary">หน้าแรกเว็บไซต์</a>
    <a href="/about"class="btn btn-danger">เกี่ยวกับฉัน</a>
    <a href="/form"class="btn btn-danger">ฟอร์มของฉัน</a>-->


    <nav class="navbar navbar-expand-lg navbar navbar-dark bg-dark" >
        <div class="container-fluid">
          <a class="navbar-brand" href="/"><S></S>TSL translation </a>
          <button class="navbar-toggler" type="button" data-bs-toggle="collapse" data-
bs-target="#navbarNav" aria-controls="navbarNav" aria-expanded="false" aria-
label="Toggle navigation">
            <span class="navbar-toggler-icon"></span>
          </button>
          <div class="collapse navbar-collapse" id="navbarNav">
            <ul class="navbar-nav">
              <li class="nav-item">
                <a class="nav-link active" aria-current="page"
href="/videolive">Translate</a>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                </li>

              </ul>
            </div>
          </div>
        </nav>
      <div class="container my-4">
          {% block content %}
          {% endblock %}
      </div>
</body>
</html>
```



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ฌ

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Client (VideoLive)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
{% extends "base.html" %}

{% block title %}

    <title>Translate</title>
    <style>

        button {
            all: unset;
            display: inline-block;
            padding: 15px;
            color: white;
            text-transform: uppercase;
            font-family: sans-serif;
            letter-spacing: 1.5px;
            font-weight: bold;
            cursor: pointer;
            border-radius: 17px;
            }

            /* Important stuff below */
            .button {
            background-color: #48484a;
            box-shadow: 0 5px 0 #000000;
            transition:.1s
            }

            .button:active {
            box-shadow: none;
            transform: translateY(5px);
            }
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
.button1 {
background-color: #4ac9e9;
box-shadow: 0 5px 0 #4daed8;
transition:.1s
}

.button1:active {
box-shadow: none;
transform: translateY(5px);
}

.button2 {
background-color: #6f7fe6;
box-shadow: 0 5px 0 #4742d7;
transition:.1s
}

.button2:active {
box-shadow: none;
transform: translateY(5px);
}

img {
display: block;
margin-left: auto;
margin-right: auto;
}
```

```
    body {
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            background:rgb(148, 186, 239);
            }
            .block{
                display: none;
                visibility: visible;
            }

    h3 { color: #ffffff; font-family: 'Raleway',sans-serif; font-size: 62px; font-weight: 800;
line-height: 72px; margin: 0 0 24px; text-align: center; text-transform: uppercase; }
    h4 { color: #000000e4; font-family: 'Raleway',sans-serif; font-size: 62px; font-weight:
800; line-height: 72px; margin: 0 0 24px; text-align: center; text-transform: uppercase; }
    </style>
{% endblock %}

{% block content %}



    <video class="block" autoplay="true"></video>
    <br />
    <img src="">

    <br />



    <button class="button" onclick="loadText()">Translation</button>
    <button class="button button1"  onclick="loadDoc()">Arrange words</button>
    <button class="button button2"  onclick="Deletetext()">Delete</button>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
<br />

<h3 id="textsign"></h3>
<h4 id="demo"></h4>



<script>
    // get video dom element
    const video = document.querySelector('video');

    // get img dom element
    const img = document.querySelector('img');


    // request access to webcam
    //navigator.mediaDevices.getUserMedia({video: {width: 426, height:
240}}).then((stream) => video.srcObject = stream);
    navigator.mediaDevices.getUserMedia({video: {width: 600, height: 300,video:
true}}).then((stream) => video.srcObject = stream);
    // returns a frame encoded in base64
    const getFrame = () => {
        const canvas = document.createElement('canvas');
        canvas.width = video.videoWidth;
        canvas.height = video.videoHeight;
        canvas.getContext('2d').drawImage(video, 0, 0);
        const data = canvas.toDataURL('image/png');
        return data;
    }
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
let url = `ws://${window.location.host}/ws/socket-server/`
//let url = 'wss://:_____.com/ws/endpoint/'
const FPS = 3;
const ws = new WebSocket(url);

ws.onopen = () => {
    console.log(`Connected to ${url}`);
    setInterval(() => {
        ws.send(getFrame());
    }, 1000 / FPS);
}

ws.onmessage = message => {
    // set the base64 string to the src tag of the image
    img.src = message.data;
}

function loadText() {
const xhttp = new XMLHttpRequest();
xhttp.onload = function() {
        if (this.readyState == 4 && this.status == 200) {
        document.getElementById("textsign").innerHTML = this.responseText;
    }
    };
let helloCounter = 1;
xhttp.open("GET", "text", true);
setInterval(() => {
    xhttp.open("GET", "text", true);
    xhttp.send()
    helloCounter++;},1000)
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    function loadDoc() {
        const xhttp = new XMLHttpRequest();
        xhttp.onload = function() {
        document.getElementById("demo").innerHTML += this.responseText;
        }
        xhttp.open("GET", "text");
        xhttp.send();
    }
    function Deletetext(){
        document.getElementById('demo').innerHTML = '';
    }



    </script>

{% endblock %}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ญ

โปรแกรมสร้างเว็บแปลภาษามือ ส่วน Client (About)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
{% extends "base.html" %}

{% block title %}
    <title>TSL translation</title>
    <style>
 @import
url('https://fonts.googleapis.com/css2?family=Oswald:wght@300&display=sw
ap');

body {
    background: #225cd8;
    padding: 0;
    margin: 0;

}

.top {
    position: relative;
    height: 10px;
    background: linear-gradient(141deg, cyan 0%, rgb(202, 182, 223) 40%,
rgb(87, 169, 232) 90%);
}

h4 {
font-family: 'Oswald', sans-serif;
color: white;
font-size: 5vw;
font-weight: 300;
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    text-transform: uppercase;
}

.text {

display: flex;
justify-content: center;
align-items: center;
height: 100vh;
}

.btn {
 position: relative;
 font-size: 17px;
 text-transform: uppercase;
 text-decoration: none;
 padding: 1em 2.5em;
 display: inline-block;
 border-radius: 6em;
 transition: all .2s;
 border: none;
 font-family: inherit;
 font-weight: 500;
 color: black;
 background-color: white;
}

.btn:hover {
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
 transform: translateY(-3px);
 box-shadow: 0 10px 20px rgba(0, 0, 0, 0.2);
}

.btn:active {
 transform: translateY(-1px);
 box-shadow: 0 5px 10px rgba(0, 0, 0, 0.2);
}

.btn::after {
 content: "";
 display: inline-block;
 height: 100%;
 width: 100%;
 border-radius: 100px;
 position: absolute;
 top: 0;
 left: 0;
 z-index: -1;
 transition: all .4s;
}

.btn::after {
 background-color: #fff;
}

.btn:hover::after {
 transform: scaleX(1.4) scaleY(1.6);
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
 opacity: 0;
}

    </style>
{% endblock %}

{% block content %}
<div class='top'></div>
<div class='text'>
    <h4>TSL translation</h4>
    <br />
    <a  class="btn" href="/videolive">Start Here</a>

</div>
{% endblock}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้